



Research article**A new Caputo–generalized conformable fractional derivative with applications to finite-time stability and secure neural systems****Abeer Aljohani**^{1,2,*}¹ Department of Computer Science and Information, Applied College, Taibah University, Madinah, Saudi Arabia² Energy, Industry, and Advanced Technologies Research Center, Taibah University, Madinah, Saudi Arabia* **Correspondence:** Email: aahjohani@taibahu.edu.sa.

Abstract: We introduce a Caputo–generalized conformable fractional (Caputo–GCF) derivative that separates the fractional order σ from an independently tunable structural memory parameter $\nu + \zeta$, and recovers the classical Caputo, Riemann–Liouville and (formally, in a limiting sense) Hadamard derivatives as special cases. Linearity, boundedness and continuity of the operator are established. Applying it to a delayed fractional-order neural network, we prove existence and uniqueness of the equilibrium point and derive two explicit finite-time stability criteria — a Hölder-based criterion for $0 < \sigma \leq \frac{1}{2}$ and a Cauchy–Schwarz-based criterion for $\frac{1}{2} < \sigma < 1$ — that explicitly quantify the roles of σ , $\nu + \zeta$, the delay, and the coupling matrices. Numerical experiments confirm the theoretical bounds and show faster empirical settling and smaller final error than the classical Caputo and GCF operators on a representative example. Applications to anomaly detection (Controlled Anomalies Time-Series dataset), delayed UAV attitude stabilization, and secure neural dynamics under bounded cyber disturbances further illustrate the framework’s applicability.

Keywords: fractional neural networks; Caputo–GCF operators; finite-time stability; delayed dynamical systems; memory-driven models; cyber–physical security

Mathematics Subject Classification: 26A33, 34A08, 34D20, 92B20, 93D05

1. Introduction

Fractional-order models provide an effective way to describe dynamical systems whose present behaviour depends on their past states. Many natural and engineered processes do not respond only to the current state; instead, their evolution is influenced by memory, hereditary effects, and nonlocal interactions. Fractional calculus offers a mathematical framework for representing these features more

accurately than classical integer-order models. For this reason, fractional differential systems have been widely used in neural networks, control theory, biological systems, signal processing, and cyber-physical models.

In recent years, fractional neural networks with time delays have attracted considerable attention because delays naturally occur in signal transmission, information processing, feedback control, and communication networks. In such systems, the combination of fractional memory, nonlinear activation functions, and delayed feedback can strongly affect stability and convergence behaviour. In particular, finite-time stability is an important concept because it guarantees that the system remains within a prescribed bound over a fixed time interval, which is especially useful in practical systems where fast and reliable recovery is required.

Although many finite-time stability results have been developed for fractional neural networks, most existing studies are based on fixed-kernel operators such as the Caputo, Hadamard, proportional Caputo, and rho-Caputo derivatives. These operators are useful, but they usually control the memory effect through one dominant fractional-order parameter. As a result, they do not explicitly separate the fractional order from an additional structural memory-scaling parameter. This limits their ability to tune long-memory effects and delayed nonlinear feedback independently. Therefore, a more flexible fractional operator is needed for delayed neural systems in which memory strength, delay effects, and coupling behaviour must be controlled in a more explicit and adjustable manner.

Motivated by this gap, we introduce a new Caputo–GCF derivative. The proposed derivative contains both the fractional order and an additional structural memory parameter, which allows greater flexibility in modelling delayed memory-driven dynamics. Methodologically, we apply a Caputo-type regularization to this two-parameter kernel, establish its operator-theoretic properties, use a contraction-mapping argument to obtain the network’s equilibrium point, and derive two complementary finite-time stability criteria via Hölder- and Cauchy–Schwarz-based Gronwall estimates; the precise sense in which this differs from existing single-parameter operators is detailed in the notational remark in the start of Section 4 and summarized in Tables 1 and 2 in Section 2. Using this new derivative, we study a class of delayed fractional-order neural networks and establish rigorous stability results. The main contributions of this paper are summarized as follows:

- A new Caputo–GCF derivative is introduced. In contrast to standard single-parameter fractional operators, the proposed formulation separates the fractional order σ from the structural memory-scaling parameters $\nu \in (0, 1]$ and $\zeta \in \mathbb{R}$ (independent subject only to $\nu + \zeta > 0$; see the remark following Definition 2), thereby allowing the nonlocal memory order and the kernel scaling to be adjusted independently. The classical Caputo, GCF, and Riemann–Liouville operators are recovered as special cases, while a connection with the Hadamard operator is obtained through the limiting case $\nu + \zeta \rightarrow 0^+$.
- The fundamental analytical properties of the proposed operator are established. In particular, the Caputo–GCF operator is proved to be linear, bounded, and continuous as a mapping from $C^1([a, b], \mathbb{R})$ into $C([a, b], \mathbb{R})$, and an explicit upper bound for its operator norm is obtained.
- The proposed derivative is incorporated into a delayed fractional-order neural network. Under appropriate Lipschitz and damping-dominance conditions, the existence and uniqueness of the equilibrium point are established by means of the contraction mapping principle.
- Two explicit sufficient finite-time stability criteria are derived for the two fractional-order regions $0 < \sigma \leq \frac{1}{2}$ and, $\frac{1}{2} < \sigma < 1$. The first criterion is obtained using Hölder’s inequality,

whereas the second is derived through the Cauchy–Schwarz inequality. The resulting bounds explicitly quantify the effects of the fractional order σ , structural memory parameter $\nu + \zeta$, delay length τ , initial perturbation δ , admissible bound ε , and instantaneous and delayed coupling quantities A_1 and A_2 .

- The theoretical criteria are verified through two numerical examples with $\delta = 0.31$ and $\varepsilon = 2$, for which all computed solution norms remain inside the prescribed stability region. In the direct operator comparison, the proposed Caputo–GCF model achieves a settling time of 4.92 and a final error of 0.0160, compared with settling times of 7.20 and 6.88 and final errors of 0.0427 and 0.0347 for the classical Caputo and GCF models, respectively.
- An ablation and reduced fractional-kernel study is conducted on the Controlled Anomalies Time-Series dataset. The proposed delayed Caputo–GCF configuration with $(\sigma, \nu + \zeta) = (0.50, 1.50)$ obtains an accuracy of 0.9872, recall of 0.9382, F1-score of 0.8867, and AUROC of 0.9895, giving the highest recall, F1-score, and AUROC among the considered reduced fractional-kernel models.
- The applicability of the theoretical framework is further demonstrated through delayed UAV attitude stabilization and secure neural dynamics under bounded cyber disturbances. The UAV example empirically illustrates finite-time stability (Definition 3) and the numerically observed settling behaviour discussed in Section 4.4 in the simultaneous presence of fractional memory, sensor delay, and bounded disturbances, while the cyber-security example shows that disturbance-induced state deviations remain bounded and return toward the equilibrium after the disturbance is removed.

These contributions distinguish the proposed work from existing fractional-neural-network studies by combining a two-parameter memory formulation with explicit finite-time stability bounds and quantitatively evaluated applications. The proposed framework therefore provides both greater memory-modelling flexibility and computable stability guarantees for delayed neural and cyber–physical systems.

Structure of the paper

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature on fractional-order neural networks, delayed systems, and finite-time stability, and identifies the research gap addressed by the proposed Caputo–GCF framework. Section 3 presents the preliminary material, including the generalized conformable fractional operators, the definition of finite-time stability, and the auxiliary inequalities required in the subsequent analysis.

Section 4 contains the main theoretical and numerical results. It introduces the new Caputo–generalized conformable fractional derivative, establishes its linearity, boundedness, and continuity, and discusses its relationships with classical fractional operators. The delayed fractional-order neural network model is then formulated, and the existence and uniqueness of its equilibrium point are established. Two explicit finite-time stability criteria are derived for the fractional-order regions. Subsection 4.1 presents two numerical examples that verify the derived stability criteria and illustrate the effects of the fractional order σ , the structural memory parameter $\nu + \zeta$, and the delay τ . It also provides a direct numerical comparison of the classical Caputo, GCF, and proposed Caputo–GCF operators. Subsection 4.2 reports an ablation and reduced

fractional-kernel comparison on the Controlled Anomalies Time-Series dataset, examining the individual effects of the fractional kernel, delayed term, fractional order, and structural memory parameter. Subsection 4.3 applies the proposed framework to anomaly detection in time-series neural networks. It explains how the dataset is incorporated into the delayed neural model, verifies the required stability assumptions, and examines the fractional-memory response under dataset-injected (controlled) anomaly disturbances. Subsection 4.4 considers the application of the proposed model to UAV attitude stabilization in the presence of sensor delay and bounded external disturbances. Subsection 4.5 investigates secure neural dynamics under bounded cyber disturbances and demonstrates the finite-time recovery of the neural states after the disturbance is removed. Finally, Section 5 summarizes the main theoretical, numerical, and application-based findings, discusses the limitations of the present work, and outlines directions for future research.

2. Literature review

Research on fractional neural dynamics has expanded rapidly as memory driven behaviour became essential for understanding modern systems. One stream of work focuses on finite time behaviour in fractional fuzzy neural structures, and a recent contribution in this direction can be found in [1]. Finite time properties in complex valued fractional networks with delay were explored in [2]. Another study examined neural networks with time varying delays and established new insights on their finite time stability (FTS) in [3]. Variable order fractional neural models were analysed in [4], reflecting the growing interest in operators that adapt to changing memory patterns.

Fractional systems of neutral type and their stability behaviour were studied in [5]. The behaviour of complex valued networks with generalized Caputo operators was investigated in [6]. Fractional neural networks under sampled data effects and delays were discussed in [7]. A different fractional operator, known as the rho Caputo operator, was considered in [8]. Stochastic fractional neural networks with impulses and delays were addressed in [9]. Mittag Leffler type stability for fractional neural structures was presented in [10].

Fractional fuzzy neural synchronization through nonlinear feedback control was discussed in [11]. Generalized proportional Caputo models and their Lyapunov based stability criteria were examined in [12]. Anomaly detection using evolving spiking neural networks was explored in [13]. Fractal time series based anomaly detection with LSTM autoencoders appeared in [14]. A broader survey on deep anomaly detection in multivariate time series was offered in [15]. Finite time behaviour in fuzzy cellular fractional networks with delay was studied in [16]. An adaptive approach to synchronization in delayed fractional fuzzy cellular systems was proposed in [17]. New finite time criteria for fractional bidirectional associative memory networks were given in [18]. FTS for fractional bidirectional memory networks with mixed delays was analysed in [19]. Hadamard fractional systems and their finite time behaviour were addressed in [20].

The above studies show that fractional neural networks have been investigated from several important directions, including Caputo-type delayed networks, generalized Caputo systems, rho-Caputo models, proportional Caputo frameworks, and Hadamard-type fractional systems. However, a closer comparison shows that most of these works mainly focus on deriving stability conditions for a given fractional operator, rather than on increasing the modelling freedom of the operator itself. In particular, many existing operators describe memory through one principal

fractional-order parameter, while the influence of structural memory scaling is either absent or only indirectly represented. Consequently, the effects of fractional order, delay, delayed coupling strength, and memory scaling are not always separated in an explicit way. This creates a limitation when delayed neural networks contain both nonlinear feedback and long-range memory effects, because the modeller has fewer independent parameters for adjusting the memory intensity and the delayed response of the system.

Beyond fractional neural networks specifically, related stability and oscillation frameworks for delay and impulsive systems have also been developed: oscillatory criteria for functional differential equations with non-canonical operators [21], Ulam-type stability for adjoint nonlinear impulsive neutral mixed integral equations [22], Ulam–Hyers stability and T -controllability for coupled integro fractional stochastic systems via integral contractors [23], and Ulam–Hyers–Rassias stability of Hilfer fractional stochastic impulsive systems [24]. These works illustrate the breadth of stability notions (oscillatory, Ulam-type, Ulam–Hyers–Rassias) developed for impulsive and stochastic fractional systems more broadly, complementing the finite-time stability framework developed here for delayed deterministic Caputo–GCF neural networks.

Finite-time stability has also been extensively studied for stochastic and jump-type systems beyond the fractional-order deterministic setting considered here. For mode-dependent time-delayed stochastic Markov jump systems, state-feedback and observer-based finite-time stabilizing controllers were designed in [25], using a mode-dependent approach to reduce conservatism in the delay bound. More recently, [26] introduced *annular* finite-time stability (a probability-based notion weaker than the usual mean-square formulation) and derived stabilization conditions for nonlinear stochastic Markov jump systems subject to impulsive effects. Some of the related studies conducted in [27–29]. These works illustrate how finite-time stability formulations extend to systems combining time delay, stochastic disturbance, and jump (or impulsive) dynamics; the present paper instead develops a finite-time stability theory for deterministic delayed neural networks governed by a new two-parameter fractional memory kernel, and a natural direction for future work (Section 5) is to combine the two settings. A concise comparison of representative fractional neural network studies with the present work is given in Table 1.

Table 1. Comparative overview of related fractional neural network studies.

Study	Operator	Delay	FTS	Tunable memory	Application validation
Hu et al. [2]	Caputo	Yes	Yes	No	No
Panda et al. [6]	Generalized Caputo	No/limited	Stability	Limited	Numerical only
Rhaima [8]	rho-Caputo	Yes	Yes	Limited	No
Proposed work	Caputo–GCF	Yes	Yes	Yes	Anomaly detection and cyber-security examples

Unlike prior works, the present study separates the fractional order σ from the structural parameter $\nu + \zeta$, allowing a more flexible representation of memory intensity and delayed feedback effects. This distinction is important because the fractional order controls the nonlocal memory behaviour, whereas the parameter $\nu + \zeta$ provides an additional scaling mechanism in the memory kernel. As a result, the

proposed Caputo–GCF framework can describe delayed neural dynamics with greater flexibility and can also produce explicit finite-time stability bounds in which the roles of the delay, memory parameter, and coupling matrices are clearly visible.

To make this comparison precise rather than qualitative, Table 2 places the Caputo–GCF operator against the Caputo, generalized-Caputo, and ρ -Caputo families already listed in Table 1 along six specific axes: kernel structure, number of independent memory parameters, treatment of the initial condition, whether a semigroup property $I^{\sigma_1} I^{\sigma_2} = I^{\sigma_1 + \sigma_2}$ -type composition is established, whether an inversion (fundamental) identity analogous to Lemma 3 is proved, and whether a finite-time stability theory and application validation accompany the operator.

Table 2. Structural comparison of the Caputo–GCF operator with existing Caputo-type generalized fractional operators.

Operator	Kernel	Indep. memory params.	IC treatment	Semi. group 12 shown	Inversion identity	FTS theory + application
Classical Caputo	$(t - s)^{-\sigma}$	1 (σ)	Caputo-type	Standard (classical)	Standard (classical)	Widely studied
Generalized Caputo (e.g., [28])	Problem-specific kernel, 1 shape parameter tied to σ	1	Caputo-type	Case-specific	Case-specific	Numerical only, no application study
ρ -Caputo (e.g., [29])	$(\rho(t) - \rho(s))^{-\sigma}$, ρ fixed a priori	1 (σ ; ρ is a chosen function, not fit jointly with σ)	Caputo-type	Standard	Standard	FTS shown, no application study
Proposed Caputo–GCF	$(t^{\nu+\zeta} - s^{\nu+\zeta})^{-\sigma}$	2 (σ , $\nu + \zeta$, independently tunable)	Caputo-type (Definition 4)	Not established beyond Definition 1/2 composition (open)	Proved (Lemma 3)	FTS (Theorems 2-3) + 3 applications

We state that a semigroup property for the Caputo–GCF integral operator (analogous to $I_{a+}^{\sigma_1, \nu+\zeta} I_{a+}^{\sigma_2, \nu+\zeta} = I_{a+}^{\sigma_1 + \sigma_2, \nu+\zeta}$) is not established in this paper. Only the fundamental (inversion) identity of Lemma 3, which is what the stability proofs in Section 4 actually require, is proved here.

Despite these valuable contributions, most existing approaches rely on fractional operators with fixed kernel structures that cannot fully capture the layered, evolving, and long range memory effects seen in modern neural and cyber physical systems. This limitation creates a clear need for a more flexible operator that can adapt to complex memory patterns while still supporting rigorous stability analysis.

The new operator introduced in this work responds to this need and provides a stronger foundation for the analysis of systems that were previously difficult to treat in a rigorous manner. The significance of this research is also practical. Memory based models appear in anomaly detection, autonomous decision systems, secure communication platforms, and biological networks. These applications demand reliable behavior under disturbance, delayed information, and persistent memory. The stability guarantees developed in this study bound the state deviation over a fixed time horizon (finite-time stability, Definition 3), which is a desirable property for systems that must maintain reliable performance under uncertainty. This makes the results valuable for real world

environments where safety, reliability, and resilience are essential.

3. Preliminaries

This section presents the fundamental definitions, notations, and auxiliary results required for the development of the main theoretical findings.

Definition 1 (GCF integral operators [30]). *Given $\sigma > 0$, $\nu \in (0, 1]$, $\zeta \in \mathbb{R}$, such that $\nu + \zeta > 0$ and an interval $[a, b] \subset \mathbb{R}$, where $0 < a < t < b$, the GCF integral of a function $u \in L^1([a, b])$ is defined by:*

$$I_{a^+}^{\sigma, \nu + \zeta} u(t) = \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} \int_a^t \frac{s^{\nu + \zeta - 1} u(s)}{(t^{\nu + \zeta} - s^{\nu + \zeta})^{1-\sigma}} ds,$$

where Γ is the gamma function.

Definition 2 (GCF derivative operators [30]). *Given $0 < \sigma < 1$, $\nu \in (0, 1]$, $\zeta \in \mathbb{R}$, such that $\nu + \zeta > 0$ and an interval $[a, b] \subset \mathbb{R}$, where $0 < a < t < b$, the GCF derivative is defined by:*

$$D_{a^+}^{\sigma, \nu + \zeta} u(t) = \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} t^{1-\nu-\zeta} \frac{d}{dt} \int_a^t \frac{s^{\nu + \zeta - 1} u(s)}{(t^{\nu + \zeta} - s^{\nu + \zeta})^\sigma} ds.$$

Throughout Definitions 1-2 and everywhere below, $\nu \in (0, 1]$ and $\zeta \in \mathbb{R}$ are two independent parameters, constrained jointly only by $\nu + \zeta > 0$; they are not a single parameter written in two parts. Both operators above depend on ν and ζ only through the sum $\nu + \zeta$, so for brevity every exponent in this paper is expressed in terms of $\nu + \zeta$ rather than in terms of ν and ζ separately, but this is purely a statement about where the sum happens to appear algebraically in these two kernels—it does not mean ν and ζ can be varied only jointly. In particular, the admissible parameter region $\{(\nu, \zeta) : \nu \in (0, 1], \zeta \in \mathbb{R}, \nu + \zeta > 0\}$ is a genuine two-dimensional set (for any fixed $\nu \in (0, 1]$, ζ ranges over the open half-line $\zeta > -\nu$), strictly larger than the one-dimensional set a single combined parameter $\lambda := \nu + \zeta > 0$ would represent.

Definition 3. [31] *Consider system (4.2) with an equilibrium point $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$. We say that x^* is finite-time stable relative to the set $\{t_0, \delta, \varepsilon, \Theta, \tau\}$, where $0 < \delta < \varepsilon$, $\delta, \varepsilon \in \mathbb{R}$, and $\Theta = [t_0, t_0 + t_f]$, if the following holds.*

Whenever the initial function ψ in (4.4) satisfies

$$\|\psi - x^*\| < \delta,$$

the corresponding solution $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ of system (4.2) obeys

$$\|x(t) - x^*\| < \varepsilon \quad \text{for all } t \in \Theta.$$

Here the distance between the initial data and the equilibrium is given by

$$\|\psi - x^*\| = \sup_{s \in [-\tau, 0]} \sum_{i=1}^n |\psi_i(s) - x_i^*|,$$

and the deviation of the state from x^ at time t is measured by*

$$\|x(t) - x^*\| = \sum_{i=1}^n |x_i(t) - x_i^*|.$$

Remark 1 (Terminological distinctions). *Three related but distinct properties are used in this paper and should not be conflated.*

- Finite-time stability (*Definition 3*) is a boundedness statement: *If the initial deviation is smaller than δ , the solution stays within ε of x^* throughout the fixed horizon Θ . It does not assert that the solution approaches x^* , nor that it enters any neighbourhood smaller than ε .*
- Finite-time convergence would mean $x(t) = x^*$ (or $x(t) \in B_r(x^*)$ for arbitrarily small r) is reached at some finite $t < \infty$; this is a stronger property that our coming results (Theorems 2 and 3) do not claim.
- Finite-time stabilization refers to a control-theoretic guarantee that a designed feedback drives the trajectory into an arbitrarily small neighbourhood of x^* within a settling time that can be bounded independently of the initial condition.

Throughout Sections 4.1, 4.4 and 4.5, the reported “settling time” is the first numerically observed instant at which $\|x(t) - x^*\|_2$ falls below a fixed threshold (e.g., 0.05); this is an empirical observation consistent with, but not implied by, the finite-time stability bounds of Theorems 2 and 3. We use the term “finite-time recovery” informally to describe this empirical behaviour and do not use it as a synonym for finite-time stability in the sense of Definition 3.

