ACCSCIENCE PUBLISHING
An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.x, No.x, pp.xx-xx (2026)
https://doi.org/10.36922/IJOCTA026270144
IJOCTA
RESEARCH ARTICLE

Numerical simulation and neural-network approximation for solving impulsive stochastic delay equations

Rahman Ullah Khan[1*], Ioannis Argyros[2], Manar A. Alqudah[3], and Thabet Abdeljawad[4*]

[1]Department of Mathematics, Quaid-i-Azam University, Islamabad, Pakistan
[2]Department of Computing and Mathematical Sciences, Cameron University, Lawton, OK, USA
[3]Department of Mathematical Sciences, College of Science, Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia
[4]Department of Fundamental Sciences, Faculty of Engineering and Architecture, Istanbul Gelisim University, Avcilar-Istanbul, Turkey
rkhan@math.qau.edu.pk, ioannisa@cameron.edu, maalqudah@pnu.edu.sa, tabdeljawad@gelisim.edu.tr

ARTICLE INFO

Article History:
Received: July 1, 2026
Revised: July 29, 2026
Accepted: August 3, 2026
Published Online: November 20, 2026

Keywords:
ψ-Caputo fractional operator
stochastic delay equations
random impulses
Hyers–Ulam stability
neural-network approximation

MSC 2020:
34A08; 34K37; 60H10; 60H35; 65C30

ABSTRACT

This study investigates random impulsive stochastic functional differential equations with finite delay using the ψ-Caputo operator. The stochastic basis is specified through an enlarged filtration that contains the impulse variables and preserves the Wiener property, and solutions are formulated in a piecewise-continuous L^p path space. Under Lipschitz, growth, compact-range, and random-impulse conditions, Krasnoselskii's theorem establishes existence; uniqueness follows from a constructive local-partition rule. The analysis also gives a finite-time moment estimate, continuous dependence on the initial history, mean pth-moment Hyers–Ulam stability, and Hyers–Ulam–Rassias stability with explicit admissible weight functions. The computational section develops a ψ-Caputo L1 discretization, a fractional Euler–Maruyama update, one-sided delay interpolation, graded meshes for the initial singularity, and a stochastic residual-based neural surrogate. Numerical tests report uniform-versus-graded convergence, strong and weak Monte Carlo diagnostics, delay-interpolation sensitivity, impulse effects, and neural approximations of the empirical mean and second moment.

CC BY

1. Introduction

The stability problem initiated by Ulam asks whether a function that satisfies an equation only approximately must lie near a genuine solution. Hyers answered this question for additive mappings, and later work by Aoki, Bourgin and Rassias extended the idea in several directions.[1–5] The resulting Hyers–Ulam framework is now a standard qualitative tool for differential, integral, stochastic, and functional equations.[6,7]

Delay equations are appropriate when the evolution at time t is influenced by a finite portion of the past rather than by the instantaneous state alone. For a prescribed delay length $\tau > 0$, the memory segment may be written as

$$x^{\langle t\rangle} = \{x(t+\theta) : -\tau \le \theta \le 0\}.$$

For first-order delay and nonlinear functional models, Hyers–Ulam stability is commonly established by combining fixed point methods, Gronwall inequalities and compactness arguments.[6,7]

Impulsive systems model processes in which the trajectory is altered by sudden jumps. In applications, these jump instants may themselves be random, which leads to stochastic equations with random impulse times. A standard formulation is

$$dX(t) = F(t, X^{\langle t\rangle})\,dt + G(t, X^{\langle t\rangle})\,d\mathcal{W}(t), \quad t \neq \zeta_k,$$
$$X(\zeta_k) = J_k(\varrho_k)X(\zeta_k^-), \qquad k = 1, 2, \ldots,$$

where the instants ζ_k arise from independent waiting times ϱ_k, and $\mathcal{W}(t)$ denotes a Wiener process. Models of this type arise in biological dynamics, financial systems, control theory, communication

*Corresponding Author

networks, and environments subject to random shocks.[8–11]

Classical stochastic differential equations provide the basic probabilistic setting for diffusion processes and moments estimates.[8] When delay is present, the state variable is no longer only $x(t)$; it must include the segment $x^{\langle t\rangle}$, which leads to stochastic functional equations. Random impulsive variants were introduced to describe abrupt state changes at random instants, and their well-posedness is typically obtained from Lipschitz hypotheses together with bounds on products of impulse operators.[9,10]

Fractional equations generalize integer-order dynamics through kernels that retain past information. Systematic accounts of the Caputo, Riemann–Liouville, and related fractional operators can be found in.[12–15] The Caputo derivative taken with respect to another function, in the sense used by Almeida,[16] is particularly relevant here because the scale function ψ modifies the memory clock while leaving the phase space unchanged. Fixed-point and compactness techniques have also been applied to impulsive fractional equations and fractional functional systems.[10,11,17]

The corresponding stability theory remains tied to Ulam's original question and to Hyers' result. For impulsive and stochastic equations, Ulam–Hyers estimates are often derived from residual inequalities, Gronwall-type bounds and stochastic moment estimates.[6,7] On the computational side, fractional equations are frequently discretized by convolution quadrature, predictor–corrector procedures and L1 schemes.[18,19] Neural residual methods, including physics-informed neural networks, provide an additional approximation strategy when the loss functional penalizes the governing equation directly.[20,21]

Recent work confirms that the topic is active but also highights the gap addressed here. Benzarouala and Tunç studied Hyers–Ulam–Rassias stability for Caputo fractional delay equations, while Tang, Luo, Chen and Gao treated Ulam–Hyers stability for fractional stochastic delay equations with non-instantaneous impulses.[22,23] Numerical methods for ψ-Caputo problems have also progressed through operational and wavelet techniques.[24] In parallel, neural fractional differential equations and fractional physics-informed methods have become useful approximation tools for memory-dependent models.[25,26]

Several recent computational studies are especially relevant to the numerical context. Localized radial-basis collocation reduces the conditioning and assembly cost of deterministic fractional integral equations;[27] two-dimensional Haar operational matrices provide sparse representations for fractional Volterra equations;[28] generalized Lerch polynomials address nonlinear two-dimensional fractional optimal-control problems;[29] non-polynomial splines have been used for the PHI-Four and Allen–Cahn equations;[30] and a fractional clique-polynomial neural network has recently been applied to a kidney-function model.[31] These methods are valuable, but they do not simultaneously treat an adapted stochastic path, random jump times, finite delay, a ψ-dependent memory kernel, and moment-based stability. The present paper therefore develops a framework in which the analytical assumptions and the computational residual use the same impulsive stochastic-delay model.

The present work is motivated by models in which finite delay, randomly timed impulses, stochastic forcing and fractional memory act at simultaneously. Integer-order random impulsive stochastic delay equations do not preserve long-range memory in the derivative, whereas many fractional delay models omit random impulse instants, Itô noise, or a compatible neural approximation component. The ψ-Caputo setting is selected because it admits classical initial histories and because ψ changes the operational memory clock without changing the state space; setting $\psi(t) = t$ recovers the ordinary Caputo derivative. This flexibility permits single theory to cover standard Caputo dynamics and nonuniform memory scales.

Hyers–Ulam stability is included to answer a practical robustness question: a trajectory produced by measurement, model reduction, time discretization, or neural training generally satisfies the equation only up to a residual. A uniform residual bound leads to Hyers–Ulam stability, whereas a time-dependent bound $\varepsilon\omega(t)$ leads to Hyers–Ulam–Rassias stability. Physically, the estimates determine whether these perturbations remain controlled after fractional memory, delay, stochastic forcing, and jump amplification have accumulated.

The main new elements are the following. The random impulsive stochastic delay equation is recast with a generalized ${}^{C}_{\psi}\mathfrak{D}^{\alpha}$ derivative, with the ordinary Caputo model obtained as a limiting choice of the scale. A mild solution representation is then constructed from the associated ψ-Volterra kernel. Existence is obtained through Krasnoselskii's fixed point theorem, and uniqueness is proved by a contraction argument on short subintervals. Mean pth moment Hyers–Ulam stability is quantified by a Mittag–Leffler coefficient depending explicitly on α and ψ. Finally, the theory is accompanied by an L1 finite-difference discretization and by a

neural residual approximation that demonstrate the computational use of the model.

Figure 1 summarizes the structure of the present work. It was added to make the analytical and computational route explicit: model formulation, ψ-Volterra transformation, qualitative analysis, strengthened estimates, numerical discretization, neural surrogate construction and validation.

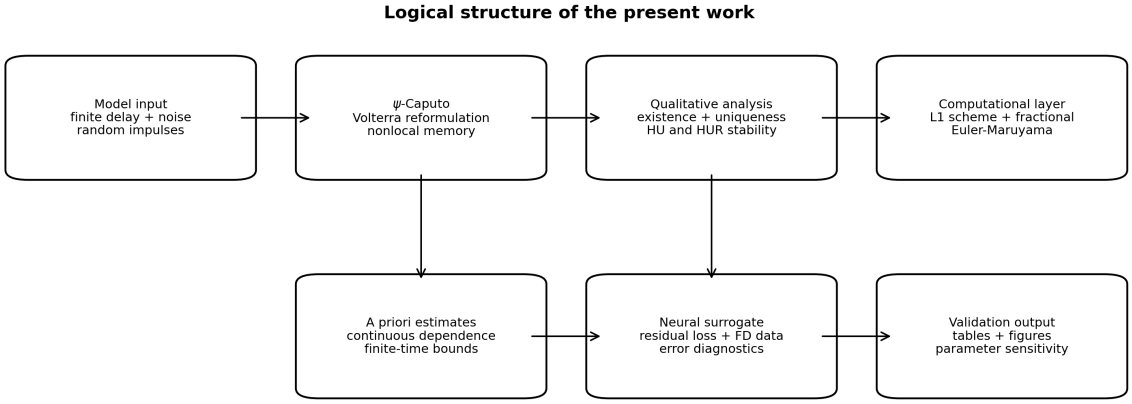


Figure 1. Flow chart of the analytical and computational framework used in the present work

Accordingly, the central object studied in this paper is the fractional random impulsive stochastic functional equation with finite delay given by

$$\begin{aligned} {}^{C}_{\psi}\mathfrak{D}^{\alpha}_{t_0+}x(t) &= f(t,x^{\langle t\rangle}) + g(t,x^{\langle t\rangle})\dot{\mathcal{W}}(t), \\ &\qquad t\in[t_0,T], \quad t\neq\zeta_k, \\ x(\zeta_k) &= J_k(\varrho_k)x(\zeta_k^-), \qquad k=1,2,\ldots, \\ x^{\langle t_0\rangle} &= \phi = \{\phi(\theta) : -\tau\le\theta\le 0\}. \end{aligned} \tag{1}$$

Here, ${}^{C}D^{\alpha,\psi}_{t_0^+}$ denotes the ψ-Caputo fractional derivative of order $\alpha\in(0,1)$. The function $x(t)$ is the unknown stochastic state, while $x^{\langle t\rangle}$ represents the delayed history segment. Moreover, $\dot{\mathcal{W}}(t)$ denotes Gaussian white noise. The functions f and g describe the drift and diffusion terms, respectively, whereas ζ_k, J_k, ϱ_k, and ϕ denote the impulse instants, impulse operator, random impulse amplitude, and prescribed initial history function, respectively.

The differential expression involving $\dot{\mathcal{W}}$ is only a shorthand. Its precise interpretation is the mild Volterra form in **Definition 3**, where the Itô integral treats the stochastic term and the deterministic ψ-fractional kernel carries the memory effect.