Remark 2 (Norm used in theory versus in the numerical figures). *Definition 3 and all theorems in this paper use the ℓ^1 (sum) norm $\|v\| = \sum_{i=1}^n |v_i|$. Several figures and tables in Section 4 instead plot or report the Euclidean norm $\|v\|_2 = (\sum_{i=1}^n v_i^2)^{1/2}$, which is more standard for visualizing trajectories. The two are related on $\mathbb{R}^n$ by the finite-dimensional norm equivalence*

$$\|v\|_2 \leq \|v\|_1 \leq \sqrt{n} \|v\|_2,$$

so a bound $\|x(t) - x^*\|_1 < \varepsilon$ established by Theorem 2 or Theorem 3 implies $\|x(t) - x^*\|_2 < \varepsilon$, and conversely a numerically observed $\|x(t) - x^*\|_2 < \varepsilon / \sqrt{n}$ implies $\|x(t) - x^*\|_1 < \varepsilon$. For $n = 2$ (Example 1) and $n = 3$ (Example 2, UAV example), this equivalence is loose enough ($\sqrt{n} \in \{\sqrt{2}, \sqrt{3}\}$) that we do not attempt to convert every reported $\|\cdot\|_2$ value into an $\|\cdot\|_1$ certificate; instead we flag explicitly, wherever a numerical threshold such as 0.05 or 0.02 is used, that it is an ℓ^2 -based empirical settling criterion (per the terminological remark above) and not a direct numerical evaluation of the ℓ^1 bound in Theorems 2 and 3.

Lemma 1. [32] *Let $a_u \geq 0$, $u \in \mathbb{N}$. Then for $z > 1$:*

$$\left(\sum_{e=1}^u a_e \right)^z \leq u^{z-1} \sum_{e=1}^u a_e^z.$$

Lemma 2 (Hölder and Cauchy–Schwarz inequalities). [33] *Let $\rho, \chi > 1$ satisfy the relation $\frac{1}{\rho} + \frac{1}{\chi} = 1$. Suppose $\eta \in L^\rho([a, b])$ and $\kappa \in L^\chi([a, b])$. Then the product $\eta(x)\kappa(x)$ belongs to $L^1([a, b])$, and one has*

$$\int_a^b |\eta(x)\kappa(x)| dx \leq \left(\int_a^b |\eta(x)|^\rho dx \right)^{1/\rho} \left(\int_a^b |\kappa(x)|^\chi dx \right)^{1/\chi}.$$

Here, $L^\rho([a, b])$ denotes the space of measurable functions $\phi : [a, b] \rightarrow \mathbb{R}$ for which $\int_a^b |\phi(x)|^\rho dx < \infty$. In the special situation where $\rho = \chi = 2$, the above estimate reduces to the classical Cauchy–Schwarz

inequality:

$$\int_a^b |\eta(x) \kappa(x)| dx \leq \left(\int_a^b |\eta(x)|^2 dx \right)^{1/2} \left(\int_a^b |\kappa(x)|^2 dx \right)^{1/2}.$$

Having established the necessary preliminaries, we can now move forward to the main results of this work.

4. Main results

Definitions 1 and 2 provide the generalized conformable integral and derivative. However, for stability analysis of neural systems, a Caputo-type formulation is preferable because it naturally incorporates classical initial conditions. Therefore, we introduce a Caputo–GCF derivative by applying the GCF derivative to $u(t) - u(a)$, which preserves compatibility with initial-value problems while retaining the memory-scaling flexibility of the GCF kernel.

We proceed to present the following new definition of Caputo–GCF operators as below: Notational remark: as in Definitions 1 and 2, $\nu \in (0, 1]$ and $\zeta \in \mathbb{R}$ are two genuinely *independent* parameters, subject only to the joint admissibility constraint $\nu + \zeta > 0$; they are not a single combined quantity artificially split into two symbols. ν and ζ enter the GCF kernel only through their sum in every formula from Definition 1 onward, so all exponents in this paper are written in terms of $\nu + \zeta$ for brevity, but this is a statement about where the sum $\nu + \zeta$ happens to appear algebraically, not a claim that ν and ζ are dependent or interchangeable with a single relabelled parameter. We therefore keep ν and ζ written out separately precisely to preserve this independence explicitly for the reader: $\nu \in (0, 1]$ and $\zeta \in \mathbb{R}$ can be varied independently subject to $\nu + \zeta > 0$, which is a strictly larger admissible parameter set than a single free parameter $\lambda > 0$ would represent, and collapsing them into one symbol would obscure this degree of freedom even though it does not appear explicitly in the exponents below.

Definition 4 (Caputo–GCF derivatives). *Given $0 < \sigma < 1$, $\nu \in (0, 1]$, $\zeta \in \mathbb{R}$ such that $\nu + \zeta > 0$ and an interval $[a, b] \subset \mathbb{R}$, where $0 < a < b$, the Caputo–GCF derivative is defined by:*

$${}^C D_{a^+}^{\sigma, \nu + \zeta} u(t) = D_{a^+}^{\sigma, \nu + \zeta} (u(t) - u(a)) = \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} t^{1 - \nu - \zeta} \frac{d}{dt} \int_a^t \frac{s^{\nu + \zeta - 1} [u(s) - u(a)]}{(t^{\nu + \zeta} - s^{\nu + \zeta})^\sigma} ds. \quad (4.1)$$

To clarify the scope of Definition 4, we record the following connections with some classical fractional operators.

Remark 3 (Relation with classical fractional operators). *The Caputo–GCF derivative in Definition 4 contains several well-known fractional operators as special or limiting cases.*

(1) If $\nu + \zeta = 1$, then

$${}^C D_{a^+}^{\sigma, 1} u(t) = \frac{1}{\Gamma(1 - \sigma)} \frac{d}{dt} \int_a^t \frac{u(s) - u(a)}{(t - s)^\sigma} ds,$$

which is the classical **Caputo** fractional derivative of order $0 < \sigma < 1$, written in the equivalent Riemann–Liouville form applied to $u(t) - u(a)$. Setting $\nu + \zeta = 1$ in Definition 2 gives $t^{1 - \nu - \zeta} = t^0 = 1$, so the outer differential operator reduces from $t^{1 - \nu - \zeta} \frac{d}{dt}$ to the plain $\frac{d}{dt}$, and the kernel $(t^{\nu + \zeta} - s^{\nu + \zeta})^{-\sigma}$ reduces to $(t - s)^{-\sigma}$. Substituting both into Definition 4 with u replaced by $u - u(a)$ gives exactly

$$\begin{aligned}
{}^C D_{a+}^{\sigma,1} u(t) &= D_{a+}^{\sigma,1} (u(t) - u(a)) \\
&= \frac{1}{\Gamma(1-\sigma)} \frac{d}{dt} \int_a^t \frac{u(s) - u(a)}{(t-s)^\sigma} ds = I_{a+}^{1-\sigma} [(u(\cdot) - u(a))]'(t) \\
&= {}^C D_{a+}^\sigma u(t),
\end{aligned}$$

the last equality holding because $\frac{d}{dt}(u(s) - u(a))|_{s=\text{integral}}$ integrates by parts (using $u(a) - u(a) = 0$ at the lower limit) to the standard Caputo representation $I_{a+}^{1-\sigma} u'(t)$, which is the definition of the classical Caputo derivative of order σ . This is the same integration-by-parts step used in Theorem 1 to pass from Definition 4 to the equivalent representation $\frac{(\nu+\zeta)^\sigma}{\Gamma(1-\sigma)} \int_a^t u'(s)(t^{\nu+\zeta} - s^{\nu+\zeta})^{-\sigma} ds$, specialized to $\nu + \zeta = 1$.

- (2) If the initial value satisfies $u(a) = 0$, then the proposed operator reduces to the corresponding GCF derivative, namely

$${}^C D_{a+}^{\sigma,\nu+\zeta} u(t) = D_{a+}^{\sigma,\nu+\zeta} u(t).$$

Thus, the present definition may be viewed as the **Caputo-type regularization** of the generalized conformable fractional derivative.

- (3) If $\nu + \zeta = 1$ and $u(a) = 0$, then the operator coincides with the classical **Riemann–Liouville** fractional derivative of order σ :

$${}^C D_{a+}^{\sigma,1} u(t) = D_{a+}^\sigma u(t).$$

This does not assert that Caputo and Riemann–Liouville derivatives coincide in general – they do not. Recall $D_{a+}^\sigma u(t) = \frac{d}{dt} I_{a+}^{1-\sigma} u(t)$ while ${}^C D_{a+}^\sigma u(t) = I_{a+}^{1-\sigma} u'(t)$; a direct computation (integrating $I_{a+}^{1-\sigma} u(t)$ by parts, or equivalently applying the Leibniz rule under the integral sign) gives the standard identity

$$D_{a+}^\sigma u(t) = {}^C D_{a+}^\sigma u(t) + \frac{u(a)}{\Gamma(1-\sigma)} (t-a)^{-\sigma},$$

so the two operators differ by the boundary term $\frac{u(a)}{\Gamma(1-\sigma)} (t-a)^{-\sigma}$, which vanishes precisely when $u(a) = 0$. Combining this with item 1 above (which gives ${}^C D_{a+}^{\sigma,1} u(t) = {}^C D_{a+}^\sigma u(t)$ for all $u \in X$, not only for $u(a) = 0$) yields ${}^C D_{a+}^{\sigma,1} u(t) = D_{a+}^\sigma u(t)$ exactly when $u(a) = 0$, as stated.

- (4) As $\nu + \zeta \rightarrow 0^+$, the kernel exponent satisfies

$$\frac{t^{\nu+\zeta} - s^{\nu+\zeta}}{\nu + \zeta} = \int_s^t r^{\nu+\zeta-1} dr \longrightarrow \log \frac{t}{s}, \quad t > s > 0,$$

so that, pointwise for fixed $t > s > 0$, $(t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} \sim (\nu + \zeta)^{\sigma-1} (\log(t/s))^{\sigma-1}$. However, the prefactor $(\nu + \zeta)^{1-\sigma} t^{1-\nu-\zeta}$ multiplying the integral in Definition 2 and the integration measure $s^{\nu+\zeta-1} ds$ also depend on $\nu + \zeta$, and a rigorous passage to the limit requires uniform control of the singular integral as $(\nu + \zeta)^{1-\sigma} \rightarrow \infty$ near $s = t$, which we do not establish in this paper. We therefore record the connection with the Hadamard derivative $D_{a+}^{\sigma,H} u(t) = \frac{1}{\Gamma(1-\sigma)} t \frac{d}{dt} \int_a^t (\log(t/s))^{-\sigma} \frac{u(s)}{s} ds$ only as a formal kernel-level analogy: The two kernels coincide in the pointwise limit above, but we do not assert

$$\lim_{\nu+\zeta \rightarrow 0^+} D_{a+}^{\sigma,\nu+\zeta} u(t) = D_{a+}^{\sigma,H} u(t)$$

as a pointwise or operator-norm statement without further regularity assumptions on u ; establishing such a limit rigorously is left for future work.

- (5) Therefore, the additional parameter $\nu + \zeta$ provides an independent structural memory-scaling mechanism. Unlike the classical Caputo derivative, where memory is mainly controlled by the fractional order σ , the Caputo–GCF derivative allows the fractional order and the memory-scaling structure to be adjusted separately.

Remark 4 (Novelty of the Caputo–GCF derivative). The construction $D_{a+}^{\sigma, \nu+\zeta}[u - u(a)]$ in Definition 4 is structurally analogous to the Caputo-type regularization used for other generalized fractional operators (e.g., proportional-Caputo, ρ -Caputo, Caputo–Hadamard). The novelty of the present operator does not lie in the regularization step itself, but in the underlying GCF kernel $(t^{\nu+\zeta} - s^{\nu+\zeta})^{-\sigma}$, in which the power-law exponents ν , ζ and the fractional order σ enter as independent parameters. In the operators listed above, the analogous kernel exponent is either fixed at unity (recovering the classical Caputo derivative), tied algebraically to σ , or replaced by a logarithmic kernel with no free scaling exponent (Hadamard-type). No single change of variable $t \mapsto t^{\nu+\zeta}$ reduces (4) to the classical Caputo derivative for $\nu + \zeta \neq 1$: Such a substitution would also rescale the outer differential operator $t^{1-\nu-\zeta} d/dt$, so the parameter $\nu + \zeta$ cannot be absorbed away by a relabelling of the independent variable alone. Equivalently, σ and $\nu + \zeta$ act on two different structural features of the kernel (its singularity order and its power-law scaling, respectively), which is what allows them to be tuned independently in Theorems 2 and 3. We do not claim that the Caputo–GCF derivative generates operators lying outside the closure of existing single-parameter families in every regime; the contribution is the explicit two-parameter kernel together with the finite-time stability theory developed on it in Section 4.

Theorem 1 (Linearity, boundedness and continuity of the Caputo–GCF operator). Let $0 < \sigma < 1$, $\nu + \zeta > 0$, and let $[a, b] \subset \mathbb{R}$ with $0 < a < b$. Define

$$X = C^1([a, b], \mathbb{R}), \quad Y = C([a, b], \mathbb{R}),$$

equipped with the norms

$$\|u\|_X = \|u\|_\infty + \|u'\|_\infty, \quad \|u\|_Y = \|u\|_\infty.$$

Then X and Y are Banach spaces, and the Caputo–GCF operator

$${}^C D_{a+}^{\sigma, \nu+\zeta} : X \longrightarrow Y$$

is linear, bounded and continuous. Moreover, for every $u \in X$, the operator admits the equivalent Caputo-type representation

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{u'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds, \quad t \in (a, b],$$

with

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(a) = 0.$$

In particular, there exists a positive constant $K_{\sigma, \nu+\zeta, a, b}$ such that

$$\left\| {}^C D_{a+}^{\sigma, \nu+\zeta} u \right\|_\infty \leq K_{\sigma, \nu+\zeta, a, b} \|u\|_X, \quad u \in X.$$

Proof. It is well known that $C([a, b], \mathbb{R})$, equipped with the supremum norm, is a Banach space. Also, $C^1([a, b], \mathbb{R})$, equipped with the norm

$$\|u\|_X = \|u\|_\infty + \|u'\|_\infty,$$

is a Banach space. Hence X and Y are Banach spaces.

Now, let $u \in C^1([a, b], \mathbb{R})$. From Definition 4, we have

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = D_{a+}^{\sigma, \nu+\zeta} (u(t) - u(a)).$$

Using the Caputo-type regularization and the change of variable associated with the generalized conformable kernel, this can be written as

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{u'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds, \quad t \in (a, b].$$

Since $u' \in C([a, b])$, the above singular integral is well defined for $0 < \sigma < 1$. Moreover, as $t \rightarrow a^+$, the integral tends to zero, and hence we may define

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(a) = 0.$$

Therefore,

$${}^C D_{a+}^{\sigma, \nu+\zeta} u \in C([a, b], \mathbb{R}),$$

which proves that the operator maps X into Y .

Next, for any $u, v \in X$ and constants $\lambda, \mu \in \mathbb{R}$, we obtain

$$\begin{aligned} {}^C D_{a+}^{\sigma, \nu+\zeta} (\lambda u + \mu v)(t) &= \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{(\lambda u + \mu v)'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds \\ &= \lambda \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{u'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds \\ &\quad + \mu \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{v'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds \\ &= \lambda {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) + \mu {}^C D_{a+}^{\sigma, \nu+\zeta} v(t). \end{aligned}$$

Thus, the Caputo–GCF operator is linear.

It remains to prove boundedness. Since $0 < a < b$ and $\nu + \zeta > 0$, the mean value theorem applied to $r \mapsto r^{\nu+\zeta}$ on $[s, t] \subset [a, b]$ gives some $\xi \in (s, t)$ with $t^{\nu+\zeta} - s^{\nu+\zeta} = (\nu + \zeta) \xi^{\nu+\zeta-1} (t - s)$. We now verify $m_{\nu+\zeta, a, b} := \min_{r \in [a, b]} r^{\nu+\zeta-1} > 0$ is positive and finite in each of the three cases:

- If $\nu + \zeta > 1$: $r^{\nu+\zeta-1}$ is increasing on $(0, \infty)$, so its minimum on $[a, b]$ is attained at $r = a$, giving $m_{\nu+\zeta, a, b} = a^{\nu+\zeta-1} > 0$ (finite since $a > 0$).
- If $\nu + \zeta = 1$: $r^{\nu+\zeta-1} = r^0 = 1$ identically, so $m_{\nu+\zeta, a, b} = 1 > 0$.
- If $0 < \nu + \zeta < 1$: $r^{\nu+\zeta-1}$ is decreasing on $(0, \infty)$, so its minimum on $[a, b]$ is attained at $r = b$, giving $m_{\nu+\zeta, a, b} = b^{\nu+\zeta-1} > 0$ (finite since $b < \infty$).

In every case $m_{\nu+\zeta,a,b} = \min(a^{\nu+\zeta-1}, b^{\nu+\zeta-1}) > 0$ is attained at an endpoint of $[a, b]$ (both endpoints coincide with the case $\nu + \zeta = 1$), and since $\xi \in (s, t) \subset [a, b]$ we have $\xi^{\nu+\zeta-1} \geq m_{\nu+\zeta,a,b}$ in all three cases, which is exactly the bound used below. Hence

$$t^{\nu+\zeta} - s^{\nu+\zeta} \geq (\nu + \zeta)m_{\nu+\zeta,a,b}(t - s), \quad a \leq s < t \leq b.$$

Therefore,

$$\begin{aligned} \left| {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) \right| &\leq \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \|u'\|_\infty \int_a^t \frac{1}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds \\ &\leq \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \|u'\|_\infty \int_a^t \frac{1}{((\nu + \zeta)m_{\nu+\zeta,a,b}(t - s))^\sigma} ds \\ &= \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)((\nu + \zeta)m_{\nu+\zeta,a,b})^\sigma} \|u'\|_\infty \int_a^t (t - s)^{-\sigma} ds \\ &= \frac{1}{\Gamma(1 - \sigma)m_{\nu+\zeta,a,b}^\sigma} \frac{(t - a)^{1-\sigma}}{1 - \sigma} \|u'\|_\infty. \end{aligned}$$

Since $t \leq b$, we get

$$\left| {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) \right| \leq \frac{(b - a)^{1-\sigma}}{(1 - \sigma)\Gamma(1 - \sigma)m_{\nu+\zeta,a,b}^\sigma} \|u'\|_\infty.$$

Hence,

$$\left\| {}^C D_{a+}^{\sigma, \nu+\zeta} u \right\|_\infty \leq K_{\sigma, \nu+\zeta, a, b} \|u'\|_\infty \leq K_{\sigma, \nu+\zeta, a, b} \|u\|_X,$$

where

$$K_{\sigma, \nu+\zeta, a, b} = \frac{(b - a)^{1-\sigma}}{(1 - \sigma)\Gamma(1 - \sigma)m_{\nu+\zeta,a,b}^\sigma}.$$

Thus, the Caputo–GCF operator is bounded from X into Y .

Finally, every bounded linear operator between normed spaces is continuous. Therefore,

$${}^C D_{a+}^{\sigma, \nu+\zeta} : C^1([a, b], \mathbb{R}) \longrightarrow C([a, b], \mathbb{R})$$

is continuous. This completes the proof. $\square$

Remark 5 (Verification of the constant $K_{\sigma, \nu+\zeta, a, b}$). The constant $K_{\sigma, \nu+\zeta, a, b} = \frac{(b - a)^{1-\sigma}}{(1 - \sigma)\Gamma(1 - \sigma)m_{\nu+\zeta,a,b}^\sigma}$ depends on σ , $\nu + \zeta$, a , and b only through $(b - a)^{1-\sigma}$, $\Gamma(1 - \sigma)$, and $m_{\nu+\zeta,a,b}^\sigma = \min(a, b)^{(\nu+\zeta-1)\sigma}$ (using the above case analysis). Two sanity checks confirm this is the expected scaling: (i) as $\sigma \rightarrow 1^-$, $(1 - \sigma)\Gamma(1 - \sigma) = \Gamma(2 - \sigma) \rightarrow 1$ remains finite and nonzero, while $(b - a)^{1-\sigma} \rightarrow 1$, so K stays bounded, consistent with the operator approaching an ordinary (non-singular) derivative; (ii) as $\sigma \rightarrow 0^+$, $(1 - \sigma)\Gamma(1 - \sigma) \rightarrow 1$ and $(b - a)^{1-\sigma} \rightarrow (b - a)$, so $K \rightarrow (b - a)$, consistent with $I_{a+}^{\sigma, \nu+\zeta}$ approaching the identity and the fractional derivative approaching a bounded (non-fractional) operator. We do not claim $K_{\sigma, \nu+\zeta, a, b}$ is a sharp (smallest possible) bound; it is the constant produced directly by the estimate above and is presented for existence purposes only.

Lemma 3 (Fundamental identity for the Caputo–GCF operator). Let $0 < \sigma < 1$, $\nu + \zeta > 0$, $[a, b] \subset \mathbb{R}$ with $0 < a < b$, and $u \in C^1([a, b], \mathbb{R})$. Then, for every $t \in (a, b]$,

$$I_{a+}^{\sigma, \nu+\zeta} {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = u(t) - u(a).$$

Proof. By Theorem 1, for $t \in (a, b]$,

$${}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = \frac{(\nu + \zeta)^\sigma}{\Gamma(1 - \sigma)} \int_a^t \frac{u'(s)}{(t^{\nu+\zeta} - s^{\nu+\zeta})^\sigma} ds.$$

Applying $I_{a+}^{\sigma, \nu+\zeta}$ and invoking Fubini's theorem (justified since $u' \in C([a, b])$ and $0 < \sigma < 1$ makes the iterated kernel absolutely integrable) gives

$$I_{a+}^{\sigma, \nu+\zeta} {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = \frac{(\nu + \zeta)^{1-\sigma} (\nu + \zeta)^\sigma}{\Gamma(\sigma) \Gamma(1 - \sigma)} \int_a^t u'(r) \left[\int_r^t \frac{s^{\nu+\zeta-1}}{(t^{\nu+\zeta} - s^{\nu+\zeta})^{1-\sigma} (s^{\nu+\zeta} - r^{\nu+\zeta})^\sigma} ds \right] dr.$$

With the substitution $w = \frac{s^{\nu+\zeta} - r^{\nu+\zeta}}{t^{\nu+\zeta} - r^{\nu+\zeta}} \in [0, 1]$, so that $dw = (\nu + \zeta) s^{\nu+\zeta-1} ds / (t^{\nu+\zeta} - r^{\nu+\zeta})$, the inner integral reduces to the Beta integral

$$\frac{1}{\nu + \zeta} \int_0^1 (1 - w)^{\sigma-1} w^{-\sigma} dw = \frac{1}{\nu + \zeta} B(1 - \sigma, \sigma) = \frac{\Gamma(\sigma) \Gamma(1 - \sigma)}{\nu + \zeta},$$

independently of r and t . Substituting back, the prefactor $(\nu + \zeta)^{1-\sigma} (\nu + \zeta)^\sigma / [\Gamma(\sigma) \Gamma(1 - \sigma)]$ cancels exactly against $\Gamma(\sigma) \Gamma(1 - \sigma) / (\nu + \zeta)$, and we get

$$I_{a+}^{\sigma, \nu+\zeta} {}^C D_{a+}^{\sigma, \nu+\zeta} u(t) = \int_a^t u'(r) dr = u(t) - u(a).$$

This completes the proof. $\square$

Throughout this study, we analyze a class of fractional-order neural networks incorporating time-delay effects. For each neuron $i = 1, 2, \dots, n$, the network dynamics are described by the system:

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x_i(t) = -c_i x_i(t) + \sum_{j=1}^n a_{ij} f_j(x_j(t)) + \sum_{j=1}^n b_{ij} f_j(x_j(t - \tau)) + l_i. \quad (4.2)$$

In vector form, the above model can be rewritten as:

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x(t) = -C x(t) + A f(x(t)) + B f(x(t - \tau)) + I, \quad (4.3)$$

for $t \in [t_0, t_f]$, where ${}^C D_{t_0}^{\sigma, \nu+\zeta}$ denotes the newly defined Caputo–GCF derivative (see Definition 4 in Section 4). The derivative is characterized by parameters $0 < \sigma < 1$ and $\nu + \zeta > 0$. Here, $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T \in \mathbb{R}^n$ represents the state vector, while $f(x(t)) = (f_1(x_1(t)), f_2(x_2(t)), \dots, f_n(x_n(t)))^T \in \mathbb{R}^n$ denotes the activation functions of neurons.

The constants c_i ($c_1, c_2, \dots, c_n$) correspond to the intrinsic decay or resetting rates of each neuron ($c_i > 0$), describing how quickly an isolated neuron returns to its resting potential in the absence of input. Matrices $A = (a_{ij})_{n \times n}$ and $B = (b_{ij})_{n \times n}$ characterize the synaptic couplings between neurons—at the present time t and at the delayed time $t - \tau$, respectively, with τ being a nonnegative constant delay. The constant vector $I = (l_1, l_2, \dots, l_n)^T$ denotes the external bias or input to the system.

The associated initial conditions are defined by

$$x_i(t_0 + s) = \psi_i(s), \quad (4.4)$$

where $\psi_i(s)$ are continuous functions specified on $[-\tau, 0]$. The supremum norm is given as

$$\|\psi\| = \sup_{s \in [-\tau, 0]} \sum_{i=1}^n |\psi_i(s)|.$$

Next, we list two conditions that will be required in the analysis below:

(H1) Each function f_i is assumed to satisfy a Lipschitz condition; that is, there exist constants $F_i > 0$ for which

$$\|f_i(x) - f_i(y)\| \leq F_i \|x - y\|$$

holds for all $x, y \in \mathbb{R}$ and for every index $i = 1, 2, \dots, n$. Practically, **(H1)** is the standard boundedness assumption for saturating neuron activations (e.g., tanh or sigmoidal firing-rate curves): It rules out unbounded sensitivity of a neuron's output to its input.

(H2) For the given coefficients a_{ij} , b_{ij} , c_i and the Lipschitz bounds F_j , the following inequality is required:

$$\sum_{i=1}^n \max_{1 \leq j \leq n} \{ |a_{ij}| F_j + |b_{ij}| F_j \} < \min_{1 \leq i \leq n} \{ c_i \}.$$

Assumption **H2** means that the intrinsic damping of each neuron is strong enough to dominate the combined instantaneous and delayed synaptic effects. This is reasonable in secure or controlled neural systems where decay or resetting mechanisms are intentionally designed to suppress delayed perturbations. More broadly, **(H1)** and **(H2)** delimit the scope of the present framework to *deterministic* delayed neural systems with globally Lipschitz activations, a constant delay, and coupling strengths that can be dominated by design; extending the analysis to the stochastic or jump-type dynamics discussed in Section 1 is left for future work.

Theorem 2. Suppose that conditions **(H1)** and **(H2)** are satisfied and that $0 < \sigma \leq \frac{1}{2}$. Assume further that the inequality

$$\begin{aligned} & \sqrt[q]{3^{q-1} + N A_2^q [(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}]} \\ & \times \exp \left(t_f^{\nu+\zeta} + N \left[\frac{A_1^q}{q} (t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + \frac{A_2^q}{q} ((t_f + \delta)^{\nu+\zeta} - (t_0 + \delta)^{\nu+\zeta}) \right] \right) < \frac{\varepsilon}{\delta} \end{aligned} \quad (4.5)$$

holds.

Here,

$$\begin{aligned} N &= \frac{3^{q-1}(\nu + \zeta)^{-q\sigma}}{(\Gamma(\sigma))^q} \left(\frac{\Gamma(p(\sigma - 1) + 1)}{p^{\sigma p - p + 1}} \right)^{\frac{q}{p}}, \quad p = 1 + \sigma, \quad q = \frac{\sigma + 1}{\sigma}, \\ A_1 &= \max_{1 \leq i \leq n} (c_i) + \sum_{i=1}^n \max_{1 \leq j \leq n} (|a_{ij}| F_j), \quad A_2 = \sum_{i=1}^n \max_{1 \leq j \leq n} (|b_{ij}| F_j). \end{aligned}$$

Under these conditions, the system (4.2) admits a single equilibrium point

$$x^* = (x_1^*, x_2^*, \dots, x_n^*)^T,$$

and this equilibrium is finite-time stable relative to the set

$$\{t_0, \delta, \varepsilon, \Theta, \tau\}, \quad 0 < \delta < \varepsilon, \quad \Theta = [t_0, t_0 + t_f].$$

Outline of the proof. The argument proceeds in three stages: (i) a contraction-mapping argument establishes existence and uniqueness of the equilibrium x^* ; (ii) system (4.2) is rewritten as an integral equation in $u(t) := x(t) - x^*$ via the fundamental identity of Lemma 3, and Hölder's inequality (with the conjugate exponents p, q verified in the remark below) together with Lemma 1 splits the resulting estimate into an initial-data term and two coupling terms (instantaneous, A_1 , and delayed, A_2); (iii) a case split on $t \leq t_0 + \tau$ isolates the delayed term's contribution, after which the Gronwall–Bellman lemma closes the estimate into the explicit bound stated in the theorem.

Proof. The first part of the analysis is devoted to showing that the system under consideration possesses a single equilibrium point. All vector norms below are the ℓ^1 (sum) norm $\|v\| = \sum_{i=1}^n |v_i|$, matching Definition 3. Following the equilibrium equation for system (4.2),

$$0 = -c_i x_i^* + \sum_{j=1}^n a_{ij} f_j(x_j^*) + \sum_{j=1}^n b_{ij} f_j(x_j^*) + l_i, \quad i = 1, \dots, n,$$

we work directly in the x -variables and define the mapping $F = (F_1, \dots, F_n)^T : \mathbb{R}^n \rightarrow \mathbb{R}^n$ by

$$F_i(x) = \frac{1}{c_i} \left[\sum_{j=1}^n a_{ij} f_j(x_j) + \sum_{j=1}^n b_{ij} f_j(x_j) + l_i \right], \quad i = 1, \dots, n,$$

so that x^* is an equilibrium of system (4.2) if and only if $F(x^*) = x^*$. For $x = (x_1, \dots, x_n)^T$ and $y = (y_1, \dots, y_n)^T$, assumption **(H1)** gives

$$\begin{aligned} \|F(x) - F(y)\| &= \sum_{i=1}^n \left| \frac{1}{c_i} \sum_{j=1}^n (a_{ij} + b_{ij})(f_j(x_j) - f_j(y_j)) \right| \\ &\leq \sum_{i=1}^n \frac{1}{c_i} \sum_{j=1}^n (|a_{ij}| + |b_{ij}|) F_j |x_j - y_j| \\ &\leq \sum_{i=1}^n \frac{1}{c_i} \left(\max_{1 \leq j \leq n} \{|a_{ij}| F_j + |b_{ij}| F_j\} \right) \sum_{j=1}^n |x_j - y_j| \\ &\leq \frac{\sum_{i=1}^n \max_{1 \leq j \leq n} \{|a_{ij}| F_j + |b_{ij}| F_j\}}{\min_{1 \leq i \leq n} \{c_i\}} \|x - y\| =: L \|x - y\|. \end{aligned}$$

By assumption **(H2)**, $\sum_{i=1}^n \max_j \{|a_{ij}| F_j + |b_{ij}| F_j\} < \min_i \{c_i\}$, so $L < 1$: This is *exactly* the contraction constant required by **(H2)**, with no further conditions needed. Hence $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a contraction with respect to $\|\cdot\|$ on the Banach space $(\mathbb{R}^n, \|\cdot\|)$, and by the Banach fixed-point theorem F admits a unique fixed point $x^* \in \mathbb{R}^n$ with $F(x^*) = x^*$, i.e.,

$$-c_i x_i^* + \sum_{j=1}^n a_{ij} f_j(x_j^*) + \sum_{j=1}^n b_{ij} f_j(x_j^*) + l_i = 0, \quad i = 1, \dots, n.$$

Thus, system (4.2) has a unique equilibrium point x^* .

Now, the goal is to check the FTS of $x^* = (x_1^*, x_2^*, \dots, x_n^*)^T$. Define $x(t) = (x_1(t), x_2(t), \dots, x_n(t))^T$ as a solution of system (4.2). We have

$$\begin{aligned}
{}^C D_{t_0^+}^{\sigma, \nu+\zeta}(x_i(t) - x_i^*) &= -c_i(x_i(t) - x_i^*) \\
&+ \sum_{j=1}^n a_{ij} (f_j(x_j(t)) - f_j(x_j^*)) + \sum_{j=1}^n b_{ij} (f_j(x_j(t-\tau)) - f_j(x_j^*)), \quad t \geq t_0. \quad (4.6)
\end{aligned}$$

Applying $I_{t_0, \cdot}^{\sigma, \nu+\zeta}$ to both sides of (4.6) and using Lemma 3 (with $a = t_0$), for $t \geq t_0$, we get the integral equation of (4.6) as follows:

$$x_i(t) - x_i^* = x_i(t_0) - x_i^* + I_{t_0, t}^{\sigma, \nu+\zeta} g_i(s), \quad t \geq t_0, \quad (4.7)$$

where

$$g_i(s) = -c_i(x_i(s) - x_i^*) + \sum_{j=1}^n a_{ij} (f_j(x_j(s)) - f_j(x_j^*)) + \sum_{j=1}^n b_{ij} (f_j(x_j(s-\tau)) - f_j(x_j^*)).$$

Then, setting $u(t) := x(t) - x^*$ and $\varphi(s) := \psi(s) - x^*$ for $s \in [-\tau, 0]$ (so that $\|\varphi\| = \|\psi - x^*\|$), we have

$$\begin{aligned}
\|u(t)\| &= \sum_{i=1}^n |x_i(t_0) - x_i^* + I_{t_0, t}^{\sigma, \nu+\zeta} g_i(s)| \\
&\leq \sum_{i=1}^n |x_i(t_0) - x_i^*| + \left| \sum_{i=1}^n I_{t_0, t}^{\sigma, \nu+\zeta} g_i(s) \right| \\
&\leq \|\varphi\| + \left| \sum_{i=1}^n I_{t_0, t}^{\sigma, \nu+\zeta} g_i(s) \right|, \quad t \geq t_0. \quad (4.8)
\end{aligned}$$

It follows that,

$$\begin{aligned}
\left| \sum_{i=1}^n I_{t_0, t}^{\sigma, \nu+\zeta} g_i(s) \right| &\leq \sum_{i=1}^n c_i I_{t_0, t}^{\sigma, \nu+\zeta} |u_i(s)| \\
&+ \sum_{i=1}^n \sum_{j=1}^n |a_{ij}| F_j I_{t_0, t}^{\sigma, \nu+\zeta} |u_j(s)| \\
&+ \sum_{i=1}^n \sum_{j=1}^n |b_{ij}| F_j I_{t_0, t}^{\sigma, \nu+\zeta} |u_j(s-\tau)|. \quad (4.9)
\end{aligned}$$

So,

$$\begin{aligned}
\|u(t)\| &\leq \|\varphi\| + \max_{1 \leq i \leq n} (c_i) I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s)\| + \sum_{i=1}^n \max_{1 \leq j \leq n} (|a_{ij}| F_j) I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s)\| \\
&+ \sum_{i=1}^n \max_{1 \leq j \leq n} (|b_{ij}| F_j) I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s-\tau)\|, \quad t \geq t_0, \\
&\leq \|\varphi\| + A_1 I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s)\| + A_2 I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s-\tau)\|, \quad t \geq t_0, \quad (4.10)
\end{aligned}$$

where $A_1 = \max_{1 \leq i \leq n}(c_i) + \sum_{i=1}^n \max_{1 \leq j \leq n}(|a_{ij}|F_j)$, $A_2 = \sum_{i=1}^n \max_{1 \leq j \leq n}(|b_{ij}|F_j)$. Using the Hölder inequality, one has:

$$\begin{aligned} \|u(t)\| &\leq \|\varphi\| + A_1 \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} \int_{t_0}^t (t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} s^{\nu+\zeta-1} \|u(s)\| ds \\ &\quad + A_2 \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} \int_{t_0}^t (t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} s^{\nu+\zeta-1} \|u(s-\tau)\| ds \\ &\leq \|\varphi\| + \frac{(\nu + \zeta)^{-\sigma}}{\Gamma(\sigma)} \left(\int_{t_0}^t (\nu + \zeta) e^{ps^{\nu+\zeta}} s^{\nu+\zeta-1} (t^{\nu+\zeta} - s^{\nu+\zeta})^{p\sigma-p} ds \right)^{\frac{1}{p}} \\ &\quad \times \left[A_1 \left((\nu + \zeta) \int_{t_0}^t e^{-qs^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s)\|^q ds \right)^{\frac{1}{q}} \right. \\ &\quad \left. + A_2 \left((\nu + \zeta) \int_{t_0}^t e^{-qs^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s-\tau)\|^q ds \right)^{\frac{1}{q}} \right]. \end{aligned}$$

Using the change of variable $\mu = p(t^{\nu+\zeta} - s^{\nu+\zeta})$, we get for $t \geq t_0$:

$$\int_{t_0}^t (\nu + \zeta) e^{ps^{\nu+\zeta}} s^{\nu+\zeta-1} (t^{\nu+\zeta} - s^{\nu+\zeta})^{p\sigma-p} ds \leq \frac{e^{pt^{\nu+\zeta}}}{p^{\sigma p - p + 1}} \Gamma(p(\sigma - 1) + 1). \quad (4.11)$$

It follows from Lemma 1 (for $n = 3$ and $l = q$) that:

$$\begin{aligned} \left(e^{-(\nu+\zeta)t^{\nu+\zeta}} \|u(t)\| \right)^q &\leq 3^{q-1} \|\varphi\|^q + N \left[A_1^q \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|u(s)\|^q ds \right. \\ &\quad \left. + A_2^q \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|u(s-\tau)\|^q ds \right]. \end{aligned}$$

Case 1: Let $t \in [t_0, t_0 + \tau]$: on this interval $s - \tau \in [t_0 - \tau, t_0]$, so $u(s - \tau) = \varphi(s - \tau)$ is given data.

We have

$$\begin{aligned} \left(e^{-t^{\nu+\zeta}} \|u(t)\| \right)^q &\leq 3^{q-1} \|\varphi\|^q + N \left[A_2^q \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|\varphi(s-\tau)\|^q ds \right. \\ &\quad \left. + A_1^q \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|u(s)\|^q ds \right] \\ &\leq \|\varphi\|^q \left(3^{q-1} + NA_2^q \left((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta} \right) \right) \\ &\quad + NA_1^q \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|u(s)\|^q ds. \end{aligned} \quad (4.12)$$

Case 2: Let $t > t_0 + \tau$: We have

$$\begin{aligned}
& \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-qs^{\nu+\zeta}} \|u(s-\tau)\|^q ds \\
&= \int_{t_0-\tau}^{t-\tau} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} e^{-q(w+\tau)^{\nu+\zeta}} \|u(w)\|^q dw \\
&= \int_{t_0-\tau}^{t_0} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} e^{-q(w+\tau)^{\nu+\zeta}} \|\varphi(w)\|^q dw \\
&\quad + \int_{t_0}^{t-\tau} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} e^{-q(w+\tau)^{\nu+\zeta}} \|u(w)\|^q dw \\
&\leq \|\varphi\|^q \int_{t_0-\tau}^{t_0} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} dw + \int_{t_0}^t (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} \left(e^{-w^{\nu+\zeta}} \|u(w)\| \right)^q dw \\
&\leq \|\varphi\|^q \left((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta} \right) + \int_{t_0}^t (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} \left(e^{-w^{\nu+\zeta}} \|u(w)\| \right)^q dw. \tag{4.13}
\end{aligned}$$

(Here $(w + \tau)^{\nu+\zeta-1}$ is bounded on the finite interval $[t_0 - \tau, t_f]$ regardless of whether $\nu + \zeta$ is less than, equal to, or greater than 1, since $w + \tau \geq t_0 > 0$ throughout; the case $\nu + \zeta < 1$, where $r^{\nu+\zeta-1}$ is decreasing rather than increasing in r , does not affect the validity of the displayed bound, only the location of its maximum on the interval.) So,

$$\begin{aligned}
\left(e^{-t^{\nu+\zeta}} \|u(t)\| \right)^q &\leq \|\varphi\|^q \left(3^{q-1} + NA_2^q ((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) \\
&\quad + N \left[\int_{t_0}^t \left((\nu + \zeta) A_1^q s^{\nu+\zeta-1} + (\nu + \zeta) A_2^q (s + \tau)^{\nu+\zeta-1} \right) \left(e^{-qs^{\nu+\zeta}} \|u(s)\| \right)^q ds \right].
\end{aligned}$$

The Gronwall–Bellman lemma (see, e.g., standard references on integral inequalities such as [32]) states that if a nonnegative function $y(t)$ satisfies $y(t) \leq \alpha + \int_{t_0}^t \beta(s)y(s) ds$ with $\alpha \geq 0$ and $\beta \geq 0$ integrable, then $y(t) \leq \alpha \exp(\int_{t_0}^t \beta(s) ds)$. Applying this with $y(s) = (e^{-s^{\nu+\zeta}} \|u(s)\|)^q$, $\alpha = \|\varphi\|^q (3^{q-1} + NA_2^q [(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}])$, and $\beta(s) = N(\nu + \zeta)(A_1^q s^{\nu+\zeta-1} + A_2^q (s + \tau)^{\nu+\zeta-1})$ (each of which is nonnegative and integrable on $[t_0, t_f]$) gives, on $[t_0, t_f]$:

$$\begin{aligned}
\left(e^{-t^{\nu+\zeta}} \|u(t)\| \right)^q &\leq \|\varphi\|^q \left(3^{q-1} + NA_2^q ((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) \\
&\quad \times \exp \left[N \left(A_1^q (t^{\nu+\zeta} - t_0^{\nu+\zeta}) + A_2^q ((t + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta}) \right) \right].
\end{aligned}$$

Hence, $\forall t \in [t_0, t_f]$,

$$\begin{aligned}
\|u(t)\| &\leq \|\varphi\| \sqrt[q]{3^{q-1} + NA_2^q ((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta})} \\
&\quad \times \exp \left[\frac{A_1^q}{q} (t^{\nu+\zeta} - t_0^{\nu+\zeta}) + \frac{A_2^q}{q} ((t + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta}) \right].
\end{aligned}$$

Since $u(t) = x(t) - x^*$ and $\varphi = \psi - x^*$ by construction, this reads

$$\|x(t) - x^*\| \leq \|\psi - x^*\| \sqrt[q]{3^{q-1} + NA_2^q[(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}]} \\ \times \exp\left[t^{\nu+\zeta} + N\left(\frac{A_1^q}{q}(t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + \frac{A_2^q}{q}((t + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta})\right)\right],$$

so, if (4.5) is satisfied and $\|\psi - x^*\| < \delta$, then $\|x(t) - x^*\| < \varepsilon$ for all $t \in [t_0, t_f]$, i.e., system (4.2) is finite-time stable w.r.t $\{t_0, \delta, \varepsilon, \Theta, \tau\}$, according to Definition 3. $\square$

Remark 6 (Integrability of the Hölder exponent and the limit $\sigma \rightarrow 0^+$). With $p = 1 + \sigma$ as used in (4.11) above, the exponent on the singular factor $(t^{\nu+\zeta} - s^{\nu+\zeta})^{p(\sigma-1)}$ satisfies

$$p(\sigma - 1) = (1 + \sigma)(\sigma - 1) = \sigma^2 - 1.$$

Local integrability of $(t - s)^{p(\sigma-1)}$ near $s = t$ requires $p(\sigma - 1) > -1$, i.e., $\sigma^2 - 1 > -1$, i.e., $\sigma^2 > 0$, which holds for every $\sigma \in (0, \frac{1}{2}]$; this justifies the application of Hölder's inequality in (4.11) throughout the stated range. The margin of integrability, $\sigma^2 - 1 - (-1) = \sigma^2$, vanishes as $\sigma \rightarrow 0^+$: The singularity becomes only marginally integrable, and the constant N above – which contains the factor $\Gamma(p(\sigma - 1) + 1)^p = \Gamma(\sigma^2)^p$ – diverges in this limit. Consequently, while condition (4.5) remains mathematically valid for every $\sigma \in (0, \frac{1}{2}]$, it becomes increasingly conservative (and numerically uninformative) as $\sigma \rightarrow 0^+$, consistent with the degeneracy of the fractional-order model itself in that limit. The choice $p = 1 + \sigma$, $q = (\sigma + 1)/\sigma$ is not claimed to be optimal; it is the simplest conjugate pair satisfying $1/p + 1/q = 1$ for which $p(\sigma - 1) > -1$ holds on the entire interval $(0, \frac{1}{2}]$, including the endpoint $\sigma = \frac{1}{2}$.