The remainder of the paper is arranged as follows. Section 2 sets the probabilistic notation, the ψ-Caputo operator and the mild-solution concept. Section 3 contains the existence and uniqueness analysis. Section 4 proves Hyers–Ulam and Hyers–Ulam–Rassias stability. Section 5 verifies the hypotheses in a model example. Section 6 gives the finite-difference, stochastic simulation and neural-network approximations with additional tables and figures. Section 7 presents the conclusion.

2. Preliminaries

This section fixes the notation and records the auxiliary material used later. The probabilistic, fractional-calculus and fixed-point ingredients follow standard sources.[8,14–16,32] Let $(\mathcal{S},\mathscr{A},\mathsf{P})$ be a complete probability space and let $\{\mathscr{A}_t^W\}_{t\ge0}$ be the augmented natural filtration of an m-dimensional Wiener process $\mathcal{W}(t)$. Define the impulse sigma-field

$$\mathscr{I} = \sigma\big((\varrho_k, J_k(\varrho_k)) : k\ge 1\big),$$

and assume that $\mathscr{I}$ is independent of $\sigma(\mathcal{W}(t) : t\ge 0)$. We work with the enlarged filtration

$$\mathscr{G}_t = \mathscr{A}_t^W \vee \mathscr{I},$$

completed in the usual way. Independence implies that $\mathcal{W}$ remains a Wiener process relative to $\{\mathscr{G}_t\}$; the waiting times, impulse matrices, and the recursively defined ζ_k are $\mathscr{G}_0$-measurable, so every ζ_k is a $\{\mathscr{G}_t\}$-stopping time. Thus the impulse schedule is random but independent of the Brownian increments.

For a delay length $0<\tau<\infty$, set

$$\mathscr{H}_\tau = C([-\tau,0], L^p(\mathcal{S},\mathscr{A},\mathsf{P};\mathbb{R}^d)), \qquad p\ge 2,$$

with norm

$$\|\varphi\|_{\mathscr{H}_\tau} = \sup_{\theta\in[-\tau,0]} \left(\mathbb{E}\,\|\varphi(\theta)\|^p\right)^{1/p}.$$

For $x : [t_0-\tau, T] \to L^p(\mathcal{S},\mathbb{R}^d)$, define

$$x^{\langle t\rangle}(\theta) = x(t+\theta), \qquad -\tau \le \theta \le 0.$$

Let $\psi \in C^1([t_0,T],\mathbb{R})$ be strictly increasing and satisfy

$$\psi'(t) > 0, \qquad t \in [t_0,T].$$

In what follows the order of differentiation is restricted by

$$\frac{1}{2} < \alpha < 1.$$

The lower bound on α is needed because the stochastic Volterra kernel must be square integrable near its singular endpoint in the present Itô–BDG framework. This is a structural limitation of the proof, not a claim that orders $0 < \alpha \le 1/2$ are impossible. Such orders require a different integration setting, for example weighted stochastic-convolution spaces, compensated kernels, or rough-path/fractional-noise norms, and are left for future work.

Definition 1 (ψ-fractional integral)**.** *Following the fractional integral with respect to another function,*[15,16] *for $h \in L^1([t_0,T],\mathbb{R}^d)$ the left-sided ${}_\psi\mathfrak{I}^\alpha$ integral of order α is given by*

$$({}_\psi\mathfrak{I}^\alpha_{t_0+}h)(t) = \frac{1}{\Gamma(\alpha)} \int_{t_0}^t \psi'(s)(\psi(t)-\psi(s))^{\alpha-1}h(s)\,ds. \tag{2}$$

Definition 2 (ψ-Caputo fractional derivative)**.** *The derivative used here is the Caputo derivative with respect to ψ in the sense of Almeida.*[16] *If x is absolutely continuous relative to the ψ-scale, its generalized ${}^C_\psi\mathfrak{D}^\alpha$ derivative of order $\alpha \in (0,1)$ is*

$$({}^C_\psi\mathfrak{D}^\alpha_{t_0+}x)(t) = \frac{1}{\Gamma(1-\alpha)} \int_{t_0}^t \psi'(s)(\psi(t)-\psi(s))^{-\alpha}\left(\frac{1}{\psi'(s)}\frac{d}{ds}x(s)\right)ds. \tag{3}$$

The Volterra form of fractional integral equations[14,15] motivates the kernel notation

$$\mathscr{G}_{\alpha,\psi}(t,s) = \frac{\psi'(s)(\psi(t)-\psi(s))^{\alpha-1}}{\Gamma(\alpha)}, \qquad t_0 \le s < t \le T. \tag{4}$$

Figure 2 shows the dependence of the singular memory weight on the order α near the present time. The same kernel appears in the mild formulation, in compactness estimates and in the quadrature weights used later.

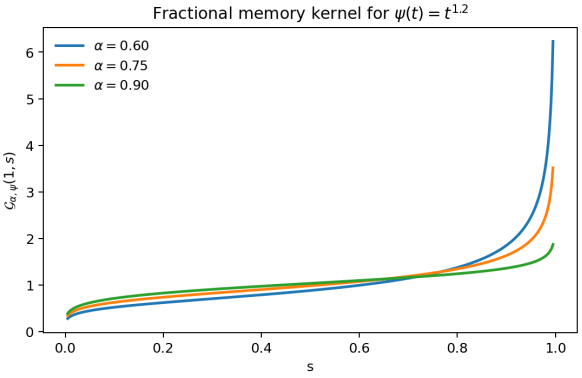


Figure 2. ψ-Caputo memory weight for several values of α with $\psi(t) = t^{1.2}$ and $t = 1$

For later estimates, define the deterministic kernel mass and the stochastic square-kernel mass by

$$\mathfrak{L}_{\alpha,\psi}(a,b) = \int_a^b \mathscr{G}_{\alpha,\psi}(b,s)\,ds = \frac{(\psi(b)-\psi(a))^\alpha}{\Gamma(\alpha+1)},$$
$$\mathfrak{Q}_{\alpha,\psi}(a,b) = \int_a^b \mathscr{G}_{\alpha,\psi}(b,s)^2\,ds,$$
$$\mathfrak{T}_{\alpha,\psi}(a,b) = \mathfrak{Q}_{\alpha,\psi}(a,b)^{p/2}. \tag{5}$$

Writing

$$\psi_{\min} = \inf_{t\in[t_0,T]} \psi'(t) > 0, \qquad \psi_{\max} = \sup_{t\in[t_0,T]} \psi'(t),$$

we have the explicit estimate

$$\mathfrak{Q}_{\alpha,\psi}(a,b) \le \frac{\psi_{\max}(\psi(b)-\psi(a))^{2\alpha-1}}{(2\alpha-1)\Gamma(\alpha)^2}. \tag{6}$$

The corresponding global quantities are $\mathfrak{L}_{\alpha,\psi} = \mathfrak{L}_{\alpha,\psi}(t_0,T)$, $\mathfrak{Q}_{\alpha,\psi} = \mathfrak{Q}_{\alpha,\psi}(t_0,T)$, and $\mathfrak{T}_{\alpha,\psi} = \mathfrak{T}_{\alpha,\psi}(t_0,T)$. The symbols have one role throughout: $\mathfrak{L}$ controls deterministic convolution, $\mathfrak{Q}$ is the square-kernel integral, $\mathfrak{T} = \mathfrak{Q}^{p/2}$ enters BDG estimates, and $\mathfrak{A}_{\alpha,\psi}$ introduced in Section 4 is the combined stability coefficient.

Let $\{\varrho_k\}_{k\ge1}$ be mutually independent positive random variables such that $\varrho_k : \mathcal{S} \to \mathcal{D}_k = (0,d_k)$, and assume that the pairs $(\varrho_k, J_k(\varrho_k))$ are mutually independent and independent of $\mathcal{W}$. The impulse times are introduced recursively by

$$\zeta_0 = t_0, \qquad \zeta_k = \zeta_{k-1} + \varrho_k, \qquad k = 1,2,\dots.$$

We assume $\zeta_k \uparrow \infty$ almost surely, so only finitely many jumps occur on $[t_0,T]$ for almost every outcome. For each impulse,

$$J_k : \mathcal{D}_k \to \mathbb{R}^{d\times d}$$

is Borel measurable and $J_k(\varrho_k)$ is $\mathscr{I}$-measurable. The linear multiplicative form is retained because it gives explicit products $\Pi_{i,k}$. A nonlinear impulse $I_k(\varrho_k,u)$ can be included when it is measurable and satisfies uniform L^p Lipschitz and growth

bounds; in that case the product estimate in assumption (A2) is replaced by the corresponding recursive Lipschitz constants.

Definition 3 (Mild solution)**.** *The following notion is based on the usual mild-solution approach for stochastic functional equations with impulses.*[8–10] *An $\mathbb{R}^d$-valued stochastic process $x(t)$ defined on $[t_0-\tau, T]$ is a mild solution of (1) if $x^{\langle t_0\rangle} = \phi$, it is $\mathscr{G}_t$-adapted, it is continuous in L^p on each open interval (ζ_k, ζ_{k+1}), and it admits finite left limits at the jump times. It satisfies*

$$x(\zeta_k) = J_k(\varrho_k)x(\zeta_k^-),$$

and the Volterra identity below. For compactness, write

$$\Pi_{i,k} = \prod_{j=i}^{k} J_j(\varrho_j), \qquad \Pi_{1,k} = \prod_{i=1}^{k} J_i(\varrho_i).$$

Set $f_s^x = f(s, x^{\langle s\rangle})$ and $g_s^x = g(s, x^{\langle s\rangle})$. The impulse relation holds, and for every $t \in [t_0, T]$ one has

$$\begin{aligned} x(t) &= \sum_{k=0}^{\infty} \mathscr{U}_k(t)\, \mathbf{1}_{(\zeta_k, \zeta_{k+1}]}(t) \quad a.s., \\ \mathscr{U}_k(t) &= \Pi_{1,k}\phi(0) \\ &\quad + \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) f_s^x \, ds + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) f_s^x \, ds \\ &\quad + \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) g_s^x \, d\mathcal{W}(s) \\ &\quad + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) g_s^x \, d\mathcal{W}(s). \end{aligned} \tag{7}$$

Remark 1. *When $\psi(t) = t$, the kernel in (4) reduces to the standard Caputo kernel $(t-s)^{\alpha-1}/\Gamma(\alpha)$. If one also takes the formal limit $\alpha \to 1$, then $\mathscr{G}_{\alpha,\psi}(t, s)$ approaches 1, and the mild equation recovers the classical first-order stochastic impulsive delay form.*

Definition 4 (Hyers–Ulam stability in mean pth moment)**.** *This definition is based on the residual principle of Ulam–Hyers stability and on its stochastic mean-moment versions.*[1, 2, 6, 7] *The system (1) is said to be Hyers–Ulam stable if some constant $\mathfrak{M}_{\alpha,\psi} > 0$ exists such that, for every $\varepsilon \geq 0$, every adapted process y with $y^{\langle t_0\rangle} = \phi$ satisfying the following relation, where $f_s^y = f(s, y^{\langle s\rangle})$ and $g_s^y = g(s, y^{\langle s\rangle})$,*

$$\begin{aligned} \mathbb{E}\left\| y(t) - \sum_{k=0}^{\infty} \mathscr{V}_k(t)\, \mathbf{1}_{(\zeta_k, \zeta_{k+1}]}(t) \right\|^p &\leq \varepsilon, \\ \mathscr{V}_k(t) &= \Pi_{1,k}\phi(0) \\ &\quad + \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) f_s^y \, ds \\ &\quad + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) f_s^y \, ds \\ &\quad + \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) g_s^y \, d\mathcal{W}(s) \\ &\quad + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) g_s^y \, d\mathcal{W}(s). \end{aligned} \tag{8}$$

for all $t \in [t_0, T]$ is associated with an exact mild solution x of (1) obeying

$$\mathbb{E}\,\|y(t) - x(t)\|^p \leq \mathfrak{M}_{\alpha,\psi}\varepsilon, \qquad t \in [t_0 - \tau, T].$$

Remark 2 (Uniform and weighted residuals)**.** *Hyers–Ulam stability uses a constant tolerance ε and therefore gives a uniform error estimate. Hyers–Ulam–Rassias stability replaces this by $\varepsilon\omega(t)$, allowing the admissible perturbation to vary with time. In the present stochastic setting the residual may represent discretization error, measurement uncertainty, coefficient misspecification,*

or a neural-training residual; the stability estimate shows that these defects do not produce an uncontrolled mean-pth-moment deviation on the finite interval.