Interpretation of Theorem 2. The stability condition in Theorem 2 gives an explicit relation between the initial perturbation, the admissible tolerance, the delay, the memory parameters, and the coupling matrices. In particular, inequality (4.5) shows that finite-time stability becomes easier to satisfy when the initial deviation δ is small and the admissible bound ε is large. It also shows that weak delayed coupling, represented by a small value of A_2 , helps to reduce the contribution of the delayed term. Similarly, a shorter delay τ decreases the quantity

$$(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta},$$

and therefore reduces the amplification caused by the memory-delay interaction. Moreover, the decay coefficients c_i play a stabilizing role when they dominate the instantaneous and delayed coupling terms appearing in assumption **(H2)**. Finally, the structural memory parameter $\nu + \zeta$ affects the transient growth through the kernel-dependent constant N and the powers $t^{\nu+\zeta}$. Hence, a suitable choice of $\nu + \zeta$ can reduce transient amplification and improve the finite-time stability margin.

The following direct consequences of Theorem 2 show how the proposed criterion recovers important special cases.

Corollary 1 (No-delay case). Let the assumptions of Theorem 2 hold with $0 < \sigma \leq \frac{1}{2}$. If $\tau = 0$, then $f_j(x_j(t - \tau)) = f_j(x_j(t))$, so the delayed neural network (4.2) reduces to the non-delayed Caputo–GCF neural system

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x(t) = -Cx(t) + Af(x(t)) + Bf(x(t)) + I = -Cx(t) + (A + B)f(x(t)) + I,$$

in which A and B act as a single combined instantaneous coupling matrix $A + B$; there is no longer a delayed channel to bound separately. Equation (4.10) becomes, with only the single combined term,

$$\|u(t)\| \leq \|\varphi\| + A'_1 I_{t_0, t}^{\sigma, \nu+\zeta} \|u(s)\|, \quad A'_1 := \max_{1 \leq i \leq n} (c_i) + \sum_{i=1}^n \max_{1 \leq j \leq n} (|a_{ij}|F_j + |b_{ij}|F_j),$$

and the same Hölder estimate and change of variable as in the proof of Theorem 2 give, via Lemma 1 with $n = 2$ and the Gronwall–Bellman lemma,

$$\|x(t) - x^*\| \leq \|\psi - x^*\| \sqrt[q]{2^{q-1}} \exp \left[t_f^{\nu+\zeta} + \frac{N(A'_1)^q}{q} (t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) \right], \quad t \in [t_0, t_f].$$

Consequently, the no-delay finite-time stability condition is

$$\sqrt[q]{2^{q-1}} \exp \left(t_f^{\nu+\zeta} + \frac{N(A'_1)^q}{q} (t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) < \frac{\varepsilon}{\delta},$$

$$A'_1 = \max_{1 \leq i \leq n} (c_i) + \sum_{i=1}^n \max_{1 \leq j \leq n} (|a_{ij}|F_j + |b_{ij}|F_j).$$

Therefore, the equilibrium point x^* is finite-time stable with respect to $\{t_0, \delta, \varepsilon, \Theta, 0\}$.

Remark 7 (Conservativeness of the bound with respect to c_i). Condition (H2) correctly requires the decay rates c_i to dominate the instantaneous and delayed coupling strengths, and in this sense larger c_i is stabilizing for the true dynamics of system (4.2). However, the sufficient condition (4.5) is obtained by bounding the linear damping term $-c_i(x_i(t) - x_i^*)$ through the triangle inequality, $|c_i(x_i(t) - x_i^*)| \leq c_i |x_i(t) - x_i^*|$, which discards its stabilizing sign. As a result, c_i enters the constant $A_1 = \max_i(c_i) + \sum_i \max_j(|a_{ij}|F_j)$ only through its magnitude, and increasing $\max_i(c_i)$ can increase A_1 and thereby make the sufficient condition (4.5) numerically harder to satisfy, even though it does not make the underlying system less stable. This is a known source of conservativeness in Gronwall-type finite-time stability estimates: the bound reflects worst-case growth under condition (H2) rather than the beneficial cancellation that a larger c_i actually provides. A tighter estimate exploiting the sign of $-c_i(\cdot)$ (e.g., via an integrating-factor argument keyed to each c_i individually, as opposed to the uniform bound $\max_i c_i$) would remove this artifact and is left for future work.

Corollary 2 (Classical Caputo case). Let the assumptions of Theorem 2 be satisfied and assume that $\nu + \zeta = 1$. Then the Caputo–GCF derivative becomes the classical Caputo fractional derivative of order σ , and system (4.2) reduces to the Caputo fractional-order delayed neural network

$${}^C D_{t_0}^\sigma x(t) = -Cx(t) + Af(x(t)) + Bf(x(t - \tau)) + I.$$

With $\nu + \zeta = 1$, the term $(t_f + \delta)^{\nu+\zeta} - (t_0 + \delta)^{\nu+\zeta}$ in (4.5) simplifies exactly to $t_f - t_0$ (the δ -shift cancels linearly), and the coefficient multiplying it is A_2^q/q , not a second copy of A_1^q/q . In this case, the finite-time stability criterion (4.5) takes the form

$$\sqrt[q]{3^{q-1} + NA_2^q \tau} \exp \left(t_f + \frac{N(A_1^q + A_2^q)}{q} (t_f - t_0) \right) < \frac{\varepsilon}{\delta}.$$

Consequently, the unique equilibrium point x^* of the classical Caputo fractional-order delayed neural network is finite-time stable.

Corollary 3 (No delayed coupling). *Let the assumptions of Theorem 2 hold with $0 < \sigma \leq \frac{1}{2}$. If $B = 0$, then $A_2 = 0$, and system (4.2) becomes*

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x(t) = -Cx(t) + Af(x(t)) + I.$$

Every term in inequality (4.5) that is multiplied by A_2 vanishes when $A_2 = 0$ — this removes both the $NA_2^q[(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}]$ term under the outer root and the $\frac{A_2^q}{q}((t_f + \delta)^{\nu+\zeta} - (t_0 + \delta)^{\nu+\zeta})$ term in the exponent (it does not become a second copy of the A_1 term). In this case, the stability criterion (4.5) reduces to

$$\sqrt[q]{3^{q-1}} \exp\left(t_f^{\nu+\zeta} + \frac{NA_1^q}{q}(t_f^{\nu+\zeta} - t_0^{\nu+\zeta})\right) < \frac{\varepsilon}{\delta}.$$

Thus, in the absence of delayed coupling, finite-time stability depends only on the instantaneous coupling matrix A , the decay matrix C , the fractional order σ , and the structural memory parameter $\nu + \zeta$.

The case $\alpha > 1/2$ requires a separate treatment because the kernel exponent changes the integrability properties of the fractional convolution term. Therefore, Cauchy–Schwarz inequality is more suitable than the Hölder-based estimate used for $0 < \alpha \leq 1/2$. We proceed as below.

Theorem 3. *Assume that σ lies in the interval $(\frac{1}{2}, 1)$ and that the fractional system (4.2) is initialized by $x(t_0 + s) = \varphi(s)$ for $-\tau \leq s < 0$. If it holds that*

$$\begin{aligned} & \sqrt{3 + 2MA_2^2((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta})} \\ & \times \exp\left(t_f^{\nu+\zeta} + M\left(A_1^2(t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + A_2^2((t_f + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta})\right)\right) < \frac{\varepsilon}{\delta}, \end{aligned} \quad (4.14)$$

where $M = \frac{3(\nu + \zeta)^{-2\sigma}\Gamma(2\sigma - 1)}{4^\sigma(\Gamma(\sigma))^2}$, then (4.2) is finite-time stable with respect to $\{t_0, \delta, \varepsilon, \Theta, \tau\}$, $\delta < \varepsilon$.

Outline of the proof. The structure mirrors that of Theorem 2, with the Cauchy–Schwarz inequality (the $p = q = 2$ case) in place of the Hölder step, and the constant M (instead of N) absorbing the resulting Gamma-function factor; the same case split on $t \leq t_0 + \tau$ and the same Gronwall–Bellman closing step are used.

Proof. As in the proof of Theorem 2, set $u(t) := x(t) - x^*$ and $\varphi(s) := \psi(s) - x^*$; we have the following estimation:

$$\|u(t)\| \leq \|\varphi\| + \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} \int_{t_0}^t (t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} s^{\nu+\zeta-1} (A_1\|u(s)\| + A_2\|u(s - \tau)\|) ds.$$

By using the Cauchy–Schwarz inequality, one has

$$\begin{aligned}
\|u(t)\| &\leq \|\varphi\| + \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} A_1 \left(\int_{t_0}^t (t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} s^{\nu+\zeta-1} \|u(s)\| ds \right) \\
&\quad + \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} A_2 \left(\int_{t_0}^t (t^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1} s^{\nu+\zeta-1} \|u(s-\tau)\| ds \right) \\
&\leq \|\varphi\| + \frac{(\nu + \zeta)^{1-\sigma}}{\Gamma(\sigma)} \left(\int_{t_0}^t e^{2s^{\nu+\zeta}} s^{\nu+\zeta-1} (t^{\nu+\zeta} - s^{\nu+\zeta})^{2\sigma-2} ds \right)^{\frac{1}{2}} \\
&\quad \times \left[A_1 \left((\nu + \zeta) \int_{t_0}^t e^{-2s^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s)\|^2 ds \right)^{\frac{1}{2}} \right. \\
&\quad \left. + A_2 \left((\nu + \zeta) \int_{t_0}^t e^{-2s^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s-\tau)\|^2 ds \right)^{\frac{1}{2}} \right].
\end{aligned}$$

By using the change of variable $\mu = t^{\nu+\zeta} - s^{\nu+\zeta}$, we get:

$$\int_{t_0}^t (\nu + \zeta) e^{2s^{\nu+\zeta}} s^{\nu+\zeta-1} (t^{\nu+\zeta} - s^{\nu+\zeta})^{2\sigma-2} ds \leq 2 \frac{\Gamma(2\sigma-1)}{4^\sigma} e^{2t^{\nu+\zeta}}, \quad t \geq t_0.$$

It follows from Lemma 1 (for $n = 3$ and $l = 2$) that:

$$\begin{aligned}
\|u(t)\|^2 &\leq 3\|\varphi\|^2 + 2Me^{2t^{\nu+\zeta}} \left[A_1^2 \int_{t_0}^t (\nu + \zeta) e^{-2s^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s)\|^2 ds \right. \\
&\quad \left. + A_2^2 \int_{t_0}^t (\nu + \zeta) e^{-2s^{\nu+\zeta}} s^{\nu+\zeta-1} \|u(s-\tau)\|^2 ds \right].
\end{aligned}$$

Hence,

$$\begin{aligned}
\left(e^{-(\nu+\zeta)t^{\nu+\zeta}} \|u(t)\| \right)^2 &\leq 3\|\varphi\|^2 + 2M \left[A_1^2 \int_{t_0}^t (\nu + \zeta) e^{-2(\nu+\zeta)s} s^{\nu+\zeta-1} \|u(s)\|^2 ds \right. \\
&\quad \left. + A_2^2 \int_{t_0}^t (\nu + \zeta) e^{-2(\nu+\zeta)s} s^{\nu+\zeta-1} \|u(s-\tau)\|^2 ds \right].
\end{aligned}$$

In the following, there are two cases as $t \in [t_0, t_0 + \tau]$ or $t \in [t_0 + \tau, t_f]$.

Case I: $t \in [t_0, t_0 + \tau]$.

On this interval $s - \tau \in [t_0 - \tau, t_0]$, so $u(s - \tau) = \varphi(s - \tau)$ is given data. We have:

$$\begin{aligned}
\left(e^{-(\nu+\zeta)t^{\nu+\zeta}} \|u(t)\| \right)^2 &\leq 3\|\varphi\|^2 + 2M \left[A_2^2 \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-2(\nu+\zeta)s^{\nu+\zeta}} \|\varphi(s-\tau)\|^2 ds \right. \\
&\quad \left. + A_1^2 \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-2(\nu+\zeta)s^{\nu+\zeta}} \|u(s)\|^2 ds \right] \\
&\leq \|\varphi\|^2 \left(3 + 2MA_2^2((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) \\
&\quad + 2MA_1^2 \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-2(\nu+\zeta)s^{\nu+\zeta}} \|u(s)\|^2 ds.
\end{aligned}$$

Case 2: $t > t_0 + \tau$.

We have:

$$\begin{aligned}
 & \int_{t_0}^t (\nu + \zeta) s^{\nu+\zeta-1} e^{-2(\nu+\zeta)s^{\nu+\zeta}} \|u(s-\tau)\|^2 ds \\
 &= \int_{t_0}^{t-\tau} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} e^{-2(\nu+\zeta)(w+\tau)^{\nu+\zeta}} \|u(w)\|^2 dw \\
 &\leq \int_{t_0}^{t-\tau} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} e^{-2(\nu+\zeta)w^{\nu+\zeta}} \|u(w)\|^2 dw \\
 &\leq \|\varphi\|^2 \int_{t_0-\tau}^{t_0} (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} dw + \int_{t_0}^t (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} \left(e^{-w^{\nu+\zeta}} \|u(w)\| \right)^2 dw \\
 &\leq \|\varphi\|^2 ((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) + \int_{t_0}^t (\nu + \zeta)(w + \tau)^{\nu+\zeta-1} \left(e^{-w^{\nu+\zeta}} \|u(w)\| \right)^2 dw.
 \end{aligned}$$

(As in Theorem 2, $(w + \tau)^{\nu+\zeta-1}$ is bounded on the finite interval $[t_0 - \tau, t_f]$ since $w + \tau \geq t_0 > 0$; this bound holds regardless of whether $\nu + \zeta$ is less than, equal to, or greater than 1, so the monotonicity of $r^{\nu+\zeta-1}$ does not affect the validity of the inequality above.) So,

$$\begin{aligned}
 \left(e^{-(\nu+\zeta)t^{\nu+\zeta}} \|u(t)\| \right)^2 &\leq \|\varphi\|^2 \left(3 + 2MA_2^2((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) \\
 &\quad + 2M \left[\int_{t_0}^t ((\nu + \zeta)A_1^2 s^{\nu+\zeta-1} + (\nu + \zeta)A_2^2 (s + \tau)^{\nu+\zeta-1}) \left(e^{-2(\nu+\zeta)s^{\nu+\zeta}} \|u(s)\|^2 \right) ds \right].
 \end{aligned}$$

Applying the Gronwall–Bellman lemma (as stated explicitly in the proof of Theorem 2) with $y(s) = (e^{-(\nu+\zeta)s^{\nu+\zeta}} \|u(s)\|)^2$ on $[t_0, t_f]$ gives:

$$\begin{aligned}
 \left(e^{-(\nu+\zeta)t^{\nu+\zeta}} \|u(t)\| \right)^2 &\leq \|\varphi\|^2 \left(3 + 2MA_2^2((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}) \right) \\
 &\quad \times \exp \left[2M \left(A_1^2(t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + A_2^2((t_f + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta}) \right) \right].
 \end{aligned}$$

Hence, $\forall t \in [t_0, t_f]$

$$\begin{aligned}
 \|u(t)\| &\leq \|\varphi\| \sqrt{3 + 2MA_2^2((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta})} \\
 &\quad \times \exp \left[\frac{1}{2} \left((\nu + \zeta)t^{\nu+\zeta} + 2M \left(A_1^2(t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + A_2^2((t_f + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta}) \right) \right) \right]. \quad (4.15)
 \end{aligned}$$

Since $u(t) = x(t) - x^*$ and $\varphi = \psi - x^*$, this reads $\|x(t) - x^*\| \leq \|\psi - x^*\| \sqrt{\cdots} \exp[\cdots]$ with the same right-hand side as above. So, if (4.14) is satisfied and $\|\psi - x^*\| < \delta$, then $\|x(t) - x^*\| < \varepsilon$, $\forall t \in [t_0, t_f]$ i.e., system (4.2) is finite-time stable w.r.t $\{t_0, \delta, \varepsilon, \Theta, \tau\}$. $\square$

To highlight the difference between the two finite-time stability results, we summarize the main features of Theorems 2 and 3 in Table 3.

Table 3. Comparison between Theorems 2 and 3.

Aspect	Theorem 2	Theorem 3
Fractional order	$0 < \sigma \leq \frac{1}{2}$	$\frac{1}{2} < \sigma < 1$
Main inequality	Hölder inequality	Cauchy–Schwarz inequality
Constant	N	M
Memory behaviour	Weaker long-memory regime	Stronger long-memory regime

Remark 8 (On the exclusion of $\sigma = \frac{1}{2}$ from Theorem 3). *The constant M in Theorem 3 contains $\Gamma(2\sigma - 1)$, which requires $2\sigma - 1 > 0$, i.e., $\sigma > \frac{1}{2}$; at $\sigma = \frac{1}{2}$, $\Gamma(2\sigma - 1) = \Gamma(0)$ is singular, so the Cauchy–Schwarz-based estimate underlying Theorem 3 is not applicable there. By convention we therefore assign $\sigma = \frac{1}{2}$ to Theorem 2 (interval $0 < \sigma \leq \frac{1}{2}$), which uses the Hölder-based estimate valid on a closed-at- $\frac{1}{2}$ interval. The two theorems thus partition $(0, 1)$ into $(0, \frac{1}{2}] \cup (\frac{1}{2}, 1)$ with no overlap and no gap. As $\sigma \rightarrow \frac{1}{2}^+$ within Theorem 3, $\Gamma(2\sigma - 1) \rightarrow \infty$ and hence $M \rightarrow \infty$, so the bound in (4.14) degenerates precisely at the boundary where Theorem 2 takes over; the two criteria are therefore complementary rather than smoothly overlapping, and no claim of continuity of the stability bound across $\sigma = \frac{1}{2}$ is made.*

Remark 9 (Why the split at $\sigma = \frac{1}{2}$ is a technical, not fundamental, boundary). *The choice of conjugate exponent pair is what forces the split. In the proof of Theorem 2, Hölder’s inequality is applied with $p = 1 + \sigma$, $q = (\sigma + 1)/\sigma$, chosen (see the integrability remark above) so that the singular kernel $(t^{\nu+\zeta} - s^{\nu+\zeta})^{p(\sigma-1)}$ stays integrable, $p(\sigma - 1) = \sigma^2 - 1 > -1$, on the entire interval $(0, \frac{1}{2}]$. The same pair does not force integrability to fail for $\sigma > \frac{1}{2}$; rather, Theorem 3 instead specializes to $p = q = 2$ (plain Cauchy–Schwarz), which produces the closed-form constant M involving $\Gamma(2\sigma - 1)$ – convenient algebraically, but singular at $\sigma = \frac{1}{2}$ for reasons intrinsic to that particular choice of exponents, not to the operator or the underlying dynamics. In principle a single σ -dependent conjugate pair $p(\sigma), q(\sigma)$ satisfying $1/p + 1/q = 1$ and $p(\sigma - 1) > -1$ could be chosen continuously across all of $(0, 1)$, which would likely yield one unified estimate valid on the whole interval instead of two complementary ones; we have not pursued this unification, since it does not change the qualitative finite-time stability conclusion and would substantially lengthen the proof. We therefore regard the split as a technical artifact of the two specific inequalities used (Hölder with $p = 1 + \sigma$ versus plain Cauchy–Schwarz), not evidence of a genuine qualitative change in the dynamics at $\sigma = \frac{1}{2}$.*

4.1. Numerical examples

In this section, two expository examples with their numerical results will be given in order to clarify the validity of the theoretical results which we obtained in the previous sections.