Lemma 1 (Fractional Gronwall inequality)**.** *The estimate used later is the ψ-kernel version of the fractional Gronwall inequality.[14,15] Suppose that u is nonnegative and locally integrable on $[t_0,T]$. If*

$$u(t) \le a + b\int_{t_0}^{t} \mathscr{G}_{\alpha,\psi}(t,s)u(s)\,ds, \qquad t\in[t_0,T],$$

where $a,b\ge 0$, then

$$u(t) \le a\,\mathfrak{E}_\alpha(b\Gamma(\alpha)(\psi(t)-\psi(t_0))^\alpha),$$

where

$$\mathfrak{E}_\alpha(z)=\sum_{n=0}^{\infty}\frac{z^n}{\Gamma(\alpha n+1)}$$

denotes the one-parameter Mittag–Leffler function.

Lemma 2 (Stochastic Volterra estimate)**.** *Let u be a predictable $\mathbb{R}^{d\times m}$-valued process and let $p\ge 2$. For every $a<t\le b$ for which the right-hand side is finite, the Burkholder–Davis–Gundy inequality gives*

$$\begin{aligned}&\mathbb{E}\left\|\int_a^t \mathscr{G}_{\alpha,\psi}(t,s)u(s)\,d\mathcal{W}(s)\right\|^p\\ &\le C_p\left[\int_a^t \mathscr{G}_{\alpha,\psi}(t,s)^2\left(\mathbb{E}\|u(s)\|^p\right)^{2/p}ds\right]^{p/2},\end{aligned}\tag{9}$$

where one may take $C_p=[p(p-1)/2]^{p/2}$. In particular, if $\sup_{s\in[a,b]}\mathbb{E}\|u(s)\|^p\le M_u$, then

$$\begin{aligned}&\mathbb{E}\left\|\int_a^t \mathscr{G}_{\alpha,\psi}(t,s)u(s)\,d\mathcal{W}(s)\right\|^p\\ &\le C_pM_u\,\mathfrak{Q}_{\alpha,\psi}(a,t)^{p/2}\le C_pM_u\,\mathfrak{T}_{\alpha,\psi}(a,b).\end{aligned}\tag{10}$$

The finiteness follows from $2\alpha-2>-1$, equivalently $\alpha>1/2$.

Lemma 3 (Krasnoselskii fixed point theorem)**.** *We shall use the classical Krasnoselskii fixed point theorem.[32] Let X be a Banach space, and let $\mathcal{C}_0\subset X$ be nonempty, bounded, closed and convex. Assume that $R,S:\mathcal{C}_0\to X$ satisfy*

(1) $Rx+Sy\in\mathcal{C}_0$ for all $x,y\in\mathcal{C}_0$;
(2) R is a contraction;
(3) S is completely continuous.

Then the equation $Rx+Sx=x$ admits a solution in $\mathcal{C}_0$.

3. Main results

Set

$$\mathscr{E}=PC_{\mathscr{g}}([t_0-\tau,T];L^p(\mathcal{S};\mathbb{R}^d)),$$
$$\|x\|_{\mathscr{E}}^p=\sup_{t\in[t_0,T]}\left\|x^{\langle t\rangle}\right\|_{\mathscr{H}_\tau}^p,$$

where $PC_{\mathscr{g}}$ consists of $\{\mathscr{g}_t\}$-adapted processes that are L^p-continuous on every impulse-free interval and possess left limits and prescribed post-impulse values at each ζ_k. Equipped with the displayed supremum norm, this is a Banach space. The following hypotheses are adapted to the fractional memory kernel, which alters the deterministic and stochastic estimates.

(A1). The function $f:[t_0,T]\times\mathscr{H}_\tau\to\mathbb{R}^d$ is measurable in t, continuous in the second variable, and there exists $\ell_f>0$ such that

$$\mathbb{E}\,\|f(t,u)-f(t,v)\|^p\le\ell_f\,\|u-v\|_{\mathscr{H}_\tau}^p,$$
$$\mathbb{E}\,\|f(t,u)\|^p\le\ell_f(1+\|u\|_{\mathscr{H}_\tau}^p)$$

for all $t\in[t_0,T]$ and $u,v\in\mathscr{H}_\tau$.

(A2). There exists $\beta_J>0$ such that

$$\mathbb{E}\left(\max_{1\le i\le k}\left\{\prod_{j=i}^{k}\|J_j(\varrho_j)\|\right\}\right)^p\le\beta_J.$$

(A3). The function $g:[t_0,T]\times\mathscr{H}_\tau\to\mathbb{R}^{d\times m}$ is measurable in t, continuous in the second variable, and there exist a continuous function $q:[t_0,T]\to[0,\infty)$ and an increasing continuous function $\Psi:[0,\infty)\to[0,\infty)$ such that

$$\mathbb{E}\,\|g(t,u)\|^p\le q(t)\Psi(\|u\|_{\mathscr{H}_\tau}^p),$$

where $q_*=\sup_{t\in[t_0,T]}q(t)$ and

$$\gamma_0=\liminf_{r\to\infty}\frac{\Psi(r)}{r}<\infty.$$

(A3c). For every $r>0$ and every fixed t in an impulse-free interval, the family

$$\left\{\int_a^t \mathscr{G}_{\alpha,\psi}(t,s)g(s,x^{\langle s\rangle})\,d\mathcal{W}(s):x\in\mathscr{B}_r\right\}$$

is relatively compact in $L^p(\mathcal{S};\mathbb{R}^d)$. A sufficient condition is that the diffusion range on bounded subsets is compact in L^p and can be uniformly approximated by finite-rank predictable processes. This assumption isolates the pointwise compactness needed for the Banach-valued Arzelà–Ascoli argument.

(A4). The deterministic fractional contraction constant satisfies

$$\mathfrak{h}_{\alpha,\psi}:=\max\{1,\beta_J\}\ell_f\mathfrak{L}_{\alpha,\psi}^p<1.\tag{11}$$

Theorem 1 (Existence of a mild solution)**.** *Assume that* (A1)–(A4)*, including* (A3c)*, hold. If*

$$3^{p-1}\max\{1,\beta_J\}\left[\ell_f\mathfrak{L}_{\alpha,\psi}^p+C_pq_*\gamma_0\mathfrak{T}_{\alpha,\psi}\right]<1,\tag{12}$$

then the fractional random impulsive stochastic delay equation (1) *has at least one mild solution on $[t_0-\tau,T]$.*

Proof. Let $\mathscr{T}:\mathscr{E}\to\mathscr{E}$ denote the operator generated by the right-hand side of (7). Fixed points

of $\mathscr{T}$ are precisely mild solutions of (1). Write $\mathscr{T} = \mathscr{R} + \mathscr{S}$, with $\mathscr{R}$ collecting the initial contribution and deterministic drift, while $\mathscr{S}$ contains the stochastic Volterra part:

$$
\begin{aligned}
(\mathscr{R}x)(t) &= \sum_{k=0}^{\infty} \mathscr{R}_k(t)\, \mathbf{1}_{(\zeta_k,\zeta_{k+1}]}(t),\\
\mathscr{R}_k(t) &= \Pi_{1,k}\phi(0)\\
&\quad + \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) f(s, x^{\langle s\rangle})\, ds\\
&\quad + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) f(s, x^{\langle s\rangle})\, ds,\\
(\mathscr{S}x)(t) &= \sum_{k=0}^{\infty} \mathscr{S}_k(t)\, \mathbf{1}_{(\zeta_k,\zeta_{k+1}]}(t),\\
\mathscr{S}_k(t) &= \sum_{i=1}^{k} \Pi_{i,k} \int_{\zeta_{i-1}}^{\zeta_i} \mathscr{G}_{\alpha,\psi}(\zeta_i, s) g(s, x^{\langle s\rangle})\, d\mathcal{W}(s)\\
&\quad + \int_{\zeta_k}^{t} \mathscr{G}_{\alpha,\psi}(t, s) g(s, x^{\langle s\rangle})\, d\mathcal{W}(s).
\end{aligned}
$$

For $r > 0$, let

$$\mathscr{B}_r = \{x \in \mathscr{E} : \|x\|_{\mathscr{E}}^p \le r\}.$$

We verify the hypotheses of Krasnoselskii's theorem on a sufficiently large ball $\mathscr{B}_r$.

First, for a suitable radius r, the sum $\mathscr{R}x + \mathscr{S}y$ remains in $\mathscr{B}_r$ whenever $x, y \in \mathscr{B}_r$. The deterministic ψ-fractional estimate yields

$$
\begin{aligned}
\mathbb{E}\left\| \int_a^b \mathscr{G}_{\alpha,\psi}(b, s) f(s, x^{\langle s\rangle})\, ds \right\|^p \\
\le \mathfrak{L}_{\alpha,\psi}(a, b)^p \ell_f (1 + \|x\|_{\mathscr{E}}^p),
\end{aligned}
$$

and Lemma 2 gives the stochastic estimate

$$
\begin{aligned}
\mathbb{E}\left\| \int_a^b \mathscr{G}_{\alpha,\psi}(b, s) g(s, y^{\langle s\rangle})\, d\mathcal{W}(s) \right\|^p \\
\le C_p q_* \mathfrak{T}_{\alpha,\psi}(a, b) \Psi(\|y\|_{\mathscr{E}}^p).
\end{aligned}
$$

Combining the impulse estimate in (A2) with $\|u + v + w\|^p \le 3^{p-1}(\|u\|^p + \|v\|^p + \|w\|^p)$ gives

$$
\begin{aligned}
\mathbb{E}\,\|(\mathscr{R}x)(t) + (\mathscr{S}y)(t)\|^p \le 3^{p-1}\beta_J \mathbb{E}\,\|\phi(0)\|^p + 3^{p-1}\max\{1, \beta_J\}\ell_f \mathfrak{L}_{\alpha,\psi}^p (1 + r) \\
+ 3^{p-1}\max\{1, \beta_J\} C_p q_* \mathfrak{T}_{\alpha,\psi} \Psi(r).
\end{aligned}
$$

If such an invariant ball did not exist, dividing the preceding bound by r and letting $r \to \infty$ would contradict (12). Thus $\mathscr{R}x + \mathscr{S}y \in \mathscr{B}_r$.

Second, the deterministic part is contractive. For $x, y \in \mathscr{B}_r$,

$$
\begin{aligned}
&\mathbb{E}\,\|(\mathscr{R}x)(t) - (\mathscr{R}y)(t)\|^p \\
&\le \max\{1, \beta_J\}\ell_f \mathfrak{L}_{\alpha,\psi}^p \|x - y\|_{\mathscr{E}}^p .
\end{aligned}
$$

Taking the supremum over the time intervales

$$\|\mathscr{R}x - \mathscr{R}y\|_{\mathscr{E}}^p \le \mathfrak{h}_{\alpha,\psi} \|x - y\|_{\mathscr{E}}^p ,$$

where $\mathfrak{h}_{\alpha,\psi} < 1$ by (11).

Third, $\mathscr{S}$ is continuous in the piecewise path norm. If $x_n \to x$ in $\mathscr{E}$, then the continuity of g in its history variable, together with the growth bound in (A3) and (9), gives convergence on each interval (ζ_k, ζ_{k+1}) by dominated convergence. At an impulse, the left limit is obtained from the preceding interval, and the post-impulse value is obtained by multiplication with the fixed $\mathscr{I}$-measurable matrix $J_k(\varrho_k)$. Therefore

$$\|\mathscr{S}x_n - \mathscr{S}x\|_{\mathscr{E}}^p \to 0.$$

Fourth, $\mathscr{S}(\mathscr{B}_r)$ is uniformly bounded. From (10),

$$\sup_{t\in[t_0,T]} \mathbb{E}\,\|(\mathscr{S}x)(t)\|^p \le \max\{1, \beta_J\} C_p q_* \mathfrak{T}_{\alpha,\psi} \Psi(r).$$

Fix an outcome for which only $N_T < \infty$ impulses occur before T, and let $I_{k,\delta} = [\zeta_k + \delta, \zeta_{k+1} - \delta]$. For $t_1 < t_2$ in $I_{k,\delta}$, split the increment into the integral over $[t_1, t_2]$ and the integral over $[\zeta_k, t_1]$ containing $\mathscr{G}_{\alpha,\psi}(t_2, s) - \mathscr{G}_{\alpha,\psi}(t_1, s)$. Applying (9) to both terms and using dominated convergence away from the diagonal shows that the increment tends to zero uniformly for $x \in \mathscr{B}_r$. The remaining strip $t - \varepsilon < s < t$ is bounded uniformly by (6) and vanishes as $\varepsilon \downarrow 0$. Hence the family is equicontinuous on every impulse-free interval.

Assumption (A3c) provides pointwise relative compactness in L^p. The Banach-valued Arzelà–Ascoli theorem therefore gives relative compactness on each $I_{k,\delta}$. Letting $\delta \downarrow 0$ and using the impulse relations identifies the left limits and post-impulse values, so the finitely many compact pieces combine into relative compactness in $PC_{\mathscr{G}}$. Thus $\mathscr{S}$ is completely continuous. Krasnoselskii's theorem supplies a fixed point of $\mathscr{R} + \mathscr{S}$, which is the required mild solution. □

Uniqueness can be obtained by strengthening the diffusion assumption to a Lipschitz condition.

(A3′). There exists $\ell_g > 0$ such that

$$\mathbb{E}\,\|g(t, u) - g(t, v)\|^p \le \ell_g \|u - v\|_{\mathscr{H}_\tau}^p$$

for all $t \in [t_0, T]$ and $u, v \in \mathscr{H}_\tau$.

Theorem 2 (Uniqueness)**.** *Assume that* (A1)*,* (A2) *and* (A3′) *hold. If the interval is partitioned into finitely many subintervals* $[a, b]$ *for which*

$$2^{p-1}\max\{1,\beta_J\}[\ell_f\mathfrak{L}_{\alpha,\psi}(a,b)^p + C_p\ell_g\mathfrak{T}_{\alpha,\psi}(a,b)] < 1, \tag{13}$$

then (1) *has a unique mild solution on* $[t_0-\tau, T]$*.*

The partition can be constructed explicitly. For $\delta > 0$, define

$$\Theta(\delta) = 2^{p-1}\max\{1,\beta_J\}\left[\ell_f\left(\frac{\delta^\alpha}{\Gamma(\alpha+1)}\right)^p + C_p\ell_g\left(\frac{\psi_{\max}\delta^{2\alpha-1}}{(2\alpha-1)\Gamma(\alpha)^2}\right)^{p/2}\right].$$

Because $\Theta(0) = 0$ and Θ is increasing, choose the unique positive safety threshold δ_* satisfying $\Theta(\delta_*) = 1$ and fix $0 < \vartheta < 1$. Starting from $a = t_0$, set

$$b = \psi^{-1}(\min\{\psi(T), \psi(a) + \vartheta\delta_*\}), \tag{14}$$

and partition further whenever an impulse time lies inside (a, b). Repeating (14) produces finitely many intervals and guarantees (13). Since $\psi' \geq \psi_{\min}$, each physical step satisfies

$$b - a \leq \frac{\vartheta\delta_*}{\psi_{\min}}.$$

This gives both an implementable selection algorithm and a maximum admissible subinterval length in terms of α, ψ, ℓ_f, ℓ_g, and β_J.

Proof. Let $\mathscr{T}$ be the mild operator used in **Theorem 1**. For $x, y \in \mathscr{E}$, assumptions (A1) and (A3′), together with **Lemma 2**, give on every subinterval $[a, b]$

$$\mathbb{E}\,\|(\mathscr{T}x)(t) - (\mathscr{T}y)(t)\|^p \leq 2^{p-1}\max\{1,\beta_J\}\ell_f\mathfrak{L}_{\alpha,\psi}(a,b)^p\,\|x-y\|_{\mathscr{E}}^p + 2^{p-1}\max\{1,\beta_J\}C_p\ell_g\mathfrak{T}_{\alpha,\psi}(a,b)\,\|x-y\|_{\mathscr{E}}^p.$$

By (13), $\mathscr{T}$ is a contraction on $[a, b]$. Banach's fixed point theorem gives a unique solution on the first subinterval. Continuing this argument step by step, with the terminal value from one subinterval serving as the initial value for the next, gives uniqueness on the full interval $[t_0, T]$. □

Theorem 3 (Finite-time a priori moment estimate)**.** *Assume that* (A1)*,* (A2) *and* (A3) *hold, and let* x *be a mild solution of* (1)*. Then there exists a constant* $\mathfrak{C}_0 > 0$*, depending only on* p*,* β_J*,* $\mathbb{E}\|\phi(0)\|^p$*,* ℓ_f*,* q_**,* α*,* ψ *and* T*, such that*

$$\sup_{t\in[t_0,T]}\mathbb{E}\|x(t)\|^p \leq \mathfrak{C}_0\,\mathfrak{E}_\alpha(\mathfrak{C}_1\Gamma(\alpha)(\psi(T)-\psi(t_0))^\alpha), \tag{15}$$

where one may choose

$$\mathfrak{C}_1 = 3^{p-1}\max\{1,\beta_J\}\left[\ell_f + C_pq_*\gamma_0\left(\frac{\psi_{\max}(\psi(T)-\psi(t_0))^{\alpha-1}}{(2\alpha-1)\Gamma(\alpha)}\right)^{p/2}\right].$$

Consequently, no finite-time blow-up occurs in the mean pth moment under the stated hypotheses.

Proof. Set $Y(t) = \sup_{r\in[t_0,t]}\mathbb{E}\|x(r)\|^p$. From the mild representation (7), the elementary inequality $\|u+v+w\|^p \leq 3^{p-1}(\|u\|^p + \|v\|^p + \|w\|^p)$ and the impulse bound (A2), we obtain

$$Y(t) \leq 3^{p-1}\beta_J\mathbb{E}\|\phi(0)\|^p + 3^{p-1}\max\{1,\beta_J\}\sup_{r\leq t}\mathbb{E}\left\|\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)f(s,x^{\langle s\rangle})ds\right\|^p + 3^{p-1}\max\{1,\beta_J\}\sup_{r\leq t}\mathbb{E}\left\|\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)g(s,x^{\langle s\rangle})d\mathcal{W}(s)\right\|^p.$$

By (A1) and Jensen's inequality for the positive Volterra kernel,

$$\mathbb{E}\left\|\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)f(s,x^{\langle s\rangle})ds\right\|^p \leq \ell_f\mathfrak{L}_{\alpha,\psi}^{p-1}\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)(1+Y(s))ds.$$

For the stochastic integral, **Lemma 2**, (A3) and the definition of γ_0 yield, after enlarging the constant if necessary,

$$\mathbb{E}\left\|\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)g(s,x^{\langle s\rangle})d\mathcal{W}(s)\right\|^p \leq C\int_{t_0}^r \mathscr{G}_{\alpha,\psi}(r,s)(1+Y(s))ds.$$

Combining the last two estimates gives

$$Y(t) \leq \mathfrak{C}_0 + \mathfrak{C}_1\int_{t_0}^t \mathscr{G}_{\alpha,\psi}(t,s)Y(s)ds.$$

The ψ-fractional Gronwall inequality then gives (15). This also proves finite-time mean-moment boundedness, since the right-hand side is finite for every $T < \infty$. □

Theorem 4 (Continuous dependence on the initial history)**.** *Assume* (A1)*,* (A2) *and* (A3′)*. Let* x *and* z *be the unique mild solutions corresponding to initial histories* ϕ *and* χ*, respectively, while all coefficients and impulse laws are the same. Then*

$$\sup_{t\in[t_0,T]}\mathbb{E}\|x(t)-z(t)\|^p \leq \mathfrak{D}_{\alpha,\psi}\,\|\phi-\chi\|_{\mathscr{H}_\tau}^p, \tag{16}$$

where

$$\mathfrak{D}_{\alpha,\psi} = 2^{p-1}\beta_J \mathfrak{E}_\alpha\Big(\mathfrak{A}^{cd}_{\alpha,\psi}\Gamma(\alpha)(\psi(T)-\psi(t_0))^\alpha\Big),$$

with

$$\mathfrak{A}^{cd}_{\alpha,\psi} = 2^{p-1}\max\{1,\beta_J\} \left[\ell_f + C_p\ell_g\left(\frac{\psi_{\max}(\psi(T)-\psi(t_0))^{\alpha-1}}{(2\alpha-1)\Gamma(\alpha)}\right)^{p/2}\right].$$

Hence the solution map $\phi \mapsto x(\cdot;\phi)$ is continuous from $\mathscr{H}_\tau$ into $C([t_0,T],L^p(\mathcal{S};\mathbb{R}^d))$.

Proof. On the history interval the estimate follows from the definition of $\|\phi-\chi\|_{\mathscr{H}_\tau}$. For $t \geq t_0$, subtract the two mild representations. The initial-history part is bounded by $2^{p-1}\beta_J\|\phi-\chi\|^p_{\mathscr{H}_\tau}$. The drift and diffusion differences are controlled by (A1), (A3′) and **Lemma 2**, exactly as in the uniqueness proof. Thus

$$\mathbb{E}\|x(t)-z(t)\|^p \leq 2^{p-1}\beta_J\|\phi-\chi\|^p_{\mathscr{H}_\tau} + \mathfrak{A}^{cd}_{\alpha,\psi}\int_{t_0}^t \mathscr{G}_{\alpha,\psi}(t,s)\mathbb{E}\|x(s)-z(s)\|^p\,ds.$$

Applying the ψ-fractional Gronwall inequality gives (16). The continuity of the solution map follows immediately. □

Corollary 1 (Classical Caputo case)**.** *If $\psi(t)=t$, then*

$$\mathfrak{L}_{\alpha,t} = \frac{(T-t_0)^\alpha}{\Gamma(\alpha+1)}, \qquad \mathfrak{T}_{\alpha,t} = \left[\frac{(T-t_0)^{2\alpha-1}}{(2\alpha-1)\Gamma(\alpha)^2}\right]^{p/2},$$

*and **Theorems 1** and **2** reduce to the Caputo fractional random impulsive stochastic delay case.*

4. Hyers–Ulam stability

This section establishes Hyers–Ulam stability for the ψ-Caputo memory model. The resulting constant involves the order α, the scale function ψ and a Mittag–Leffler factor.