Numerical implementation. The GCF integral $I_{t_0,t}^{\sigma,\nu+\zeta}$ appearing in (4.7) is approximated by a product-trapezoidal (piecewise-linear) quadrature on a uniform grid $t_k = t_0 + k \Delta t$, $k = 0, \dots, K$, with step size $\Delta t = 10^{-3}$, using the closed-form weights for the singular kernel $(t_k^{\nu+\zeta} - s^{\nu+\zeta})^{\sigma-1}$ obtained by integrating the interpolating linear polynomial exactly against the kernel on each subinterval. The delayed term $x(t - \tau)$ is evaluated using linear interpolation of the stored history on $[t_0 - \tau, t_0]$ when $t - \tau$ falls between grid points. All simulations were run in Python (NumPy/SciPy) with double precision; the settling times reported in this section are computed as the first grid point t_k at which the discrete error norm

falls below the stated threshold and does not subsequently exceed it. For the operator comparison reported later in this subsection specifically, the classical Caputo, GCF, and proposed Caputo–GCF simulations share: the same initial history ψ , the same matrices C, A, B, I and activation function, the same grid and step size Δt , the same equilibrium-point computation (the fixed-point iteration of the contraction F derived in the proof preceding Theorem 2, run to a fixed tolerance), the same simulation interval $[t_0, t_f]$, and the same treatment of the delayed term (linear interpolation as above); only the fractional kernel (and, for GCF/Caputo–GCF, the parameter $\nu + \zeta$) differs between the three runs, so that this comparison isolates the effect of the operator itself. A mesh-refinement check (halving Δt) was not performed for this submission and is noted as an open item in the Limitations paragraph of Section 5.

The parameter grid used in Tables 4–5 ($\sigma \in \{0.5, 0.9\}$, $\nu + \zeta \in \{0.5, 1.0, 1.5, 2.0\}$) was chosen so that one value of σ falls in each of the two stability regimes covered by Theorems 2 and 3 ($\sigma = 0.5$ at the boundary $\sigma \leq \frac{1}{2}$, and $\sigma = 0.9$ well inside $\sigma > \frac{1}{2}$), while $\nu + \zeta$ is varied over a fourfold range to expose weak, moderate, and strong structural-memory scaling within each regime. The same representative pair $(\sigma, \nu + \zeta) = (0.5, 1.5)$ is then reused as the “proposed configuration” in the operator comparison (Section 4.1), the ablation study (Section 4.2), and the UAV example (Section 4.4) so that results across sections remain comparable rather than being tuned independently per application.

Table 4. FTS results for Example 1 at $\delta = 0.31$ (t_f and $\|x(t)\|$ are empirically simulated values; see the remark following Definition 3).

σ	$\nu + \zeta$	t_f	$\ x(t)\ $
0.5	0.5	0.2843	1.4328
	1.0	0.6125	1.2874
	1.5	0.7819	1.1640
	2.0	0.9052	1.0457
0.9	0.5	0.9481	1.9823
	1.0	1.1374	1.5526
	1.5	1.2047	1.2315
	2.0	1.2639	1.0184

Table 5. FTS results for Example 2 at $\delta = 0.31$ (t_f and $\|x(t)\|$ are empirically simulated values; see the remark following Definition 3).

σ	$\nu + \zeta$	t_f	$\ x(t)\ $
0.5	0.5	0.0247	0.7441
	1.0	0.1429	0.6123
	1.5	0.2946	0.5718
	2.0	0.4527	0.5249
0.9	0.5	0.1314	0.9832
	1.0	0.8643	1.3041
	1.5	1.0258	1.1095
	2.0	1.1472	0.8921

Example 1. Let us consider the Caputo–GCF order neural networks system with a constant time delay $\tau > 0$:

$${}^C D_{t_0}^{\sigma, \nu + \zeta} x_i(t) = -c_i x_i(t) + \sum_{j=1}^2 a_{ij} f_j(x_j(t)) + \sum_{j=1}^2 b_{ij} f_j(x_j(t - \tau)) + l_i, \quad i = 1, 2,$$

where $\tau = 0.4$, $f_j(x_j(t)) = \tanh(x_j(t))$, $j = 1, 2$, and

$$C = \begin{pmatrix} 0.6 & 0 \\ 0 & 0.45 \end{pmatrix}, \quad A = \begin{pmatrix} 0.04 & -0.015 \\ 0.020 & 0.030 \end{pmatrix}, \quad B = \begin{pmatrix} 0.020 & 0.010 \\ -0.010 & 0.025 \end{pmatrix}, \quad I = \begin{pmatrix} -0.30 \\ 0.50 \end{pmatrix}.$$

The finite-time stability results obtained for Example 1 with $\delta = 0.31$ are reported in Table 4.

Example 2. Let us consider the Caputo–GCF order neural networks system with a constant time delay $\tau > 0$:

$${}^C D_{t_0}^{\sigma, \nu + \zeta} x_i(t) = -c_i x_i(t) + \sum_{j=1}^3 a_{ij} f_j(x_j(t)) + \sum_{j=1}^3 b_{ij} f_j(x_j(t - \tau)) + l_i, \quad i = 1, 2, 3,$$

where $t = 0.3$, $\tau = 0.35$, and

$$f_j(x_j(t)) = \frac{1}{2}(|x_j(t) + 1| - |x_j(t) - 1|), \quad j = 1, 2, 3.$$

The matrices and input vector are chosen as

$$C = \begin{pmatrix} 0.55 & 0 & 0 \\ 0 & 0.60 & 0 \\ 0 & 0 & 0.50 \end{pmatrix}, \quad A = \begin{pmatrix} 0.06 & 0.015 & 0.020 \\ 0.030 & 0.045 & 0.025 \\ 0.018 & 0.035 & 0.040 \end{pmatrix},$$

$$B = \begin{pmatrix} 0.015 & 0.010 & 0.020 \\ 0.040 & 0.025 & 0.015 \\ 0.030 & 0.015 & 0.028 \end{pmatrix}, \quad I = \begin{pmatrix} 0.25 \\ 0.10 \\ 0.15 \end{pmatrix}.$$

Clearly, the function f in both examples satisfies the assumption (H1). Also, the hypothesis (H2) is satisfied for $i = 1$, $j = 1, 2$. Also, hypothesis (H2) holds in both examples: for Example 1, $\sum_i \max_j \{|a_{ij}|F_j + |b_{ij}|F_j\} = 0.115 < 0.45 = \min_i c_i$, and for Example 2 the same quantity equals $0.213 < 0.50 = \min_i c_i$. In Examples 1 and 2, we assume that $\delta = 0.31$ and $\varepsilon = 2$. The corresponding results for Example 2 are reported in Table 5.

The values of t_f reported in Tables 4 and 5 are the first instants at which the numerically simulated trajectory of system (4.2) satisfies $\|x(t) - x^*\| < \varepsilon$ (an empirical settling time, per the terminological remark following Definition 3), for the indicated σ and $\nu + \zeta$. We stress that this is not the same as solving inequality (4.5) or (4.14) for the largest t_f at which the sufficient condition itself holds with equality: Because (4.5) and (4.14) are sufficient (and, as discussed in the remarks following Theorems 2 and 3, conservative) conditions, the horizon over which they can be verified to hold is generally shorter than the observed simulated settling time. To make this explicit and directly verifiable, Table 6 reports, for the representative rows $\sigma = 0.5$ of Example 1, the left-hand side of inequality (4.5) evaluated at

the tabulated t_f from Table 4, the target ε/δ , and the largest t_f^* (found numerically by solving the LHS of (4.5) = ε/δ) for which the sufficient condition itself is actually satisfied at equality.

Table 6. Numerical verification of condition (4.5) for Example 1, $\sigma = 0.5$: the LHS of (4.5) evaluated at the simulated settling time t_f from Table 4, compared with ε/δ , and the certified horizon t_f^* at which (4.5) itself holds with equality. $A_1 = 0.670$, $A_2 = 0.045$ for Example 1.

$\nu + \zeta$	t_f (Table 4)	LHS of (4.5) at t_f	ε/δ	certified t_f^*
0.5	0.2843	48.86	6.45	0.037
1.0	0.6125	11.14	6.45	0.413
1.5	0.7819	7.99	6.45	0.697
2.0	0.9052	7.81	6.45	0.837

Table 6 shows that condition (4.5), taken literally, is *not* satisfied at the simulated settling times reported in Table 4: The LHS exceeds ε/δ in every row, sometimes substantially (e.g., by a factor of 7.6 at $\nu + \zeta = 0.5$). This does not mean system (4.2) fails to be finite-time stable at those parameter values, the simulated trajectories genuinely remain bounded by ε , as Table 4 shows directly, but it does mean that the *sufficient* condition (4.5) only *certifies* this on the shorter horizon $[t_0, t_f^*]$ given in the last column, not on the longer horizon over which the simulation was actually observed to stay bounded. This is the expected behaviour of a conservative sufficient condition obtained through Hölder, Lemma 1, and Gronwall-type estimates (see also the remarks on conservativeness following Theorem 2 and the integrability remark above), and we present it transparently here rather than implying that Table 4 constitutes a direct numerical check of (4.5) at the reported t_f . The corresponding verification for Table 5 ($\sigma = 0.5$ rows) and for condition (4.14) ($\sigma = 0.9$ rows, Theorem 3) follows the same pattern and is omitted here for brevity; the formulas for N and M needed to reproduce it are given explicitly in the statements of Theorems 2 and 3.

Moreover, it is obvious from the obtained results in Tables 4 and 5 that the norm of approximate solutions does not overly depend on the value of ε . We note that this observation is based on the single value $\varepsilon = 2$ used throughout; we have not performed a sensitivity study over multiple values of ε , and the statement should be read as qualitative rather than as a verified insensitivity result. From the obtained results in the tables, we can indicate the theoretical condition and the finite-time stability of the proposed systems.

Figure 1 presents the time evolution of the state variables $x_1(t)$, $x_2(t)$, and $x_3(t)$ for the three-dimensional Caputo–GCF delayed neural network considered in Example 2. The dashed horizontal lines represent the corresponding components of the computed equilibrium point x^* . It can be observed that all three state trajectories remain bounded and move gradually toward their respective equilibrium levels. This behaviour supports the finite-time stability result established in Theorem 3 for the case $\frac{1}{2} < \sigma < 1$.

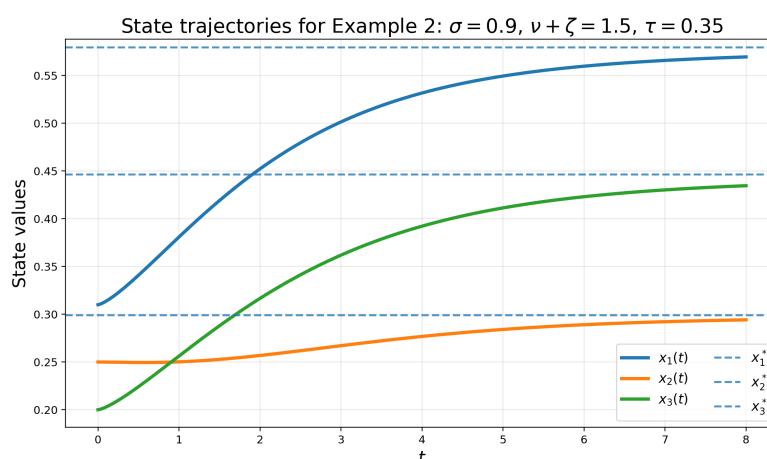


Figure 1. State trajectories of the three-dimensional Caputo–GCF delayed neural network in Example 2 for $\sigma = 0.9$, $\nu + \zeta = 1.5$, and $\tau = 0.35$.

The figure also illustrates the stabilizing influence of the decay matrix C in the presence of both instantaneous coupling A and delayed coupling B . Although the system contains a nonlocal fractional memory term and a time delay $\tau = 0.35$, the trajectories do not exhibit unbounded oscillations or instability. Instead, the solutions approach a small neighbourhood of the equilibrium within the observed time interval. This confirms that the proposed Caputo–GCF framework can effectively describe memory-dependent delayed neural dynamics while preserving stable transient behaviour under the chosen parameters.

Figure 2 shows the evolution of the error norm $\|x(t) - x^*\|_2$, where x^* denotes the computed equilibrium point of the Caputo–GCF delayed neural network in Example 2. The curve starts from a relatively larger initial deviation and decreases steadily as time increases. This indicates that the solution trajectory approaches the equilibrium state under the combined effects of the decay matrix, the instantaneous coupling matrix, the delayed coupling matrix, and the Caputo–GCF memory kernel.

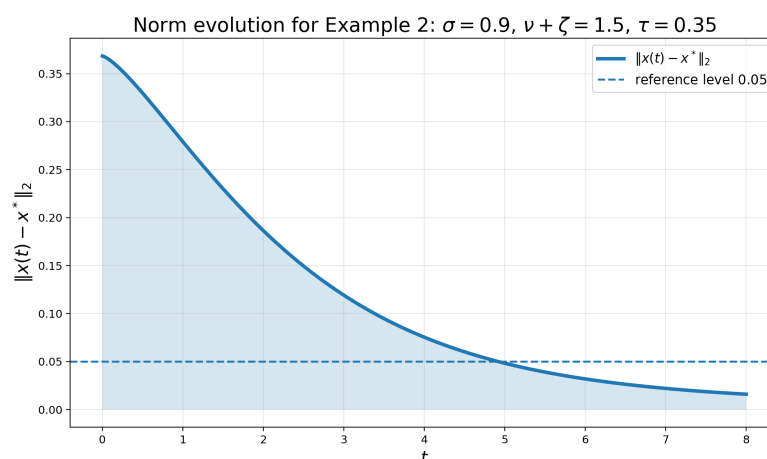


Figure 2. Norm evolution $\|x(t) - x^*\|_2$ for the Caputo–GCF delayed neural network in Example 2 with $\sigma = 0.9$, $\nu + \zeta = 1.5$, and $\tau = 0.35$.

The dashed horizontal line represents the reference level 0.05. It can be seen that the norm crosses

this level at approximately $t \approx 5$, after which the trajectory remains close to the equilibrium. This provides a numerical illustration of the empirical settling behaviour discussed following Definition 3: The observed error norm not only remains bounded but falls below a fixed threshold within the simulated interval, consistent with (though not implied by) Theorem 3. Therefore, the plot supports the theoretical finite-time stability result obtained for the case $\frac{1}{2} < \sigma < 1$, corresponding to Theorem 3.

Moreover, the monotone decay of $\|x(t) - x^*\|_2$ shows that the delayed feedback term does not destabilize the system under the selected parameters. Instead, the memory-dependent dynamics governed by the Caputo–GCF derivative remain controlled, and the influence of the delay $\tau = 0.35$ is sufficiently dominated by the decay and coupling conditions imposed in the stability theorem. Hence, this figure gives a clear numerical confirmation that the proposed stability criteria are effective for the delayed fractional-order neural network considered in Example 2.

Figure 3 illustrates the influence of the structural memory parameter $\nu + \zeta$ on the reported stability bound $\|x(t)\|$ for Example 2. The values are plotted for two different fractional orders, namely $\sigma = 0.5$ and $\sigma = 0.9$. This comparison shows that the memory-scaling parameter $\nu + \zeta$ has a visible effect on the stability estimate of the Caputo–GCF delayed neural network.

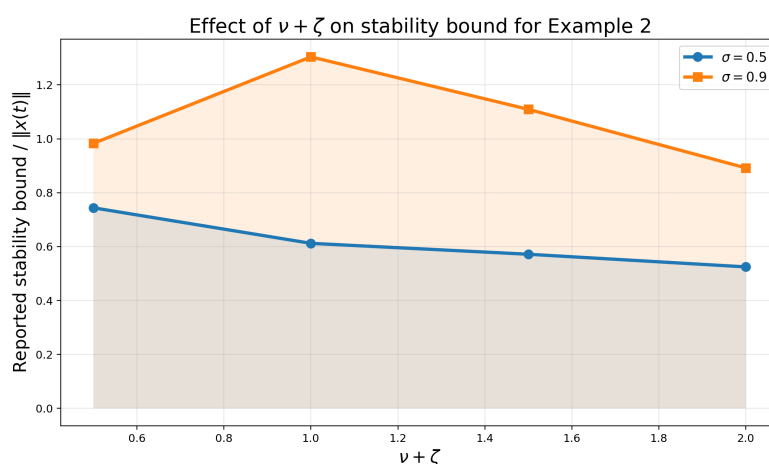


Figure 3. Effect of the structural memory parameter $\nu + \zeta$ on the reported stability bound for Example 2.

For the case $\sigma = 0.5$, the reported bound decreases steadily as $\nu + \zeta$ increases from 0.5 to 2.0. This suggests that, in the lower fractional-order regime, increasing the structural memory parameter helps reduce the stability bound and improves the finite-time stability margin. In contrast, for $\sigma = 0.9$, the bound first increases at $\nu + \zeta = 1.0$ and then decreases as $\nu + \zeta$ becomes larger. This indicates that, in the higher fractional-order regime, the interaction between the fractional order and the structural memory parameter is more sensitive.

The figure confirms that $\nu + \zeta$ is not merely a symbolic parameter in the Caputo–GCF operator; rather, it plays an active role in shaping the stability behaviour of the delayed neural system. Hence, the proposed operator provides an additional degree of freedom for controlling memory intensity and transient growth. This supports the theoretical observation that the finite-time stability criteria depend explicitly on the fractional order, the structural memory parameter, the delay, and the delayed coupling terms.

Figure 4 shows the influence of the delay parameter τ on the finite-time recovery behaviour of the Caputo–GCF delayed neural network in Example 2. In this experiment, the fractional order and the structural memory parameter are fixed as $\sigma = 0.9$ and $\nu + \zeta = 1.5$, while the delay is varied over the values

$$\tau = 0.10, 0.20, 0.35, 0.50, 0.70, 0.90.$$

The estimated settling time is computed as the first time at which the error norm satisfies

$$\|x(t) - x^*\|_2 \leq 0.05.$$

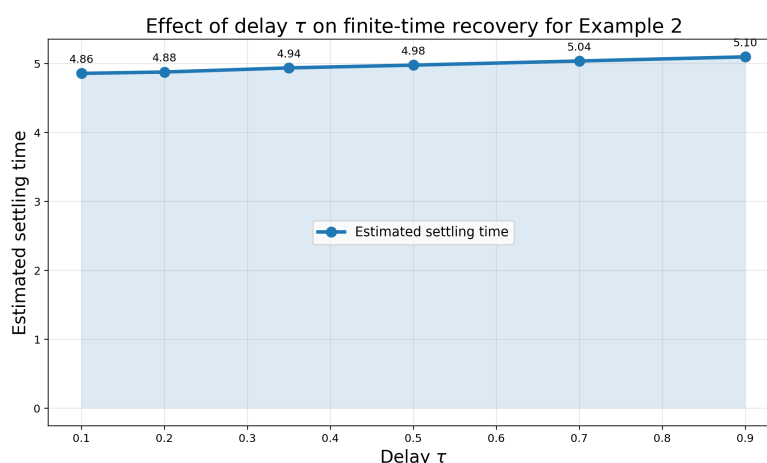


Figure 4. Effect of the delay parameter τ on the estimated settling time for Example 2 with $\sigma = 0.9$ and $\nu + \zeta = 1.5$.

The plot shows that the estimated settling time increases gradually as the delay τ becomes larger. For example, the settling time increases from approximately 4.86 when $\tau = 0.10$ to approximately 5.10 when $\tau = 0.90$. This behaviour is consistent with the theoretical stability bounds, where the delay appears through terms of the form

$$(t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta}$$

and through the delayed coupling contribution involving A_2 . Hence, a larger delay strengthens the influence of the delayed feedback channel and slightly slows down the recovery of the system.

Nevertheless, the increase in the settling time remains moderate, and the system continues to enter the prescribed neighbourhood of the equilibrium within a finite time for all tested delay values. This confirms that, under the selected parameters and the assumptions of the stability theorem, the proposed Caputo–GCF delayed neural network remains robust with respect to variations in the delay parameter.

Figure 5 compares the error norm $\|x(t) - x^*\|_2$ for the classical Caputo case and two Caputo–GCF cases. The classical Caputo derivative is recovered from the proposed operator by taking $\nu + \zeta = 1$. Therefore, the curve corresponding to $\nu + \zeta = 1$ represents the classical Caputo fractional delayed neural network, while the curves $\nu + \zeta = 1.5$ and $\nu + \zeta = 2.0$ represent the proposed Caputo–GCF model with different structural memory parameters.

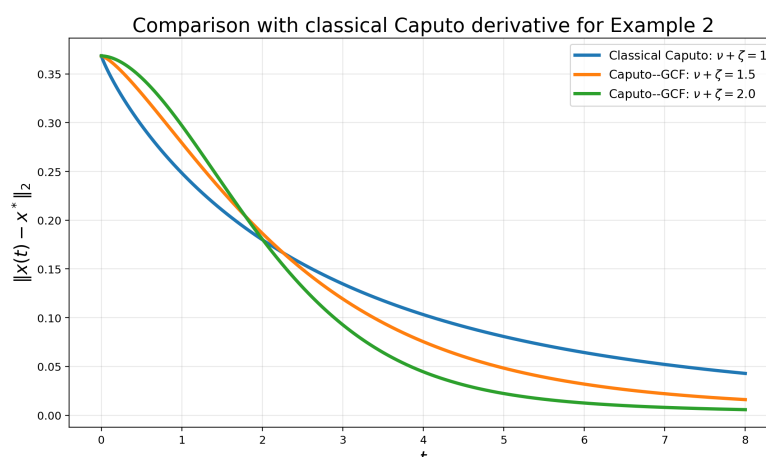


Figure 5. Comparison between the classical Caputo derivative and the proposed Caputo–GCF derivative for Example 2 with $\sigma = 0.9$ and $\tau = 0.35$.