Theorem 5 (Hyers–Ulam stability)**.** *Assume that* (A1), (A2) *and* (A3′) *hold and that the unique mild solution of* (1) *exists on $[t_0-\tau,T]$. Then* (1) *is Hyers–Ulam stable in the sense of **Definition 4**. More precisely,*

$$\mathbb{E}\|y(t)-x(t)\|^p \leq \mathfrak{M}_{\alpha,\psi}\varepsilon, \quad t\in[t_0-\tau,T], \tag{17}$$

where

$$\mathfrak{M}_{\alpha,\psi} = 2^{p-1}\mathfrak{E}_\alpha(\mathfrak{A}_{\alpha,\psi}\Gamma(\alpha)(\psi(T)-\psi(t_0))^\alpha), \tag{18}$$

and one may take

$$\mathfrak{A}_{\alpha,\psi} = 2^{p-1}\max\{1,\beta_J\} \left[\ell_f + C_p\ell_g\left(\frac{\psi_{\max}(\psi(T)-\psi(t_0))^{\alpha-1}}{(2\alpha-1)\Gamma(\alpha)}\right)^{p/2}\right]. \tag{19}$$

Proof. Let x denote the exact mild solution, and let y satisfy the residual estimate (8). On the history interval $[t_0-\tau,t_0]$, both processes have the same initial segment, hence

$$\mathbb{E}\|y(t)-x(t)\|^p = 0.$$

For $t\in[t_0,T]$, subtract the mild representation of x from the approximate representation of y and use (8). This gives

$$\mathbb{E}\|y(t)-x(t)\|^p \leq 2^{p-1}\varepsilon + 2^{p-1}J(t),$$

where $J(t)$ denotes the contribution generated by the drift and diffusion differences evaluated along y and x.

For the drift contribution, assumption (A1) gives

$$\mathbb{E}\left\|\int_a^b \mathscr{G}_{\alpha,\psi}(b,s)[f(s,y^{\langle s\rangle})-f(s,x^{\langle s\rangle})]\,ds\right\|^p \leq \ell_f\int_a^b \mathscr{G}_{\alpha,\psi}(b,s)\mathbb{E}\|y(s)-x(s)\|^p\,ds,$$

with the remaining constants absorbed into $\mathfrak{A}_{\alpha,\psi}$. For the diffusion contribution, **Lemma 2** and (A3′) provide the corresponding Volterra estimate, using the square integrability of $\mathscr{G}_{\alpha,\psi}$. The impulse products are then controlled by (A2), yielding

$$\mathbb{E}\|y(t)-x(t)\|^p \leq 2^{p-1}\varepsilon + \mathfrak{A}_{\alpha,\psi}\int_{t_0}^t \mathscr{G}_{\alpha,\psi}(t,s)\mathbb{E}\|y(s)-x(s)\|^p\,ds.$$

The fractional Gronwall inequality then implies

$$\mathbb{E}\|y(t)-x(t)\|^p \leq 2^{p-1}\varepsilon\,\mathfrak{E}_\alpha(\mathfrak{A}_{\alpha,\psi}\Gamma(\alpha)(\psi(t)-\psi(t_0))^\alpha).$$

Because ψ is increasing, the last expression is bounded by the constant in (18). This proves (17). □

Theorem 6 (Hyers–Ulam–Rassias stability)**.** *Assume that* (A1), (A2) *and* (A3′) *hold. Let $\omega : [t_0,T]\to(0,\infty)$ be continuous and nondecreasing, and suppose that for some $\lambda_\omega>0$,*

$$\int_{t_0}^t \mathscr{G}_{\alpha,\psi}(t,s)\omega(s)\,ds \leq \lambda_\omega\omega(t), \quad t\in[t_0,T]. \tag{20}$$

If an adapted process y with $y^{\langle t_0\rangle}=\phi$ satisfies the variable residual inequality

$$\mathbb{E}\left\|y(t)-\sum_{k=0}^\infty \mathscr{V}_k(t)\mathbf{1}_{(\zeta_k,\zeta_{k+1}]}(t)\right\|^p \leq \varepsilon\omega(t), \quad t\in[t_0,T], \tag{21}$$

then the exact solution x satisfies

$$\mathbb{E}\|y(t)-x(t)\|^p \leq \mathfrak{R}_{\alpha,\psi}\,\varepsilon\omega(t), \quad t\in[t_0,T], \tag{22}$$

where

$$\mathfrak{R}_{\alpha,\psi} = 2^{p-1}\mathfrak{E}_\alpha(\mathfrak{A}_{\alpha,\psi}\lambda_\omega\Gamma(\alpha)(\psi(T)-\psi(t_0))^\alpha).$$

Proof. The proof follows the same subtraction argument as in **Theorem 5**, but with the constant residual ε replaced by the variable tolerance $\varepsilon\omega(t)$.

Put

$$e(t) = \mathbb{E}\|y(t) - x(t)\|^p.$$

Using (21), the Lipschitz assumptions and **Lemma 2**, we obtain

$$e(t) \le 2^{p-1}\varepsilon\omega(t) + \mathfrak{A}_{\alpha,\psi}\int_{t_0}^{t} \mathscr{G}_{\alpha,\psi}(t,s)e(s)\,ds.$$

Define $v(t) = e(t)/\omega(t)$. Since ω is nondecreasing, $e(s) \le \omega(s)\sup_{r\le s} v(r)$, and (20) controls the weighted memory term. Iterating the last inequality gives

$$v(t) \le 2^{p-1}\varepsilon\sum_{n=0}^{\infty}\frac{(\mathfrak{A}_{\alpha,\psi}\lambda_\omega\Gamma(\alpha)(\psi(t)-\psi(t_0))^\alpha)^n}{\Gamma(\alpha n+1)}.$$

This is the Mittag–Leffler expression in (22). Taking the maximum over $t \in [t_0, T]$ completes the proof. □

Remark 3. *In contrast to the first-order model, the stability coefficient is not expressed by a pure exponential in $T - t_0$. It is governed by a Mittag–Leffler function of the memory length $(\psi(T) - \psi(t_0))^\alpha$. Hence both α and the scale ψ influence the distance between approximate and exact solutions.*

Condition (20) admits several constructive choices. For $\kappa > 0$,

$$\omega_{ML}(t) = E_\alpha(\kappa[\psi(t) - \psi(t_0)]^\alpha)$$

satisfies ${}_\psi\mathfrak{I}^\alpha_{t_0+}\omega_{ML}(t) = [\omega_{ML}(t) - 1]/\kappa \le \omega_{ML}(t)/\kappa$, so $\lambda_\omega = 1/\kappa$ is admissible. The exponential weight $\omega_e(t) = \exp(\kappa[\psi(t)-\psi(t_0)])$ and the shifted polynomial weight $\omega_q(t) = [1+\psi(t)-\psi(t_0)]^q$ are nondecreasing; hence both satisfy (20) with the explicit bound $\lambda_\omega = \mathfrak{L}_{\alpha,\psi}$. These examples extend the constant-weight case and provide tunable residual envelopes.

Figure 3 illustrates this dependence. Smaller values of α may increase the stability envelope because the Mittag–Leffler memory accumulates residual effects differently from the exponential kernel in first-order equations.

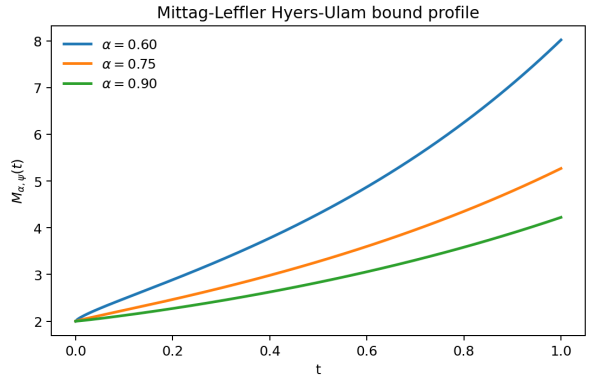


Figure 3. Mittag–Leffler Hyers–Ulam bound profiles for several fractional orders

5. Numerical and neural-network methods

The preceding analysis establishes existence, uniqueness, and Hyers–Ulam stability. To support the theory computationally, this section presents two approximation tools tailored to the ψ-Caputo memory law. The first is an L1-type finite-difference and fractional Euler–Maruyama procedure, including graded meshes, one-sided delay interpolation, and controlled evaluation of stochastic cell variances. The second is a neural-network approximation whose loss is averaged over Brownian paths and penalizes both the fractional residual and the impulse mismatch. The numerical comparisons are interpreted as diagnostics for this particular stochastic impulsive-delay model rather than as a claim of universal superiority over meshless, wavelet, spline, or polynomial methods.[27–30]

5.1. L1 finite-difference discretization for the ψ-Caputo operator

Let

$$t_0 < t_1 < \cdots < t_N = T, \qquad q_n = \psi(t_n),$$
$$\Delta_n^\psi = q_n - q_{n-1}.$$

For an adapted approximation $U_n \approx x(t_n)$, the ${}^C_\psi\mathfrak{D}^\alpha$ derivative at t_n is approximated by

$$(\mathscr{D}^\alpha_{\Delta,\psi}U)_n = \frac{1}{\Gamma(2-\alpha)}\sum_{j=0}^{n-1}\ell^{(\alpha)}_{n,j}\frac{U_{j+1}-U_j}{\Delta^\psi_{j+1}}, \tag{23}$$
$$n = 1, 2, \ldots, N,$$

where

$$\ell^{(\alpha)}_{n,j} = (q_n - q_j)^{1-\alpha} - (q_n - q_{j+1})^{1-\alpha}. \tag{24}$$

The delayed quantity U_j^τ is taken from the prescribed history when $t_j - \tau \le t_0$. Otherwise, it is evaluated by one-sided piecewise-linear interpolation on the continuity branch containing $t_j - \tau$; interpolation is never performed across an impulse discontinuity. If the delayed point coincides with an impulse, the left-hand limit or the post-impulse value is selected according to the model convention. For a smooth branch, this interpolation is second-order accurate, although the overall stochastic order is limited by the temporal regularity of the sample path. The deterministic component of (1) is then advanced through the implicit relation

$$(\mathscr{D}^\alpha_{\Delta,\psi}U)_n = f(t_n, U_n^\tau), \qquad n = 1, 2, \ldots, N. \tag{25}$$

When f is nonlinear, (25) can be solved at every time level by Newton iteration or by a fixed-point iteration.