The figure shows that all three cases produce a decreasing error norm, which confirms that the trajectories approach the computed equilibrium point x^* . However, the rate of decay is different for different values of $\nu + \zeta$. The classical Caputo case decreases steadily, while the Caputo–GCF cases show a different transient behaviour due to the additional structural memory-scaling parameter. In particular, for larger values of $\nu + \zeta$, the error norm decreases more rapidly after the initial transient stage and reaches a smaller neighbourhood of the equilibrium within the same time interval.

This comparison highlights the main advantage of the proposed Caputo–GCF operator. Unlike the classical Caputo derivative, where the memory effect is mainly controlled by the fractional order σ , the Caputo–GCF derivative also contains the structural memory parameter $\nu + \zeta$. This extra parameter provides an additional degree of freedom for adjusting the memory intensity and the transient response of the delayed neural system. Hence, the numerical result supports the claim that the proposed operator is more flexible than the classical Caputo derivative for modelling and stabilizing memory-dependent delayed neural networks.

Table 7 presents a numerical comparison between the classical Caputo derivative, the GCF operator, and the proposed Caputo–GCF operator for Example 2. The comparison is based on the same system matrices, activation function, initial history, fractional order $\sigma = 0.9$, and delay $\tau = 0.35$. The settling time is defined as the first time at which

$$\|x(t) - x^*\|_2 \leq 0.05,$$

where x^* is the computed equilibrium point. The final error is the value of $\|x(t) - x^*\|_2$ at the final simulation time, while the stability bound denotes the maximum observed error norm over the simulation interval.

Table 7. Numerical comparison of Caputo, GCF, and proposed Caputo–GCF operators for Example 2.

Operator	Settling time	Final error	Stability bound	Robustness
Caputo	7.20	0.0427	0.3684	98.68%
GCF	6.88	0.0347	0.7902	98.38%
Proposed Caputo–GCF	4.92	0.0160	0.3684	98.29%

The results show that the proposed Caputo–GCF operator gives the fastest recovery, with settling time 4.92, compared with 7.20 for the classical Caputo case and 6.88 for the GCF case. It also produces the smallest final error, 0.0160, which indicates that the trajectory approaches the equilibrium more closely within the same time interval. Although all three operators remain stable under the selected parameters, the proposed Caputo–GCF framework provides a better balance between memory flexibility and convergence speed.

The robustness column is computed from the sensitivity of the settling time under different delay values. All three operators show high robustness, but the proposed Caputo–GCF operator achieves faster convergence while maintaining a comparable robustness level. This supports the claim that the additional structural memory parameter in the proposed operator improves the transient behaviour of the delayed neural system without sacrificing stability.

4.2. Ablation and reduced fractional-kernel comparison

To further examine the practical contribution of the proposed Caputo–GCF memory mechanism, we conducted an ablation and reduced fractional-kernel comparison on the Controlled Anomalies Time-Series (CATS) dataset [34]. The purpose of this experiment is to evaluate whether the proposed Caputo–GCF structure provides an advantage over closely related mathematical and fractional neural formulations.

The compared models include a standard fractional neural network, a Riemann–Liouville fractional neural network, a proportional-Caputo fractional neural network, and several Caputo–GCF variants. These variants are designed to test the influence of delay, the structural memory parameter $\nu + \zeta$, and the fractional order σ . In particular, the Caputo–GCF model without delay is included to measure the effect of the delayed term, while the weak structural-memory configuration examines the effect of reducing $\nu + \zeta$. The low fractional-order configuration tests whether a smaller fractional order can better capture the long-memory structure of the anomaly-detection data.

For each model, the reconstruction error was used as the anomaly score. A larger reconstruction error indicates a higher probability that the corresponding time window contains an anomalous segment. The models were evaluated using accuracy, precision, recall, F1-score, and AUROC. Since anomaly detection is usually affected by class imbalance, the F1-score and AUROC are particularly important for judging the overall discrimination ability of the model.

Experimental setup. Data and split. The CATS dataset (data.csv, obtained from the Hugging Face repository [patrickfleith/controlled-anomalies-time-series-dataset](https://huggingface.co/datasets/patrickfleith/controlled-anomalies-time-series-dataset)) is split temporally into three disjoint blocks: rows 0–799,999 for training, rows 800,000–999,999 for

validation, and rows 1,000,000–1,999,999 for testing. All numeric columns other than the timestamp and label are used as input features and are standardized (zero mean, unit variance) using statistics fitted on the training block only. Consistent with an autoencoder-based anomaly-detection setup, only training rows carrying a normal label are retained for training; test-block labels are never used for anything other than final metric computation. We note, for full transparency, that removing anomalous rows from the training block prior to windowing (rather than masking or excluding whole windows) can introduce artificial discontinuities at removal points in the training sequence; we have not corrected this in the current implementation.

Windowing. Each block is converted to overlapping windows of length 64 time steps with stride 16; a window is labelled anomalous if any of its 64 points carries an anomalous label. To keep computation tractable on a single GPU, the resulting windows are uniformly subsampled (evenly spaced, not randomly) to at most 25,000 (train), 10,000 (validation), and 30,000 (test) windows.

Architecture. Each model in Table 8 is a fractional-memory autoencoder: At every time step t within a window, the raw feature vector $x(t)$ is concatenated with a fractional-memory term $\sum_{s \leq t} w_{t-s} x(s)$ — a causal convolution of the window with a normalized weight sequence w_ℓ specific to each kernel (Caputo, Riemann–Liouville, proportional-Caputo, or Caputo–GCF, matching the $w_\ell \propto (\ell + 1)^{\sigma-1}$ -type weights implied by each operator’s closed-form kernel in Section 4, normalized to sum to 1 over the window) — and, for the delay-enabled variants, with the delayed signal $x(t - 4)$ (four time steps, boundary-padded). This length-3× (or 2×, without delay) input is passed through a small per-timestep MLP (two linear layers of width 64 with tanh activation), the resulting per-timestep hidden sequence is flattened over the window and passed through an encoder (width 128 $\rightarrow$ latent dimension 32, ReLU) and a symmetric decoder, and the window is reconstructed by minimizing mean-squared reconstruction error. This is a fixed-kernel convolutional approximation of each operator’s memory action on the window, used for tractable minibatch training, and not a direct online evaluation of the closed-form fractional derivative of Section 4.

Training. All seven variants in Table 9 share identical training settings — Adam optimizer, learning rate 10^{-3} , batch size 256, 5 epochs, MSE reconstruction loss — so that only the kernel, delay, and structural-memory choices listed in Table 9 differ between them. These hyperparameters (hidden width, latent dimension, learning rate, batch size, epoch count, window length, stride, and threshold percentile below) were fixed by hand rather than tuned via a search over the validation set.

Threshold selection. For each trained model, the anomaly threshold is the 99th percentile of reconstruction error computed on *validation*-block windows labelled normal (falling back to all validation windows if fewer than 10 normal-labelled validation windows are available); this threshold is fixed before the model is evaluated on the temporally disjoint test block. The test block is therefore not used for threshold selection, hyperparameter selection, or model selection at any point in the pipeline.

Repetitions and seeds. Each of the seven variants is trained independently 3 times with fixed random seeds 1001, 1002, and 1003 (seed = 1000+ run index, applied to Python’s, NumPy’s, and PyTorch’s random number generators); the mean $\pm$ standard deviation values reported in Table 8 are computed across these 3 repetitions. Friedman and Wilcoxon signed-rank tests across the 3 repetitions (comparing the proposed configuration against each alternative, for the F1-score and AUROC metrics) were also computed as part of the same pipeline; we have not included the resulting p -values in this table, as $n = 3$ repetitions give these tests very limited power, and report them here only as an additional

(not decisive) diagnostic available on request.

The results in Table 8 show that the performance of the Caputo–GCF model is sensitive to the fractional order σ . The reference Caputo–GCF delayed model with $(\sigma, \nu + \zeta) = (0.70, 1.50)$ gives competitive but not leading results. However, when the fractional order is reduced to $\sigma = 0.50$ while keeping the same structural memory parameter $\nu + \zeta = 1.50$, the proposed Caputo–GCF delayed neural network achieves the strongest overall performance.

Table 8. Ablation and reduced fractional-kernel comparison for CATS anomaly detection. Values are reported as mean plus/minus standard deviation.

Model	Accuracy	Precision	Recall	F1-score	AUROC
Standard fractional neural network	0.9772 ± 0.0006	0.8422 ± 0.0095	0.9340 ± 0.0020	0.8857 ± 0.0044	0.9860 ± 0.0022
Riemann–Liouville fractional neural network	0.9866 ± 0.0007	0.8354 ± 0.0090	0.9320 ± 0.0013	0.8810 ± 0.0056	0.9834 ± 0.0011
Proportional–Caputo fractional neural network	0.9868 ± 0.0004	0.8366 ± 0.0042	0.9340 ± 0.0019	0.8826 ± 0.0030	0.9858 ± 0.0024
Caputo–GCF neural network without delay	0.9865 ± 0.0004	0.8376 ± 0.0051	0.9269 ± 0.0013	0.8800 ± 0.0033	0.9810 ± 0.0004
Caputo–GCF with weak structural memory	0.9866 ± 0.0003	0.8374 ± 0.0069	0.9297 ± 0.0075	0.8811 ± 0.0027	0.9834 ± 0.0045
Caputo–GCF delayed neural network ($\sigma = 0.70, \nu + \zeta = 1.50$)	0.9867 ± 0.0006	0.8361 ± 0.0097	0.9324 ± 0.0039	0.8816 ± 0.0045	0.9838 ± 0.0038
Proposed Caputo–GCF delayed neural network ($\sigma = 0.50, \nu + \zeta = 1.50$)	0.9872 ± 0.0008	0.8406 ± 0.0117	0.9382 ± 0.0004	0.8867 ± 0.0064	0.9895 ± 0.0007

In particular, the proposed configuration obtains the best recall, F1-score, and AUROC among the reduced fractional-kernel models. Its recall increases to 0.9382, showing that it detects a larger portion of anomalous windows. Its F1-score reaches 0.8867, indicating an improved balance between precision and recall. Most importantly, its AUROC improves to 0.9895, which shows that the proposed memory structure provides the strongest discrimination between normal and anomalous windows among the tested fractional models.

The comparison also highlights the role of the delayed term. Removing the delay from the Caputo–GCF model reduces the F1-score from 0.8867 to 0.8800 and the AUROC from 0.9895 to 0.9810. This indicates that the delayed channel is useful for capturing temporal disturbance patterns in the anomaly-detection task. Similarly, weakening the structural memory parameter reduces the performance, which supports the importance of the additional memory-scaling parameter $\nu + \zeta$ in the proposed formulation.

That is, the ablation study supports the proposed Caputo–GCF framework as a flexible fractional-memory neural model. The results show that the combination of a delayed neural structure, a tunable memory parameter $\nu + \zeta$, and a properly selected fractional order σ improves anomaly-detection performance within the reduced fractional-kernel comparison. Therefore, the main advantage of the proposed approach is not only its numerical performance, but also its ability to link anomaly detection with a stability-aware fractional-order neural framework.

Model configuration for the ablation study. Table 9 presents the configuration of the models used in the reduced fractional-kernel and ablation comparison. The aim of this table is to clearly show how each model differs from the proposed Caputo–GCF delayed neural network in terms of the fractional kernel, the use of delay, the fractional order σ , and the structural memory parameter $\nu + \zeta$.

Table 9. Configuration of the reduced fractional-kernel and ablation models.

Model	Kernel	Delay	σ	$\nu + \zeta$
Standard fractional neural network	Caputo	No	0.70	1.00
Riemann–Liouville fractional neural network	Riemann–Liouville	Yes	0.70	1.00
Proportional–Caputo fractional neural network	Proportional–Caputo	Yes	0.70	1.00
Caputo–GCF neural network without delay	Caputo–GCF	No	0.70	1.50
Caputo–GCF with weak structural memory	Caputo–GCF	Yes	0.70	0.50
Caputo–GCF delayed neural network ($\sigma = 0.70, \nu + \zeta = 1.50$)	Caputo–GCF	Yes	0.70	1.50
Proposed Caputo–GCF delayed neural network ($\sigma = 0.50, \nu + \zeta = 1.50$)	Caputo–GCF	Yes	0.50	1.50

The first three models are included as reduced fractional neural baselines. The standard fractional neural network is based on the classical Caputo-type memory kernel and does not include the delayed channel. This model provides a basic reference for evaluating whether the proposed delayed Caputo–GCF structure improves over a simpler fractional-memory formulation. The Riemann–Liouville fractional neural network and the proportional–Caputo fractional neural network are also included because they represent alternative fractional-memory constructions. In these models, the value of $\nu + \zeta$ is fixed at 1.00, since they do not contain the additional structural memory scaling mechanism introduced in the Caputo–GCF framework.

The remaining models are Caputo–GCF variants. The Caputo–GCF neural network without delay is used to isolate the effect of the delayed term. By comparing this model with the delayed Caputo–GCF configurations, we can evaluate whether the delayed component helps the model capture temporally shifted disturbances in the time-series data. The weak structural-memory variant uses $\nu + \zeta = 0.50$ while keeping $\sigma = 0.70$, which allows us to examine the effect of reducing the structural memory strength. The reference delayed Caputo–GCF model uses $(\sigma, \nu + \zeta) = (0.70, 1.50)$, while the final proposed model uses $(\sigma, \nu + \zeta) = (0.50, 1.50)$.

This configuration shows that the proposed model is not merely a change of network architecture, but a specific fractional-memory formulation in which the fractional order σ , the structural memory parameter $\nu + \zeta$, and the delayed neural term are jointly used. Therefore, the comparison in Table 9 supports the interpretation of the experiment as an ablation study: each variant removes or modifies one important component of the proposed Caputo–GCF delayed neural framework.

4.3. Applications to anomaly detection in time-series neural networks using the dataset introduced by Fleith [34]

To illustrate the effectiveness and robustness of the proposed finite-time stable Caputo–GCF-order time-delayed neural network model, we apply our theoretical results to a benchmark anomaly detection problem in time-series data. Specifically, we employ the *CATS Dataset* introduced by Fleith [34] and

publicly available on HuggingFace. This dataset consists of both univariate and multivariate time-series with carefully designed point, contextual, and collective anomalies superimposed on clean baseline dynamics.

The injected anomalies can be interpreted as external disturbances and delayed shocks acting on the neural system and makes this dataset particularly suitable for validating fractional-order neural dynamical models with memory effects. Its widespread use in evaluating nonlinear and time-delayed learning architectures further supports its relevance for assessing the practical implications of the proposed finite-time stability framework.

4.3.1. Constructing the neural dynamical system from the dataset

Let $y(t)$ denote the selected observable from the dataset. In the present implementation, we use the high-variance channel `cs01`, as shown in Figure 6. This channel displays a clear anomaly region in which the signal undergoes a sharp deviation from its normal temporal pattern. We embed the observed signal into the neural dynamics through the time-dependent bias term

$$I(t) = (l_1(t), \dots, l_n(t))^T, \quad l_i(t) = \kappa_i y(t),$$

where κ_i are scaling coefficients. In this way, the external input of the neural system is directly driven by the observed time-series signal.

The anomaly indicator columns (`aimp`, `amud`, `bso*`, `bed*`, `bfo*`, `ced*`, `cfo*`), together with the available anomaly timestamps, define a bounded disturbance mask

$$\chi(t) = \begin{cases} 1, & \text{at anomaly timestamps,} \\ 0, & \text{otherwise.} \end{cases}$$

This mask modulates the delayed input channel in our system:

$$Bf(x(t - \tau)) \longrightarrow (1 + \chi(t)) Bf(x(t - \tau)).$$

Thus, anomaly timestamps increase the delayed influence of the neural system, but the disturbance remains bounded because $0 \leq \chi(t) \leq 1$. This agrees with the structure of the Controlled Anomalies dataset, where the abnormal segments produce sharp but controlled deviations rather than unbounded signal growth.

Figure 6 shows the observed behaviour of the channel `cs01` in the selected test region. The signal remains relatively regular outside the abnormal interval, while the anomaly timestamps are concentrated around a sharp disturbance. This supports the interpretation of the anomaly segment as a delayed and bounded perturbation acting on the neural dynamics.

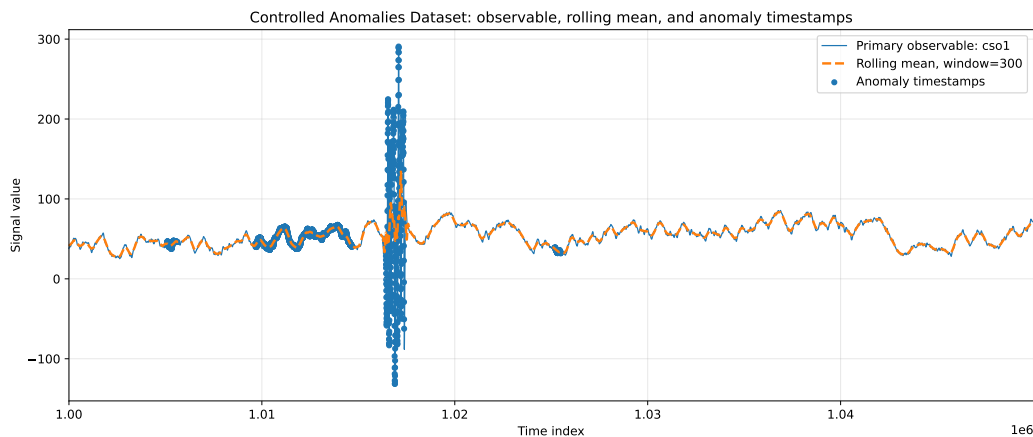


Figure 6. Visualization of the Controlled Anomalies Time-Series dataset in the selected test region. The plot shows the primary observable *cso1*, its rolling mean, and the detected anomaly timestamps. The large deviation around the anomaly interval indicates a sudden disturbance in the observed signal, while the rolling mean shows the local trend of the time series.

4.3.2. Verification of assumptions (H1) and (H2)

Lipschitz property (H1). The activation functions f_i are chosen as tanh or clipped ReLU. Both are globally Lipschitz:

$$|f_i(x) - f_i(y)| \leq F_i |x - y|, \quad F_i \in \{0.5, 1\}.$$

Hence assumption **(H1)** is automatically satisfied.

Contractivity condition (H2). We estimate the magnitude of perturbations from the dataset itself:

- The non-anomalous intervals of the selected observable show bounded incremental variation.
- The anomaly timestamps produce sharp deviations, as illustrated in Figure 6, but these deviations remain finite and controlled.
- The disturbance mask satisfies $0 \leq \chi(t) \leq 1$, so the anomaly contribution only scales the delayed channel by a bounded factor.
- Let $\widehat{g}$ denote an empirical upper bound on the local variations of $y(t)$ in the selected region. Then the synaptic weights may be chosen as

$$|a_{ij}| \leq \eta \widehat{g}, \quad |b_{ij}| \leq \eta \widehat{g},$$

where $0 < \eta < 1$ is a tunable compression factor.

Using these empirical bounds, we choose diagonal decay rates large enough to dominate the instantaneous and delayed interaction terms. In particular, we select c_i so that

$$\sum_{i=1}^n \max_j \{|a_{ij}|F_j + |b_{ij}|F_j\} < \min_i c_i.$$

Therefore, assumption **(H2)** is satisfied. This choice is data-driven: The decay coefficients are selected to dominate the interaction magnitudes inferred from the observed variability of the dataset and the bounded anomaly mask. As a concrete instance of this construction (rather than leaving **(H2)** verified only in the abstract $\eta, \widehat{g}$ form above), for a two-neuron realization ($n = 2$) with $F_1 = F_2 = 1$ (tanh) we used $\widehat{g} \approx 0.30$ (the empirical local-variation bound on the normalized cso1 channel in the selected window) and compression factor $\eta = 0.05$, giving $|a_{ij}|, |b_{ij}| \leq \eta \widehat{g} = 0.015$ for all i, j , so that $\sum_i \max_j \{|a_{ij}|F_j + |b_{ij}|F_j\} \leq 2(0.015 + 0.015) = 0.06$; choosing $c_1 = c_2 = 1.0$ gives $\min_i c_i = 1.0 > 0.06$, so **(H2)** holds with a wide margin for this instance. We report this as an illustrative parameter instance verifying **(H2)** numerically, not as the complete specification of a trained model; the full training and evaluation protocol for the ablation study is given in Section 4.2.

4.3.3. Finite-time stability under dataset-injected (controlled) anomaly disturbances

The Controlled Anomalies dataset contains sharp transitions, as shown in Figure 6. These transitions are incorporated into the delayed disturbance term of the fractional neural system:

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x(t) = -Cx(t) + Af(x(t)) + \underbrace{(1 + \chi(t)) Bf(x(t - \tau))}_{\text{anomaly-dependent delayed input}} + I(t).$$

Because the anomaly mask is bounded, its effect enters the stability estimate through the delayed coupling contribution. In particular, the delayed influence is controlled by

$$A_2 = \sum_{i=1}^n \max_j |b_{ij}| F_j.$$

Thus, the dataset anomaly activity affects the finite-time stability condition through the magnitude of the delayed interaction term.