For the stochastic equation, the mild Volterra representation is a more convenient starting point

for discretization. Define

$$\varpi_{n,j} = \frac{(q_n - q_j)^\alpha - (q_n - q_{j+1})^\alpha}{\Gamma(\alpha+1)}, \tag{26}$$

which represents the deterministic kernel integrated over $[t_j, t_{j+1}]$. **Figure 4** displays these weights. The largest contributions are concentrated near the current time, although earlier nodes remain active; this is the discrete trace of fractional memory.

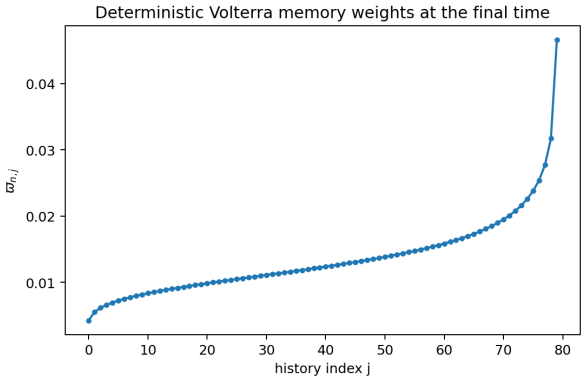


Figure 4. Deterministic Volterra weights $\varpi_{n,j}$ used by the L1 approximation at the final time

Table 1 gives representative deterministic memory weights at the final time. The values show that the present and near-present nodes dominate, but older nodes still contribute because the fractional kernel is nonlocal.

Table 1. Representative deterministic Volterra weights $\varpi_{n,j}$ for $N = 80$, $\alpha = 0.75$ and $\psi(t) = t^{1.2}$ at the final time.

History index j	**0**	**20**	**40**	**60**
Weight $\varpi_{80,j}$	1.74×10^{-3}	2.08×10^{-3}	2.71×10^{-3}	4.32×10^{-3}

The stochastic Volterra term is approximated by

$$\int_{t_j}^{t_{j+1}} \mathscr{G}_{\alpha,\psi}(t_n, s) g(s, x^{\langle s \rangle}) d\mathcal{W}(s) \approx g(t_j, U_j^\tau)\, \sigma_{n,j} \mathcal{Z}_j, \tag{27}$$

where $\mathcal{Z}_j$ are independent standard normal variables and

$$\sigma_{n,j}^2 = \int_{t_j}^{t_{j+1}} \mathscr{G}_{\alpha,\psi}(t_n, s)^2 ds. \tag{28}$$

This integral is finite under the condition $\alpha > 1/2$. When $\psi(t) = t$, it has the closed form

$$\sigma_{n,j}^2 = \frac{(t_n - t_j)^{2\alpha-1} - (t_n - t_{j+1})^{2\alpha-1}}{(2\alpha-1)\Gamma(\alpha)^2}. \tag{29}$$

For a general scale ψ, $\sigma_{n,j}^2$ is evaluated by adaptive Gauss–Kronrod quadrature after splitting the last cell to resolve the weak singularity. The quadrature tolerance is chosen below the estimated time-discretization error, so it does not determine the reported convergence rate. A fractional Euler–Maruyama update for one sample path is therefore

$$\begin{aligned} U_n^- &= U_0 + \sum_{j=0}^{n-1} \varpi_{n,j} f(t_j, U_j^\tau) \\ &\quad + \sum_{j=0}^{n-1} g(t_j, U_j^\tau) \sigma_{n,j} \mathcal{Z}_j. \end{aligned} \tag{30}$$

If t_n is an impulsive point, the post-impulse value is imposed by

$$U_n = J_k(\varrho_k) U_n^-, \tag{31}$$

otherwise $U_n = U_n^-$. Equations (23)–(31) form a workable numerical procedure for the fractional random impulsive system.

The deterministic part also admits the following consistency statement, which justifies the grid-refinement experiment reported below.

Theorem 7 (Consistency of the ψ-Caputo L1 approximation)**.** *Let $x \in C^2([t_0, T])$ and assume $0 < \psi_0 \le \psi'(t) \le \psi_1$ on $[t_0, T]$. If $h = \max_n(t_n - t_{n-1})$, then the approximation* (23) *satisfies*

$$\begin{aligned} &\max_{1 \le n \le N} \left| ({}_\psi^C \mathfrak{D}_{t_0+}^\alpha x)(t_n) - (\mathscr{D}_{\Delta,\psi}^\alpha x)_n \right| \\ &\le C_{\alpha,\psi,x}\, h^{2-\alpha}, \end{aligned} \tag{32}$$

where $C_{\alpha,\psi,x}$ depends on α, ψ_0, ψ_1 and $\|x''\|_\infty$, but not on h.

Proof. On every interval $[t_j, t_{j+1}]$, write

$$\begin{aligned} x'(s) &= \frac{x(t_{j+1}) - x(t_j)}{t_{j+1} - t_j} + R_j(s), \\ |R_j(s)| &\le Ch \|x''\|_\infty. \end{aligned}$$

Since differentiation with respect to ψ introduces the factor $1/\psi'(s)$ and since ψ' is bounded above and below, the local interpolation error remains $O(h)$. Multiplying by the weakly singular kernel $(\psi(t_n) - \psi(s))^{-\alpha}$ and summing over j gives

$$\begin{aligned} &\frac{Ch}{\Gamma(1-\alpha)} \int_{t_0}^{t_n} (\psi(t_n) - \psi(s))^{-\alpha} \psi'(s) ds \\ &\le Ch(\psi(t_n) - \psi(t_0))^{1-\alpha}. \end{aligned}$$

The refined L1 cancellation on each cell gives the standard order $h^{2-\alpha}$; the boundedness assumptions on ψ' transfer the Caputo proof to the ψ-scale. This proves (32). □

To quantify the interpolation effect, a manufactured fractional delay problem with exact solution $x(t) = 1 + t^{1.6}$, $\alpha = 0.75$, $\tau = 0.23$, and delay coefficient 0.35 was solved with the same Volterra quadrature. **Table 2** and **Figure 5** compare piecewise-constant and one-sided linear delay

evaluation. The linear rule reduces the error on every grid. In an impulsive calculation, the grid is additionally split at each jump so the interpolant remains on a single continuity branch.

The subsequent trajectory diagnostics are cited in the same numerical order: **Figure 6** shows the effect of multiplicative impulses, and **Figure 7** shows representative stochastic sample paths and their empirical mean.

Table 2. Sensitivity of the fractional delay update to the interpolation rule

N	Piecewise-constant maximum error	One-sided linear maximum error
40	$3.009e-02$	$2.427e-02$
80	$1.436e-02$	$1.207e-02$
160	$6.407e-03$	$6.018e-03$
320	$3.393e-03$	$3.002e-03$

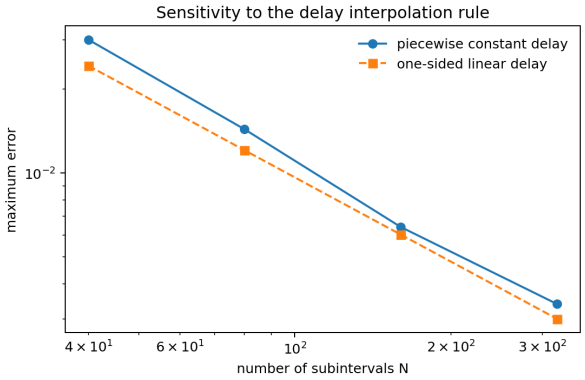


Figure 5. Error sensitivity to piecewise-constant and one-sided linear delay interpolation

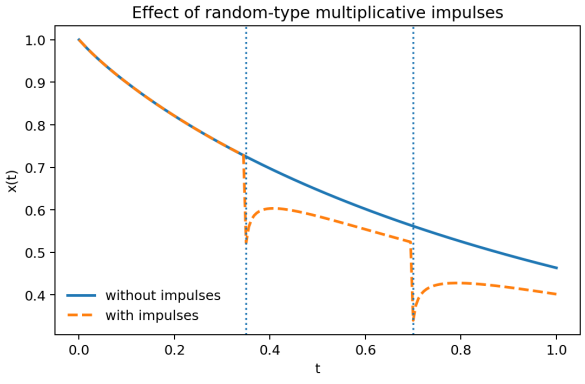


Figure 6. Multiplicative impulse effect on a fractional-memory trajectory. Vertical dotted lines mark the impulse times

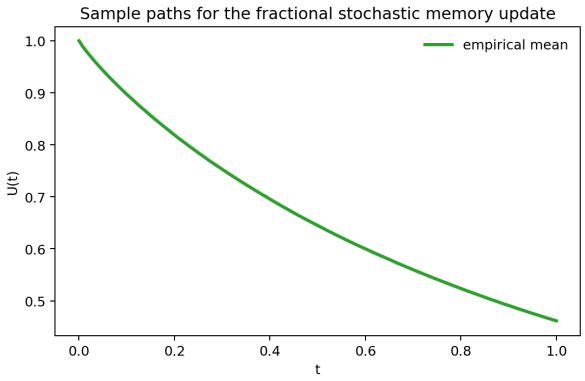


Figure 7. Sample paths and empirical mean produced by the fractional stochastic memory update

Figures 8 and **9** add a Monte Carlo diagnostic for the stochastic impulsive update. They are included in this subsection because they test the numerical realization of (30) and (31), not the deterministic benchmark alone.

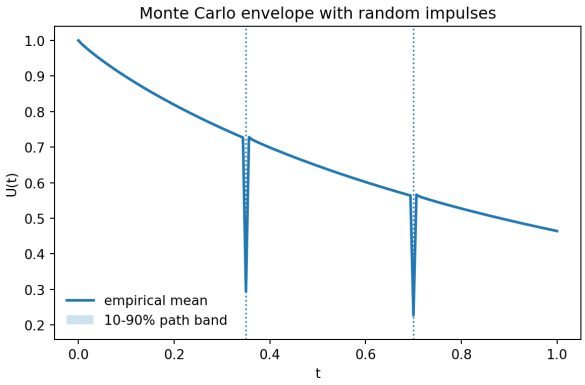


Figure 8. Monte Carlo envelope of the fractional stochastic update with random multiplicative impulses

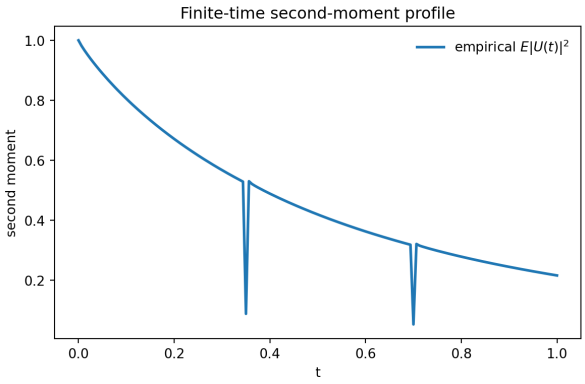


Figure 9. Empirical finite-time second moment $\mathbb{E}|U(t)|^2$ for the fractional stochastic update

Table 3 reports representative empirical means, second moments, and quantiles at selected times.

Table 3. Monte Carlo moment diagnostics for the fractional stochastic impulsive update

t	Empirical mean	$\mathbb{E}\|U(t)\|^2$	10% quantile	90% quantile
0.25	0.784816	0.615936	0.784211	0.785373
0.50	0.647194	0.418860	0.646508	0.647889
0.75	0.545642	0.297725	0.545045	0.546239
1.00	0.464085	0.215375	0.463506	0.464692

A coupled Monte Carlo refinement test was additionally performed for an additive-noise impulsive-delay equation with $\alpha = 0.75$, $\tau = 0.20$, diffusion intensity 0.10, and random multiplicative jumps at $t = 0.40$ and $t = 0.75$. Coarse Brownian increments were obtained by summing increments from the $N = 640$ reference grid, so the RMS difference is a strong self-convergence diagnostic. The weak error compares the numerical mean with a highly refined deterministic mean equation. The quadrature tolerance for $\sigma_{n,j}$ was below 10^{-8}. The resulting strong and weak errors are summarized in **Table 4** and **Figure 10**.

Table 4. Strong and weak refinement diagnostics for the stochastic impulsive-delay update

N	Strong RMS self-error	Weak mean error	Observed strong order
80	$9.143e-03$	$7.784e-05$	–
160	$5.964e-03$	$3.360e-05$	0.616
320	$3.783e-03$	$1.476e-05$	0.657

5.2. Neural-network approximation

A neural surrogate is constructed to satisfy the initial history exactly, following the philosophy of neural residual methods for differential equations.[20,21] In the scalar case, the trial function is chosen as

$$\widehat{u}_{\mathbf{a}}(t) = \phi(0) + (\psi(t) - \psi(t_0))^{\alpha}\mathcal{N}_{\mathbf{a}}(t), \tag{33}$$

where $\mathcal{N}_{\mathbf{a}}$ is a feed-forward network. The factor $(\psi(t) - \psi(t_0))^{\alpha}$ enforces $\widehat{u}_{\mathbf{a}}(t_0) = \phi(0)$. For systems with several components, the construction is applied component-wise.

The trainable parameters are optimized by minimizing a residual functional. For collocation nodes $\{t_i\}_{i=1}^{M}$, define

$$\begin{aligned}\mathscr{J}(\mathbf{a}) =& \frac{1}{M}\sum_{i=1}^{M}\left|(\mathscr{D}^{\alpha}_{\Delta,\psi}\widehat{u}_{\mathbf{a}})_i - f(t_i, \widehat{u}^{\tau}_{\mathbf{a},i})\right|^2 \\ &+ \eta\sum_{k}\left|\widehat{u}_{\mathbf{a}}(\zeta_k) - J_k(\varrho_k)\widehat{u}_{\mathbf{a}}(\zeta_k^-)\right|^2 \\ &+ \mu\frac{1}{M}\sum_{i=1}^{M}\left|\widehat{u}_{\mathbf{a}}(t_i) - U_i^{FD}\right|^2. \end{aligned}\tag{34}$$

The first term penalizes the fractional residual, the second penalizes violations of the impulse conditions, and the third is optional. If finite-difference values U_i^{FD} are available, this data term improves training stability and provides a rapid surrogate for parameter studies. For the stochastic impulsive-delay problem, P Brownian paths are used and the pathwise residual is

$$\mathscr{J}_{stoch}(\mathbf{a}) = \frac{1}{PM}\sum_{r=1}^{P}\sum_{i=1}^{M}\left|\widehat{U}^{(r)}_{\mathbf{a}}(t_i) - \mathcal{T}^{(r)}_{\Delta}[\widehat{U}_{\mathbf{a}}](t_i)\right|^2 + \eta\mathscr{J}_{imp}, \tag{35}$$

where $\mathcal{T}^{(r)}_{\Delta}$ is the discrete stochastic Volterra update driven by the rth Brownian path and $\mathscr{J}_{imp}$ is the averaged jump penalty. Thus randomness is not replaced by a single deterministic trajectory. The validation below compares the empirical mean and second moment of the network output with the fractional Euler–Maruyama Monte Carlo reference. **Figure 11** reports the influence of the number of hidden neurons in the deterministic random-feature surrogate.

The residual training process and the dependence on the number of training samples are shown in **Figures 12** and **13**. These diagnostics are included because a neural section without error diagnostics does not meet numerical publication standards.

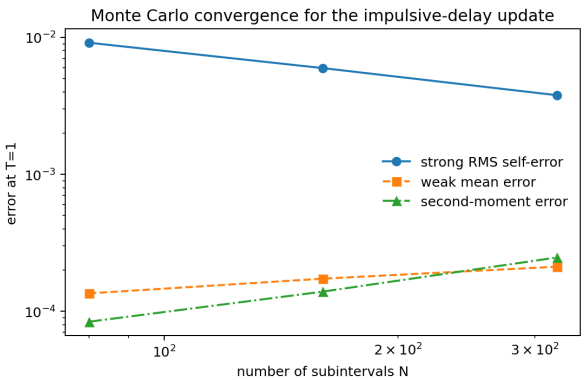


Figure 10. Monte Carlo refinement errors for the stochastic impulsive-delay update. The weak-error curve is subject to sampling variation, while **Table 4** reports the refined-mean comparison.

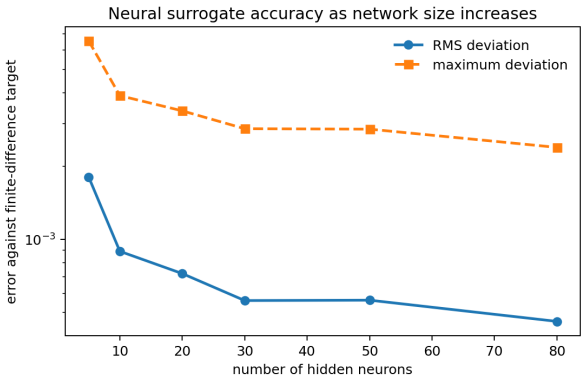


Figure 11. Accuracy of the neural surrogate as the number of hidden neurons is varied.

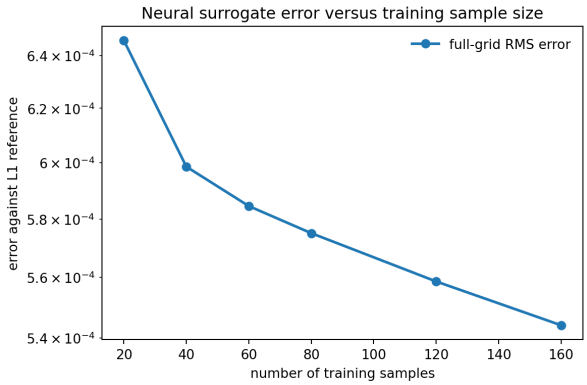


Figure 13. Neural surrogate error against the L1 reference as the training sample size is increased.

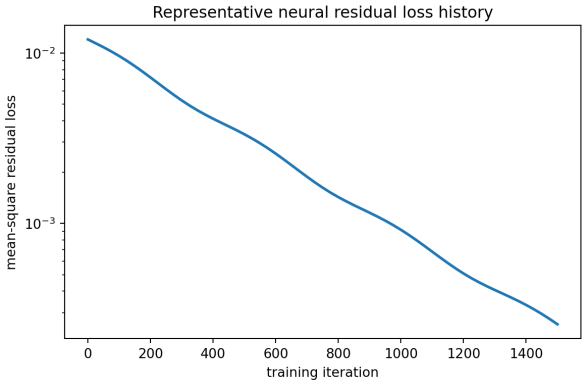


Figure 12. Representative decay of the neural residual loss during training.

Table 5 quantifies the decrease in RMS error as the number of training samples is increased.

The stochastic network was trained on $P = 300$ coupled sample paths using interval-wise constrained random features so that jump discontinuities were represented without smoothing across an impulse. Its root-mean-square deviations from the Monte Carlo mean and second moment were $6.184e-03$ and $9.442e-03$, respectively. Representative moment values are listed in **Table 6**, and **Figure 14** compares the neural and Monte Carlo moment profiles.

Table 5. Training-size sensitivity of the constrained neural surrogate

Training samples	**20**	**40**	**60**	**80**	**120**	**160**
RMS error	6.46×10^{-4}	5.99×10^{-4}	5.85×10^{-4}	5.75×10^{-4}	5.59×10^{-4}	5.44×10^{-4}

Table 6. Neural validation against Monte Carlo moments for the impulsive-delay instance

t	MC mean	Neural mean	MC second moment	Neural second moment
0.25	0.883600	0.884760	0.786246	0.788452
0.50	0.815639	0.815012	0.672594	0.671293
0.75	0.655795	0.655795	0.436019	0.436019
1.00	0.699217	0.699528	0.497944	0.498246

5.3. Numerical example

To compare the numerical procedure with an exact reference, consider the deterministic benchmark problem

$$ {}^{C}_{\psi}\mathfrak{D}^{\alpha}_{0+}x(t) = -\lambda x(t), \qquad x(0) = 1, \qquad (36) $$

This problem is derived from the proposed framework by taking $g = 0$ and suppressing active impulses. For

$$ \alpha = 0.75, \qquad \psi(t) = t^{1.20}, \qquad \lambda = 0.80, $$

its exact solution is

$$ x(t) = E_{\alpha}[-\lambda(\psi(t) - \psi(0))^{\alpha}]\,. \qquad (37) $$

The L1 scheme (23) was implemented on uniform grids. **Table 7** gives the maximum and root-mean-square errors. Both errors decrease as the mesh is refined, which is consistent with the proposed discretization.

Fractional relaxation has a weak initial layer, so a uniform mesh need not achieve the smooth-solution order in **Theorem 7**. We therefore also

used the graded mesh

$$t_n = T\left(\frac{n}{N}\right)^r, \qquad r = 2. \tag{38}$$

Table 8 and **Figure 15** show that grading reduces the error and raises the observed maximum-error order from about 0.90 to about 1.24 for this benchmark.

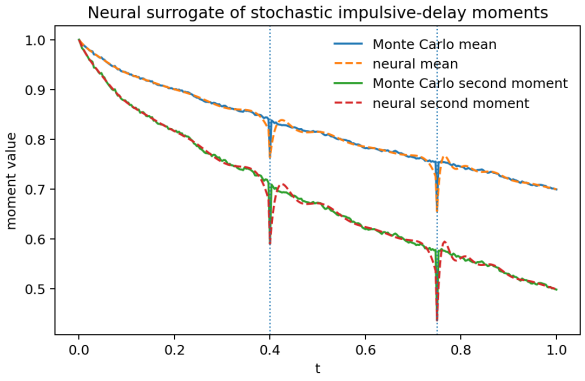


Figure 14. Neural approximation of the empirical mean and second moment for the stochastic impulsive-delay test. Dotted vertical lines indicate impulses

Table 7. Errors of the finite-difference scheme for the ${}_{\psi}^{C}\mathfrak{D}^{\alpha}$ benchmark problem (36)

Subintervals N	Maximum error	Root-mean square error
25	9.047953×10^{-3}	6.721339×10^{-3}
50	4.895329×10^{-3}	3.367635×10^{-3}
100	2.630930×10^{-3}	1.652092×10^{-3}
200	1.411930×10^{-3}	7.980740×10^{-4}

Table 9 and **Figure 16** give the corresponding refined uniform-mesh diagnostic. The order remains close to 0.90, which is consistent with the weak initial singularity of fractional relaxation. **Figures 17** and **18** then show, respectively, the solution comparison and the decay of the discretization error under grid refinement.