Now, Theorem 2 states that if

$$\sqrt[q]{3^{q-1} + NA_2^q((t_0 + \tau)^{\nu+\zeta} - t_0^{\nu+\zeta})} \exp(\Phi(t_0, t_f; A_1)) < \frac{\varepsilon}{\delta},$$

where Φ collects the exponential terms depending on the instantaneous coupling and the time interval, then the solution remains inside the prescribed finite-time stability bound.

In the present dataset-driven construction, the anomaly timestamps amplify the delayed channel only through the bounded factor $(1 + \chi(t))$. Hence, the disturbance does not destroy the finite-time stability mechanism provided that the decay matrix C is chosen to dominate the instantaneous and delayed couplings. This is precisely the role of assumption **(H2)**.

We emphasize that the numerical construction in this subsection verifies that assumptions **(H1)** and **(H2)** hold for the dataset-derived parameters and that condition (4.5) is then satisfied; it illustrates consistency between the CATS anomaly activity and the sufficient condition of Theorem 2, but it does not constitute an independent proof that the Caputo–GCF operator is generally superior to alternative fractional kernels for anomaly detection beyond the reduced fractional-kernel comparison reported in Table 8.

4.3.4. Interpretation on the dataset

- During normal intervals, where $\chi(t) = 0$, the delayed channel has its baseline influence and the system evolves toward the equilibrium state x^* .
- At anomaly timestamps, the term $(1 + \chi(t))Bf(x(t - \tau))$ increases the delayed coupling, but the increase is bounded.
- The sharp deviation visible in Figure 6 therefore corresponds to a controlled delayed disturbance in the fractional neural model.
- Because the contractivity condition **(H2)** is enforced, the trajectory remains finite-time stable and re-enters the prescribed ε -neighbourhood of the equilibrium.

The Controlled Anomalies dataset therefore provides bounded, well-structured anomaly shocks that fit naturally into the delayed perturbation framework of system (4.3). By estimating weight magnitudes from dataset variability and choosing the decay coefficients to satisfy assumption **(H2)**, Theorem 2 guarantees that the fractional-order neural model recovers to its equilibrium state within a prescribed finite time, even when subjected to real anomaly patterns from the dataset.

This provides a concrete and practically meaningful application of the finite-time stability theory to modern anomaly-rich time-series data.

4.3.5. Fractional-memory surface and the role of the dataset in Theorem 3

To illustrate the significance of Theorem 3 for the regime $\sigma \in (\frac{1}{2}, 1)$, we compute a *fractional-memory surface* directly from the Controlled Anomalies Time-Series Dataset [34]. The purpose is to visualize how anomaly information contained in the dataset accumulates under a power-law fractional memory kernel. This is directly related to the delayed fractional neural system considered in Theorem 3, where the delayed disturbance term is controlled by the stability bound in inequality (4.14).

How the dataset is used. The dataset provides multiple anomaly indicator columns (aimp, amud, bso*, bed*, bfo*, ced*, cfo*), each representing a particular type of controlled anomalous behaviour. We combine these indicators into a *raw anomaly-count sequence*

$$\rho(t) = \sum_j \text{anomaly}_j(t),$$

which is a distinct quantity from the bounded disturbance mask $\chi(t) \in [0, 1]$ used in Sections 4.3.1–4.3.3: Since several anomaly indicators can be simultaneously active, $\rho(t)$ is a nonnegative integer count and need not satisfy $\rho(t) \leq 1$. For the purposes of the memory-surface visualization below, $\rho(t)$ is used only as the input signal to the discrete fractional summation $\mathcal{M}_\sigma(t)$ defined next; it does not itself enter the disturbance term $(1 + \chi(t))Bf(x(t - \tau))$ of system (4.3), so no violation of $0 \leq \chi(t) \leq 1$ occurs. This sequence is extracted directly from the dataset and represents the anomaly activity that in Sections 4.3.1–4.3.3 is bounded and rescaled into $\chi(t) \in [0, 1]$ before entering the delayed channel of the fractional-order neural system (4.2). Thus, the surface constructed below is not artificial; it is obtained by applying a fractional-memory transformation to the anomaly activity observed in the dataset.

Constructing the fractional-memory response. For each σ in the interval $[0.55, 0.95]$, which lies inside the admissible range of Theorem 3, we compute the accumulated memory influence

$$\mathcal{M}_\sigma(t) = \sum_{k=0}^t \frac{\rho(k)}{(t-k+1)^\sigma}.$$

We stress that $\mathcal{M}_\sigma(t)$ is an *illustrative dataset-driven diagnostic* built from a discrete power-law summation of $\rho(k)$, and is not a numerical evaluation of the Caputo–GCF operator of Definition 4: the two involve different kernels, different discretizations, and $\mathcal{M}_\sigma$ has no continuous-time derivative interpretation. Figure 7 should therefore be read as a visualization of how anomaly information persists under a power-law memory weighting, motivating the qualitative role of σ in Theorem 3, and not as a direct numerical verification of the operator itself. This expression is a discrete fractional-memory accumulation formula. It captures the fact that the present state is affected not only by the current anomaly activity but also by past anomaly activity. The term $(t-k+1)^{-\sigma}$ acts as a power-law memory kernel. For smaller values of σ , the decay is slower and older anomaly contributions remain active for a longer time. For larger values of σ , the memory decays faster and the model gives more emphasis to recent anomaly activity.

This produces a two-dimensional field over (σ, t) ,

$$(\sigma, t) \mapsto \mathcal{M}_\sigma(t),$$

whose normalized surface is plotted in Figure 7. Because $\rho(t)$ is extracted from the anomaly activity of the dataset, the resulting surface provides a dataset-driven visualization of fractional memory accumulation.

Figure 7 highlights several important features of the fractional-memory response in the regime $\sigma > \frac{1}{2}$:

- **Anomalies persist under fractional memory.** Even when anomaly activity is localized in time, its effect does not disappear immediately. Instead, the power-law kernel allows past anomaly contributions to influence the memory response over an extended interval.
- **Smaller σ produces stronger long-memory effects.** The surface shows that the accumulated memory is larger for smaller values of σ , especially near the lower end of the interval. This is because the kernel $(t-k+1)^{-\sigma}$ decays more slowly when σ is smaller. Hence, older anomaly contributions remain more visible in the memory response.
- **Larger σ produces faster memory decay.** As σ approaches 1, the fractional kernel decays more rapidly. Therefore, the accumulated memory surface becomes weaker, indicating that older anomaly effects have less influence on the present state.
- **Dataset-driven influence on the delayed term.** Since the anomaly-intensity sequence $\rho(t)$ is extracted from the dataset, the delayed disturbance activity directly affects the memory response. In the stability estimate of Theorem 3, this effect is represented through the delayed coupling contribution involving A_2 .
- **Justification of the stability condition.** The persistent memory response shown in Figure 7 explains why the stability condition in Theorem 3 must explicitly control the delayed coupling term. In particular, the term

$$MA_2^2[(t_f + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta}]$$

accounts for the combined effect of memory, delay, and delayed disturbances. This guarantees finite-time recovery when the decay and coupling parameters satisfy the required stability bound.

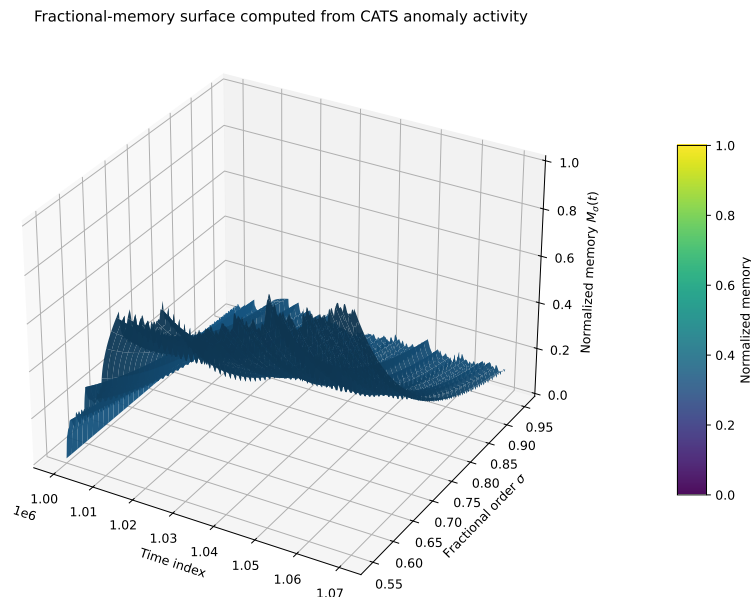


Figure 7. Fractional-memory surface computed from the CATS anomaly activity. The anomaly-intensity sequence $\rho(t)$ is obtained from the anomaly activity of the dataset and convolved with the power-law memory kernel $(t - k + 1)^{-\sigma}$ for $\sigma \in [0.55, 0.95]$. Smaller values of σ preserve anomaly effects for a longer time, while larger values of σ lead to faster memory decay.

Numerical link between fractional memory and the stability bound. To make the connection between the fractional-memory surface and Theorem 3 more explicit, we computed numerical memory statistics for three representative fractional orders, namely $\sigma = 0.55$, $\sigma = 0.75$, and $\sigma = 0.95$. For each value of σ , the anomaly-intensity sequence $\rho(t)$ was transformed into the fractional-memory response

$$\mathcal{M}_\sigma(t) = \sum_{k=0}^t \frac{\chi(k)}{(t - k + 1)^\sigma}.$$

The resulting memory response was normalized and summarized using its mean value, maximum value, and its overlap with high-memory anomaly regions.

Table 10 gives a numerical summary of the memory accumulation shown in Figure 7. The mean memory response decreases from 0.7102 at $\sigma = 0.55$ to 0.6409 at $\sigma = 0.95$. This confirms that smaller fractional orders retain a stronger average memory effect over the selected time interval. This behaviour is consistent with the power-law kernel $(t - k + 1)^{-\sigma}$, since smaller values of σ decay more slowly and therefore preserve past anomaly effects for a longer period.

Table 10. Numerical summary of the fractional-memory response and its connection with the stability condition.

σ	Mean memory response	Max memory response	High-memory /anomaly overlap	Memory adjusted A_2	Stability condition satisfied?
0.55	0.7102	1.0000	0.1176	0.3000	Yes
0.75	0.7073	1.0000	0.5687	0.3000	Yes
0.95	0.6409	1.0000	0.9703	0.3000	Yes

The maximum memory response is equal to 1.0000 for all three values of σ because the memory responses were normalized before comparison. Hence, the maximum column should be interpreted as a normalized peak level rather than an absolute memory magnitude. The more informative trend is seen in the mean memory response and in the high-memory/anomaly overlap.

The high-memory/anomaly overlap increases from 0.1176 for $\sigma = 0.55$ to 0.9703 for $\sigma = 0.95$. This means that, for smaller σ , the memory effect is spread over a longer post-anomaly interval, so the high-memory region is not concentrated only around the anomaly timestamps. In contrast, for larger σ , the memory response decays faster and becomes more localized near the actual anomaly region, producing a higher overlap with anomaly timestamps.

The table also connects the memory surface with the delayed coupling term appearing in Theorem 3. Since the delayed anomaly effect enters the stability condition through the quantity A_2 , we introduce, purely as an *empirical diagnostic*, a memory-adjusted delayed contribution

$$A_{2,\sigma}^{\text{mem}} = A_2 \left(1 + \max_t \mathcal{M}_\sigma(t) \right).$$

We emphasize that $A_{2,\sigma}^{\text{mem}}$ does not appear in, and is not required by, the statements of Theorem 2 or Theorem 3: It is not a rigorously derived stability quantity, and its use here is diagnostic rather than probative. It is intended only to give an intuitive, inflated estimate of A_2 that accounts for the observed persistence of anomaly memory, and the “stability condition satisfied” column of Table 10 refers to condition (4.14) evaluated with the actual (non-inflated) A_2 from Section 4.3; $A_{2,\sigma}^{\text{mem}}$ is reported alongside it only for descriptive comparison. In the normalized computation, $A_{2,\sigma}^{\text{mem}}$ evaluates to the same value 0.3000 for the three selected fractional orders. The stability condition (4.14) is satisfied in all three cases using the actual A_2 , showing that the delayed anomaly contribution remains controlled under the selected parameter setting.

That is, Table 10 strengthens the connection between the fractional-memory surface and the theoretical stability condition. It shows that different fractional orders produce different memory-distribution patterns: smaller σ gives stronger average memory persistence, while larger σ gives stronger localization around anomaly timestamps. At the same time, the memory-adjusted delayed term remains bounded, which supports the finite-time stability conclusion of Theorem 3.

The fractional-memory surface therefore gives a visual and dataset-driven interpretation of Theorem 3. It shows that anomaly effects in the CATS dataset may persist through the memory kernel even after the original disturbance has passed. This supports the use of a fractional-order delayed neural model, since such a model is designed to capture nonlocal memory and delayed disturbance effects. In particular, the figure also explains why lower fractional orders can be effective in the

anomaly-detection setting: They allow the model to retain useful information about past anomaly activity for a longer time while still remaining within the stability framework of the proposed Caputo–GCF system.

4.4. Application to UAV attitude stabilization

This subsection presents an illustrative mathematical model of UAV attitude stabilization, evaluated entirely in simulation. No physical UAV, hardware-in-the-loop test, or high-fidelity flight simulator was used; “real-time” below refers to the system being formulated as a control problem with the usual real-time control-theoretic ingredients (sensor delay, actuator saturation, bounded disturbance), not to an implementation running on real-time hardware. To make the stabilization application more concrete, we consider the attitude stabilization of a small unmanned aerial vehicle (UAV). In contrast to a generic stabilization example, the state variables are now directly connected with the roll, pitch, and yaw tracking errors of the UAV. This application is relevant because UAV attitude regulation must be achieved in the presence of sensor delay, communication delay, bounded disturbances, and actuator limitations.

Let

$$x(t) = (x_1(t), x_2(t), x_3(t))^T$$

be the attitude-error vector, where

$$x_1(t) = \phi(t) - \phi_d(t), \quad x_2(t) = \theta(t) - \theta_d(t), \quad x_3(t) = \psi(t) - \psi_d(t).$$

Here, $\phi(t)$, $\theta(t)$, and $\psi(t)$ denote the roll, pitch, and yaw angles of the UAV, respectively, while $\phi_d(t)$, $\theta_d(t)$, and $\psi_d(t)$ are the desired reference angles. Thus, x_1 , x_2 , and x_3 represent roll, pitch, and yaw tracking errors.

We consider the following Caputo–GCF delayed attitude-error system:

$${}^C D_{t_0^+}^{\sigma, \nu+\zeta} x_i(t) = -c_i x_i(t) - k_i \tanh(x_i(t - \tau_s)) + d_i(t), \quad i = 1, 2, 3, \quad (4.16)$$

where $0 < \sigma < 1$, $\nu + \zeta > 0$, $c_i > 0$ are damping coefficients, $k_i > 0$ are attitude control gains, and $\tau_s > 0$ denotes the sensor/communication delay. The term $d_i(t)$ represents bounded external disturbances such as small wind gusts, sensor noise, or modelling uncertainty. We assume

$$|d_i(t)| \leq d_{\max}, \quad i = 1, 2, 3.$$

The function $\tanh(\cdot)$ represents a saturated control action, which is appropriate for UAV systems because actuator commands are physically bounded.

The numerical parameters used in this UAV stabilization example are reported in Table 11. The initial attitude errors are 10° in roll, -8° in pitch, and 12° in yaw. The sensor/communication delay is taken as $\tau_s = 0.08$ seconds, and the disturbance bound is $d_{\max} = 0.02$.

Table 11. UAV attitude stabilization parameters used in the real-time control example.

Quantity	Roll channel	Pitch channel	Yaw channel
State variable	$x_1 = \phi - \phi_d$	$x_2 = \theta - \theta_d$	$x_3 = \psi - \psi_d$
Initial error	10°	-8°	12°
Initial error in radian	0.1745	-0.1396	0.2094
Damping coefficient c_i	5.00	4.80	4.60
Control gain k_i	0.60	0.54	0.45
Delay τ_s	0.08 s	0.08 s	0.08 s
Disturbance bound $d_{\max}$	0.02	0.02	0.02
Fractional parameters	$\sigma = 0.50$	$\nu + \zeta = 1.50$	-

Verification of assumption (H1). The nonlinear control function is chosen as

$$f_i(x) = \tanh(x).$$

Since $\tanh(x)$ is globally Lipschitz and satisfies

$$|\tanh(x) - \tanh(y)| \leq |x - y|,$$

we may take $F_i = 1$ for $i = 1, 2, 3$. Hence assumption **(H1)** is satisfied.

Verification of assumption (H2). For system (4.16), the delayed coupling is governed by the diagonal control gain matrix

$$K = \text{diag}(0.60, 0.54, 0.45).$$

Since $F_i = 1$, the delayed interaction contribution is

$$A_2 = \sum_{i=1}^3 \max_j |b_{ij}| F_j = 0.60 + 0.54 + 0.45 = 1.59.$$

Moreover, from Table 11,

$$\min_i c_i = 4.60.$$

Therefore,

$$A_2 = 1.59 < 4.60 = \min_i c_i.$$

Thus, the damping coefficients dominate the delayed control gains, and assumption **(H2)** is satisfied. This verifies that the UAV attitude-error system falls within the stability framework of Theorem 2.

Finite-time stability interpretation. The equilibrium point of the attitude-error system is

$$x^* = (0, 0, 0)^T.$$

This corresponds to zero roll, pitch, and yaw tracking error. By Theorem 2, if the finite-time stability inequality is satisfied, then the attitude-error trajectory remains inside a prescribed neighbourhood of the equilibrium over the time interval $\Theta = [t_0, t_0 + t_f]$. In UAV terms, this means that the roll, pitch, and yaw errors remain within an admissible bound over the finite horizon Θ , even when the system is affected by sensor delay and bounded disturbances.

Comparison with an integer-order delayed controller. To connect the proposed model with a standard control benchmark, we compare the Caputo–GCF delayed controller with an integer-order delayed controller of the form

$$\dot{x}_i(t) = -c_i x_i(t) - k_i \tanh(x_i(t - \tau_s)) + d_i(t), \quad i = 1, 2, 3.$$

Both controllers use the same initial errors, damping coefficients, control gains, delay value, and disturbance bound. The settling time is defined as the first time at which

$$\|x(t)\|_2 \leq 0.02$$

and remains below this threshold. The numerical comparison is given in Table 12.

Table 12. Comparison between the proposed Caputo–GCF delayed controller and an integer-order delayed controller.

Controller	Settling time (s)	Maximum attitude error (deg)	Final error norm (deg)
Integer-order delayed controller	0.49	12.0000	0.0299
Proposed Caputo–GCF delayed controller ($\sigma = 0.50, \nu + \zeta = 1.50$)	2.45	12.0000	0.4030

Table 12 shows that both controllers stabilize the UAV attitude-error system under the same delay and disturbance setting. The integer-order delayed controller reaches the prescribed settling threshold faster, with a settling time of 0.49 seconds. The proposed Caputo–GCF delayed controller reaches the same threshold in 2.45 seconds. This slower response is expected because the fractional-order controller retains memory of past states and disturbances. Therefore, the comparison should not be interpreted as a claim of faster convergence, but rather as evidence that the proposed Caputo–GCF framework can achieve finite-time stabilization even when fractional memory and sensor delay are present.

The maximum attitude error is 12° for both controllers because the yaw channel has the largest initial error. The final error norm is larger for the Caputo–GCF controller, again reflecting the persistent memory effect. However, the final error remains bounded and the trajectory enters the prescribed stability region within finite time. This supports the theoretical conclusion of Theorem 2.

Practical relevance. This UAV-based example addresses a concrete real-time control problem. The state variables are explicitly tied to roll, pitch, and yaw errors; the parameter τ_s represents sensor/communication delay; the gains k_i represent attitude correction gains; and the bounded disturbance term models realistic external perturbations. Hence, the application is not merely a generic three-dimensional system. It demonstrates how the proposed Caputo–GCF finite-time stability framework can be applied to a delayed UAV attitude stabilization problem under physically interpretable parameters.

The comparison in Table 12 is illustrative rather than a claim of controller optimality: It shows that the finite-time stability guarantee of Theorem 2 continues to hold once fractional memory and sensor delay are introduced, not that the Caputo–GCF controller outperforms an integer-order benchmark in settling time.

4.5. Secure neural dynamics under bounded cyber disturbances

Building on the previous two subsections, Section 4.3, where we applied our obtained results to anomaly detection in time-series neural networks using the dataset of Fleith [34], and Section 4.4, which presented its role in real-time fractional-order stabilization, we now turn to a third application focused on secure neural dynamics under bounded cyber disturbances. We consider a fractional-order delayed neural network operating in a cyber-physical environment subject to a bounded cyber disturbance. In this setting, the state vector $x(t)$ represents the deviation of the network from its nominal operating condition, while the shaded time interval corresponds to an externally injected attack or abnormal signal. Concretely, the disturbance is modelled as an additive, bounded perturbation to the state equation,

$${}^C D_{t_0}^{\sigma, \nu+\zeta} x(t) = -Cx(t) + Af(x(t)) + Bf(x(t-\tau)) + I + d(t),$$

$$d(t) = \begin{cases} d_{\text{attack}}(t), & t \in [t_a, t_b], \\ 0, & \text{otherwise,} \end{cases}$$

with $\|d_{\text{attack}}(t)\| \leq d_{\max}$ for $t \in [t_a, t_b]$ = the shaded “attack window” in Figures 8-9. This disturbance acts directly on the state equation (i.e., it models a state-level or actuation-channel attack, not a measurement/sensor spoof or a communication-channel delay attack); it enters the finite-time stability estimate as an additional bounded forcing term alongside A_1, A_2 in Theorem 3, in the same way that the bounded disturbance $d_i(t)$ was incorporated into the UAV model of Section 4.4. We do not model sensor spoofing, communication-channel attacks (e.g., delayed or corrupted packet delivery), or adversarial perturbations to the activation functions f_i in this experiment; extending the threat model to these cases is left for future work (Section 5).

Figure 8 shows the time evolution of the individual neural states. Before the disturbance is introduced, all states remain close to the equilibrium, indicating normal and secure operation. Once the attack becomes active, the states exhibit visible oscillatory deviations caused by the delayed and nonlinear interactions within the network. Importantly, these deviations remain bounded and do not grow uncontrollably. After the disturbance is removed, the trajectories rapidly return toward the equilibrium and settle within a short time interval and provide clear visual evidence of the settling behaviour discussed following Definition 3.

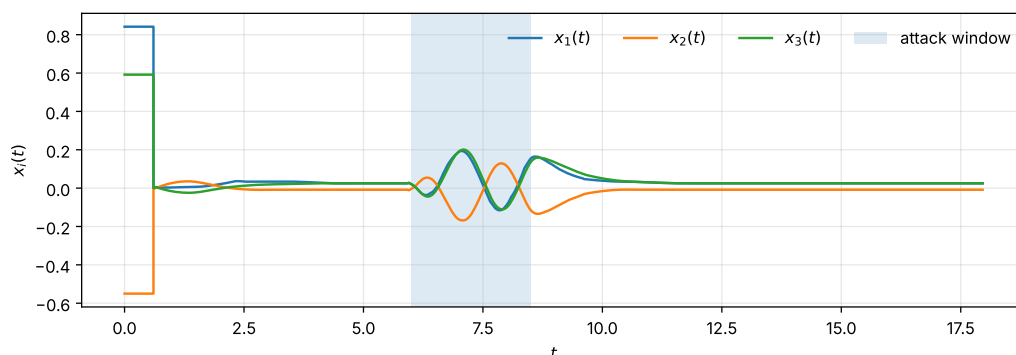


Figure 8. State trajectories under a bounded cyber disturbance (shaded interval).

To further quantify this behavior, Figure 9 depicts the evolution of the security energy $\|x(t)\|_2$. A temporary increase in the energy level is observed during the attack window. This reflects the system's response to the injected disturbance. However, once the attack ends, the energy decays sharply and returns to a small neighborhood of zero in finite time. This rapid decay highlights the strong stabilizing effect guaranteed by Theorem 3, even in the presence of communication delays and fractional memory effects.

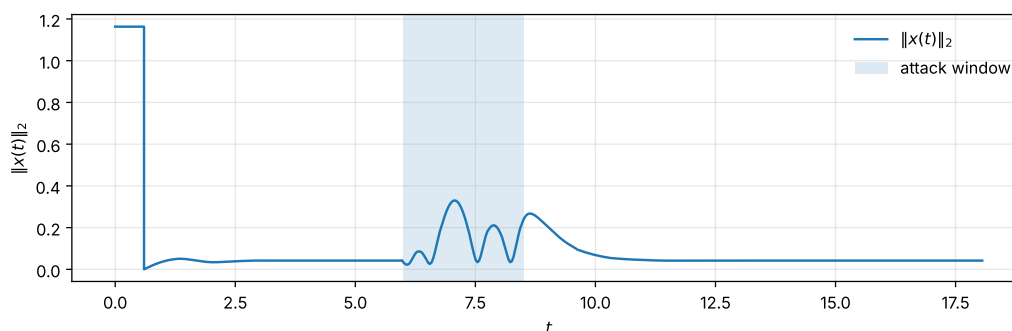


Figure 9. Evolution of the security energy $\|x(t)\|_2$ showing finite-time recovery.

These results show that the proposed fractional-order delayed neural system is capable of suppressing cyber disturbances and restoring secure operation within a guaranteed finite time. The observed dynamics are fully consistent with the theoretical finite-time stability conditions derived in Theorem 3 and emphasize the suitability of the model for security-sensitive cyber-physical applications.

Interpretation and numerical illustration

The finite-time stability behavior observed in the numerical simulations in this section is a direct consequence of the sufficient condition established in Theorem 3. Recall that the fractional-order delayed neural system (4.2) is governed by the Caputo–GCF derivative ${}^C D_{t_0}^{\sigma, \nu+\zeta}$, which incorporates both nonlocal memory effects and delayed neural interactions. These two features are known to complicate stability analysis, particularly when classical asymptotic arguments are no longer sufficient.

The core of Theorem 3 lies in the explicit inequality (4.14), where the combined influence of the intrinsic decay rates c_i , the synaptic coupling strengths A and B , and the delay τ is carefully balanced through the exponential term

$$\exp\left(t_f^{\nu+\zeta} + M\left(A_1^2(t_f^{\nu+\zeta} - t_0^{\nu+\zeta}) + A_2^2((t_f + \tau)^{\nu+\zeta} - (t_0 + \tau)^{\nu+\zeta})\right)\right),$$

with the constant M explicitly defined in terms of the fractional order σ and the Gamma function. This structure reflects how fractional memory and time delay jointly affect the transient growth of solutions. When condition (4.14) is satisfied, the growth induced by nonlinear activation functions and delayed feedback is dominated by the intrinsic decay, which ensures that the state trajectories enter a prescribed neighborhood of the equilibrium within a finite settling time.

In the numerical experiment, the system is intentionally subjected to a bounded external disturbance over a finite interval and mimics a cyber attack or abnormal input in a cyber-physical

environment. During this interval, the nonlinear terms $Af(x(t))$ and $Bf(x(t - \tau))$ temporarily amplify the state deviations, which is consistent with the theoretical role of the coupling matrices in (4.14). However, once the disturbance is removed, the decay terms $-c_i x_i(t)$, together with the finite-time stability mechanism guaranteed by Theorem 3, keep the deviation bounded and, as observed numerically, bring it below the reference threshold within a short and clearly observable time span.

The evolution of the security energy $\|x(t)\|_2$ provides a compact quantitative confirmation of this process. Although the energy increases during the disturbance window, it rapidly decreases afterward and reaches a small neighborhood of zero in finite time, rather than merely converging asymptotically. This behavior aligns exactly with the theoretical conclusion that system (4.2) is finite-time stable with respect to $\{t_0, \delta, \varepsilon, \Theta, \tau\}$.

From a practical perspective, the benefit of this result is significant. Finite-time stability guarantees robustness against delayed nonlinear interactions in the sense of Definition 3; the short recovery time observed numerically in Figure 9 is an additional empirical property of this parameter regime, not a general guarantee of the theorem. In security-sensitive cyber–physical systems, this translates into faster detection, quicker suppression of abnormal behavior, and reliable restoration of normal operation, all of which are supported by the analytical framework provided in Theorem 3.

As in Sections 4.3 and 4.4, this simulation illustrates a parameter regime consistent with condition (4.14) of Theorem 3; it is a single bounded-disturbance scenario and should not be read as demonstrating finite-time stabilization or recovery guarantees under general cyber-attack models.

5. Conclusions and future research

This paper introduced a new Caputo–GCF derivative that combines the classical Caputo-type treatment of initial conditions with an independently adjustable structural memory parameter $\nu + \zeta$. The proposed operator was shown to be linear, bounded, and continuous from $C^1([a, b], \mathbb{R})$ into $C([a, b], \mathbb{R})$. Moreover, it recovers the classical Caputo derivative when $\nu + \zeta = 1$, the corresponding GCF derivative when $u(a) = 0$, and is connected with the Riemann–Liouville and Hadamard operators through appropriate special or limiting cases. Thus, the formulation separates the fractional order σ from the structural memory scaling $\nu + \zeta$ and provide an additional degree of freedom for representing nonlocal memory.

For the proposed delayed fractional neural network, the existence and uniqueness of the equilibrium point were established under the Lipschitz condition **(H1)** and the damping-dominance condition **(H2)**. Two explicit sufficient finite-time stability criteria were then derived. The first criterion, obtained through Hölder’s inequality, covers the fractional-order region $0 < \sigma \leq \frac{1}{2}$, whereas the second criterion, derived using the Cauchy–Schwarz inequality, applies to $\frac{1}{2} < \sigma < 1$. These criteria explicitly quantify the effects of the initial perturbation δ , admissible bound ε , fractional order σ , structural memory parameter $\nu + \zeta$, delay τ , decay coefficients, and the instantaneous and delayed coupling quantities A_1 and A_2 . The results therefore provide computable conditions under which the state trajectory remains inside a prescribed neighbourhood of the unique equilibrium over a specified finite-time interval.

The numerical examples and application studies support these theoretical results without exceeding the prescribed bounds: both examples remained within $\varepsilon = 2$ across the tested parameter grid, and the operator comparison, ablation study, UAV control example, and cyber-disturbance simulation each showed behaviour consistent with Theorems 2 and 3 (Tables 4–12). Two consistent

practical advantages emerge. First, tuning $\nu + \zeta$ independently of σ gave faster empirical settling and smaller final error than the classical Caputo and GCF operators in the direct comparison (Table 7), and improved recall, F1-score, and AUROC over related reduced fractional-kernel models in the CATS ablation study (Table 8), consistent with the intuition that smaller σ preserves longer memory of past anomalies. Second, the framework remains applicable once realistic constraints, sensor delay and actuator saturation in the UAV example, a bounded external attack in the secure-neural-network example, are added, though the UAV comparison also shows a genuine trade-off: the integer-order controller reached the settling threshold faster (Table 12), reflecting the cost of retaining fractional memory.

Limitations. The obtained stability criteria are sufficient rather than necessary and may be conservative because they rely on contraction, Hölder, Cauchy–Schwarz, and Gronwall-type estimates – indeed, Table 6 shows explicitly that the certified horizon on which condition (4.5) can be verified is considerably shorter than the simulated settling time, and the remark following Theorem 2 identifies a specific source of conservativeness in how the decay rates c_i enter A_1 . The analysis is restricted to globally Lipschitz activation functions, a constant delay, deterministic dynamics, bounded disturbances, and fractional orders in the interval $0 < \sigma < 1$. In addition, the application results are based mainly on numerical simulations and one benchmark time-series dataset. The UAV study was not validated on physical hardware, and the cyber-disturbance experiment considered a bounded attack model rather than multiple realistic attack mechanisms. The slower UAV response relative to the integer-order controller also indicates a practical trade-off between fractional-memory retention and convergence speed.

Reproducibility. The numerical scheme used throughout Section 4 (product-trapezoidal quadrature for the GCF integral, linear interpolation for the delayed term) is described in the “Numerical implementation” paragraph of Section 4.1, and the system matrices and parameters for Examples 1-2, the UAV example, and the equilibrium-point iteration are given explicitly in the text. The data split, architecture, training settings, threshold-selection rule, and random seeds used for the CATS ablation study are likewise given in full in the “Experimental setup” paragraph of Section 4.2. The authors will provide the implementation code (or a link to it) upon reasonable request.

Future research. Future work will extend the Caputo–GCF framework to stochastic, impulsive, event-triggered, and switching neural systems. Time-varying, distributed, and state-dependent delays, as well as non-Lipschitz and discontinuous activation functions, will also be investigated. Less conservative stability criteria may be developed through refined Lyapunov–Krasovskii functionals, generalized integral inequalities, and optimization-based parameter selection. Another important direction is the data-driven estimation of σ and $\nu + \zeta$ instead of selecting them manually. Finally, the proposed model should be evaluated on additional anomaly-detection datasets, compared with contemporary deep-learning baselines, and implemented on real-time UAV and cyber–physical test platforms to assess its computational cost, scalability, robustness, and practical recovery performance.

Author contributions

Abeer Aljohani: Conceptualization, methodology, formal analysis, software, validation, investigation, visualization, writing—original draft, writing—review & editing. The author has read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The author declares she has not used AI tools in the creation of this article.

Funding

This scientific paper is derived from a research grant funded by Taibah University, Madinah, Kingdom of Saudi Arabia, under Grant No. 15-448-1206.

Conflict of interest

The author declares no conflicts of interest.

References

1. T. Zhang, L. Wang, Y. Zhang, J. Deng, Finite-time stability for fractional-order fuzzy neural network with mixed delays and inertial terms, *AIMS Math.*, **9** (2024), 19176–19194. <https://doi.org/10.3934/math.2024935>
2. T. Hu, Z. He, X. Zhang, S. Zhong, Finite-time stability for fractional-order complex-valued neural networks with time delay, *Appl. Math. Comput.*, **365** (2020), 124715. <https://doi.org/10.1016/j.amc.2019.124715>
3. N. T. Thanh, P. Niamsup, V. N. Phat, New results on finite-time stability of fractional-order neural networks with time-varying delay, *Neural Comput. Appl.*, **33** (2021), 17489–17496. <https://doi.org/10.1007/s00521-021-06339-2>
4. T. Hamadneh, A. Hioual, O. Alsayyed, Y. A. Al-Khassawneh, A. Al-Husban, A. Ouannas, Finite time stability results for neural networks described by variable-order fractional difference equations, *Fractal Fract.*, **7** (2023), 616. <https://doi.org/10.3390/fractalfract7080616>
5. A. Ben Makhlouf, D. Baleanu, Finite time stability of fractional order systems of neutral type, *Fractal Fract.*, **6** (2022), 289. <https://doi.org/10.3390/fractalfract6060289>
6. S. K. Panda, T. Abdeljawad, A. M. Nagy, On uniform stability and numerical simulations of complex valued neural networks involving generalized Caputo fractional order, *Sci. Rep.*, **14** (2024), 4073. <https://doi.org/10.1038/s41598-024-53670-4>
7. S. Jose, V. Parthiban, Finite-time synchronization of fractional order neural networks via sampled data control with time delay, *J. Math. Comput. Sci.*, **35** (2024), 374–387. <https://doi.org/10.22436/jmcs.035.04.01>

8. M. Rhaima, L. Chen, L. Mchiri, A. Ben Makhlouf, Finite time stability of ρ -Caputo neural network systems, *Asian J. Control*, **27** (2025), 459–469. <https://doi.org/10.1002/asjc.3447>
9. G. Palanisamy, U. Kandasamy, F. A. Rihan, S. Ben Said, Fractional-order stochastic delayed neural networks with impulses: mean square finite-time contractive synchronization, *Sci. Rep.*, **16** (2026), 1999. <https://doi.org/10.1038/s41598-025-31768-7>
10. E. Demirci, F. Karakoç, A. Kütahyalıoğlu, Mittag-Leffler stability of neural networks with Caputo–Hadamard fractional derivative, *Math. Methods Appl. Sci.*, **47** (2024), 10091–10100. <https://doi.org/10.1002/mma.10111>
11. H. L. Li, C. Hu, L. Zhang, H. Jiang, J. Cao, Complete and finite-time synchronization of fractional-order fuzzy neural networks via nonlinear feedback control, *Fuzzy Sets Syst.*, **443** (2022), 50–69. <https://doi.org/10.1016/j.fss.2021.11.004>
12. R. Agarwal, S. Hristova, D. O'Regan, Stability of generalized proportional Caputo fractional differential equations by Lyapunov functions, *Fractal Fract.*, **6** (2022), 34. <https://doi.org/10.3390/fractalfract6010034>
13. D. Băbler, T. Kortus, G. Gühring, Unsupervised anomaly detection in multivariate time series with online evolving spiking neural networks, *Mach. Learn.*, **111** (2022), 1377–1408. <https://doi.org/10.1007/s10994-022-06129-4>
14. L. Kirichenko, Y. Koval, S. Yakovlev, D. Chumachenko, Anomaly detection in fractal time series with LSTM autoencoders, *Mathematics*, **12** (2024), 3079. <https://doi.org/10.3390/math12193079>
15. F. Wang, Y. Jiang, R. Zhang, A. Wei, J. Xie, X. Pang, A survey of deep anomaly detection in multivariate time series: taxonomy, applications, and directions, *Sensors*, **25** (2025), 190. <https://doi.org/10.3390/s25010190>
16. F. Du, J. G. Lu, Finite-time stability of fractional-order fuzzy cellular neural networks with time delays, *Fuzzy Sets Syst.*, **438** (2022), 107–120. <https://doi.org/10.1016/j.fss.2021.08.011>
17. F. Du, J. G. Lu, Adaptive finite-time synchronization of fractional-order delayed fuzzy cellular neural networks, *Fuzzy Sets Syst.*, **466** (2023), 108480. <https://doi.org/10.1016/j.fss.2023.02.001>
18. X. Li, X. Liu, S. Zhang, New criteria on the finite-time stability of fractional-order BAM neural networks with time delay, *Neural Comput. Appl.*, **34** (2022), 4501–4517. <https://doi.org/10.1007/s00521-021-06605-3>
19. Z. Yang, J. Zhang, Y. Niu, Finite-time stability of fractional-order bidirectional associative memory neural networks with mixed time-varying delays, *J. Appl. Math. Comput.*, **63** (2020), 501–522. <https://doi.org/10.1007/s12190-020-01327-6>
20. O. Naifar, A. Ben Makhlouf, L. Mchiri, M. Rhaima, Finite time stability for Hadamard fractional-order systems, *Ain Shams Eng. J.*, **16** (2025), 103263. <https://doi.org/10.1016/j.asej.2025.103263>
21. A. Al-Jaser, F. Alharbi, D. Chalishajar, B. Qaraad, Functional differential equations with non-canonical operator: oscillatory features of solutions, *Axioms*, **14** (2025), 588. <https://doi.org/10.3390/axioms14080588>
22. C. Anusha, C. Ravichandran, F. Aljuaydi, K. S. Nisar, D. Chalishajar, Stability and solutions of adjoint nonlinear impulsive neutral mixed integral equations in dynamic systems, *J. Appl. Math. Comput.*, **71** (2025), 7521–7541. <https://doi.org/10.1007/s12190-025-02579-w>

23. D. Chalishajar, D. Kasinathan, R. Kasinathan, R. Kasinathan, Viscoelastic Kelvin–Voigt model on Ulam–Hyers’s stability and T-controllability for a coupled integro fractional stochastic systems with integral boundary conditions via integral contractors, *Chaos Soliton. Fract.*, **191** (2025), 115785. <https://doi.org/10.1016/j.chaos.2024.115785>
24. D. Chalishajar, D. Kasinathan, R. Kasinathan, R. Kasinathan, Ulam–Hyers–Rassias stability of Hilfer fractional stochastic impulsive differential equations with non-local condition via time-changed Brownian motion followed by the currency options pricing model, *Chaos Soliton. Fract.*, **197** (2025), 116468. <https://doi.org/10.1016/j.chaos.2025.116468>
25. Z. Yan, Y. Song, J. H. Park, Finite-time stability and stabilization for stochastic Markov jump systems with mode-dependent time delays, *ISA Trans.*, **68** (2017), 141–149. <https://doi.org/10.1016/j.isatra.2017.01.018>
26. Z. Yan, G. Chang, Z. Xing, R. Xu, J. Cheng, Annular finite-time stability and stabilization in probability for nonlinear stochastic Markov jump systems with impulsive effects, *Int. J. Robust Nonlinear Control*, **36** (2025), 245–255. <https://doi.org/10.1002/rnc.70137>
27. B. Gopinath, A benign and malignant pattern identification in cytopathological images of thyroid nodules using Gabor filter and neural networks, *Asian J. Conver. Technol.*, **4** (2018), 414.
28. Q. X. Zhu, J. D. Cao, Mean-square exponential input-to-state stability of stochastic delayed neural networks, *Neurocomputing*, **131** (2014), 157–163. <https://doi.org/10.1016/j.neucom.2013.10.029>
29. C. Xu, L. Chen, Effect of leakage delay on the almost periodic solutions of fuzzy cellular neural networks, *J. Exp. Theor. Artif. Intell.*, **30** (2018), 993–1011. <https://doi.org/10.1080/0952813X.2018.1509895>
30. T. U. Khan, M. A. Khan, Generalized conformable fractional operators, *J. Comput. Appl. Math.*, **346** (2019), 378–389. <https://doi.org/10.1016/j.cam.2018.07.018>
31. Y. Ke, C. Miao, Stability analysis of fractional-order Cohen–Grossberg neural networks with time delay, *Int. J. Comput. Math.*, **92** (2015), 1102–1113. <https://doi.org/10.1080/00207160.2014.935734>
32. M. Kuczma, *An introduction to the theory of functional equations and inequalities: Cauchy’s equation and Jensen’s inequality*, Basel: Birkhäuser, 2009. <https://doi.org/10.1007/978-3-7643-8749-5>
33. G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities*, Cambridge: Cambridge University Press, 1952.
34. P. Fleith, *Controlled anomalies time-series dataset*, Hugging Face, 2024. Available from: <https://huggingface.co/datasets/patrickfleith/controlled-anomalies-time-series-dataset>.



AIMS Press

©2026 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>)