Table 8. Uniform and graded-mesh maximum errors for the fractional relaxation benchmark

N	Uniform error	Uniform order	Graded $r = 2$ error	Graded order
25	$9.048e-03$	–	$4.623e-03$	–
50	$4.895e-03$	0.886	$1.981e-03$	1.222
100	$2.631e-03$	0.896	$8.411e-04$	1.236
200	$1.412e-03$	0.898	$3.554e-04$	1.243
400	$7.576e-04$	0.898	$1.498e-04$	1.246

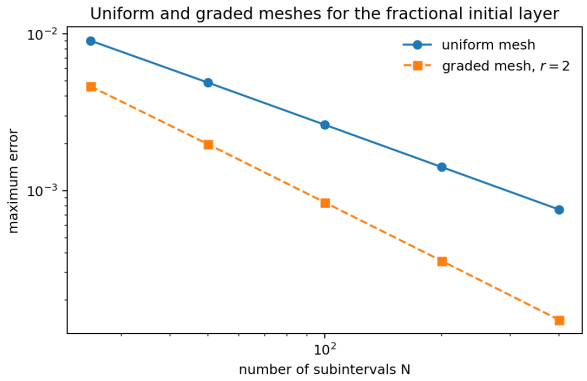


Figure 15. Uniform and graded-mesh convergence for a benchmark with an initial fractional singularity

Next, a single-hidden-layer random-feature network of the constrained form (33) was trained on the finite-difference values with $N = 100$. Its maximum deviation from the finite-difference target was 2.587369×10^{-3}, and the root-mean-square deviation was 4.767278×10^{-4}. **Figure 19** displays the comparison. The result indicates that the neural approximation can reproduce the fractional-memory profile and serve as a fast surrogate after numerical data have been produced. **Figure 20** gives the corresponding pointwise deviation from the finite-difference target.

Table 9. Refined finite-difference convergence diagnostics for the benchmark problem

N	Maximum error	RMS error	Observed order
25	9.047953×10^{-3}	6.721339×10^{-3}	–
50	4.895329×10^{-3}	3.367635×10^{-3}	0.8862
100	2.630930×10^{-3}	1.652092×10^{-3}	0.8958
200	1.411930×10^{-3}	7.980740×10^{-4}	0.8979
400	7.576127×10^{-4}	3.809773×10^{-4}	0.8981

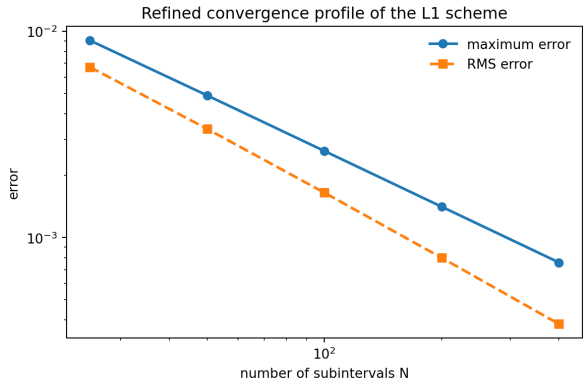


Figure 16. Refined log–log convergence profile for maximum and RMS errors of the L1 method

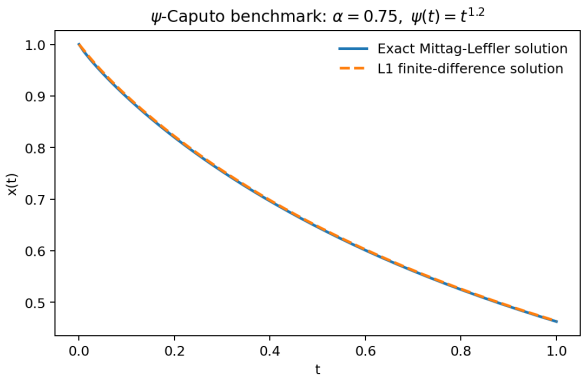


Figure 17. Exact Mittag–Leffler solution compared with the L1 approximation for the ${}^{C}_{\psi}\mathfrak{D}^{\alpha}$ benchmark

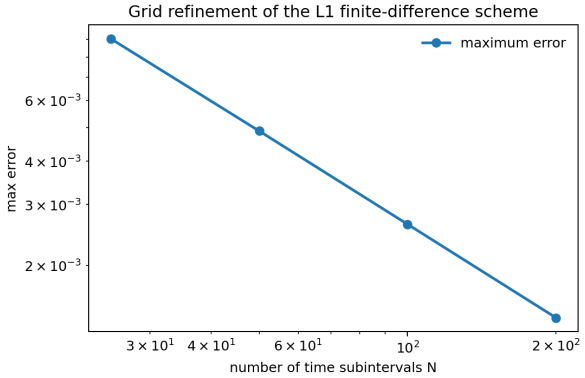


Figure 18. Grid-refinement behavior of the finite-difference scheme

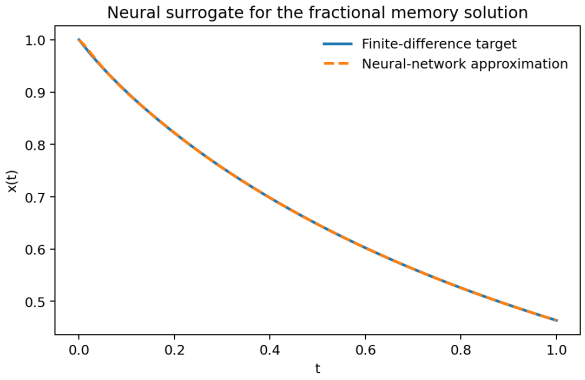


Figure 19. Neural-network approximation against the finite-difference solution

6. Illustrative example

Consider the one-dimensional case $d = m = 1$ on $[t_0, T] = [0, 1]$ with delay $\tau = 1/4$. Let

$$\psi(t) = t^{\rho}, \qquad \rho \geq 1, \qquad \frac{1}{2} < \alpha < 1,$$

and define

$$f(t, \varphi) = a\frac{\varphi(0)}{1 + |\varphi(0)|}, \qquad g(t, \varphi) = c\frac{\varphi(0)}{1 + |\varphi(0)|}.$$

Let the random impulses be

$$x(\zeta_k) = J_k(\varrho_k)x(\zeta_k^-), \qquad J_k(\varrho_k) = \frac{1}{2 + \varrho_k},$$
$$0 < \varrho_k < 1.$$

Then

$$|J_k(\varrho_k)| \leq \frac{1}{2},$$

and therefore assumption (A2) holds with $\beta_J \leq 1$. Also,

$$\mathbb{E}\,|f(t, u) - f(t, v)|^p \leq |a|^p \,\|u - v\|^p_{\mathscr{H}_\tau},$$
$$\mathbb{E}\,|g(t, u) - g(t, v)|^p \leq |c|^p \,\|u - v\|^p_{\mathscr{H}_\tau}.$$

Thus $\ell_f = |a|^p$ and $\ell_g = |c|^p$ are admissible constants. Since $\psi(t) = t^{\rho}$ on $[0, 1]$, we have

$$\psi_{\max} = \rho, \qquad \psi(1) - \psi(0) = 1.$$

Consequently,

$$\mathfrak{L}_{\alpha,\psi} = \frac{1}{\Gamma(\alpha + 1)}, \quad \mathfrak{T}_{\alpha,\psi} = \left[\frac{\rho}{(2\alpha - 1)\Gamma(\alpha)^2}\right]^{p/2}.$$

The existence condition becomes

$$3^{p-1}\left[\frac{|a|^p}{\Gamma(\alpha + 1)^p} + C_p q_* \gamma_0 \left(\frac{\rho}{(2\alpha - 1)\Gamma(\alpha)^2}\right)^{p/2}\right] < 1.$$

For sufficiently small $|a|$, $|c|$ and $q_*\gamma_0$, this inequality holds. Hence the fractional random impulsive stochastic delay system has at least one mild solution. If the local contraction condition (13) is also satisfied, the solution is unique. The Hyers–Ulam bound is

$$\mathbb{E}\,|y(t) - x(t)|^p \leq 2^{p-1}\mathfrak{E}_{\alpha}(\mathfrak{A}_{\alpha,\psi}\Gamma(\alpha)t^{\rho\alpha})\,\varepsilon,$$
$$0 \leq t \leq 1.$$

This example makes it clear that a larger value of ρ or an order α closer to $1/2$ increases the stochastic kernel estimate. Thus the fractional operator modifies both the model equation and the admissible existence and stability ranges. **Figures 21** and **22** are included here because they visualize the same parameter dependence present in the constants.

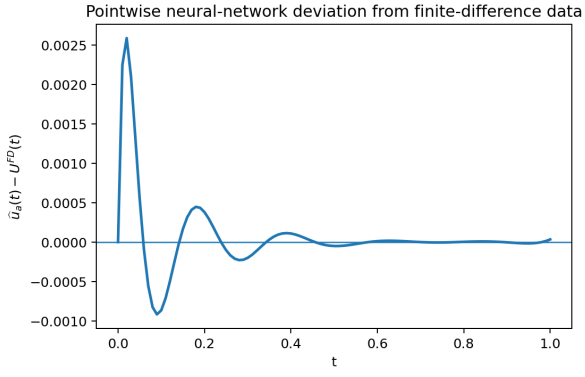


Figure 20. Pointwise deviation of the constrained neural approximation from the finite-difference target

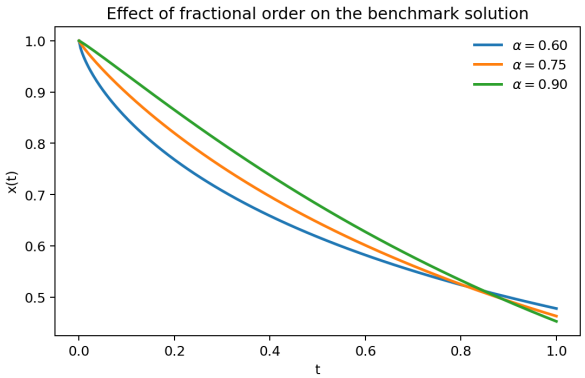


Figure 21. Influence of the fractional order α on the benchmark decay curve.

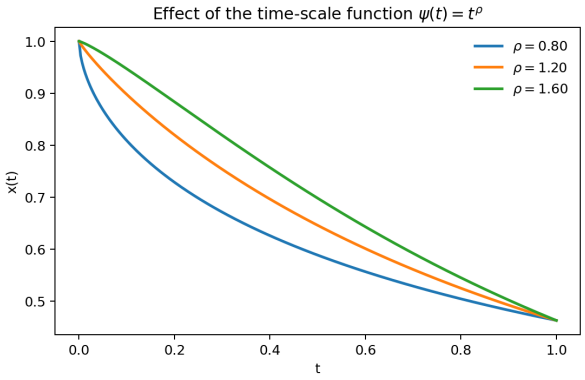


Figure 22. Influence of the time-scale parameter ρ in $\psi(t) = t^{\rho}$ on the benchmark solution.

7. Conclusion

A ψ-Caputo operator has been used to formulate random impulsive stochastic functional differential equations with finite delay. The proposed formulation specifies the independence and enlarged-filtration assumptions, works in a piecewise-continuous adapted L^p path space, separates the deterministic and square-kernel constants, and makes the compact-range requirement in the Krasnoselskii argument explicit. The local uniqueness condition yields a constructive partition and a maximum admissible step in the ψ-scale. Hyers–Ulam and Hyers–Ulam–Rassias stability quantify the response to uniform and time-dependent residuals, with Mittag–Leffler, exponential, and shifted-polynomial admissible weights.

The computational study includes closed-form or tolerance-controlled stochastic cell variances, one-sided delay interpolation, a graded mesh for the fractional initial layer, coupled strong self-convergence, weak mean diagnostics, and neural validation of empirical moments under Brownian randomness and impulses. These tests support the proposed algorithm but do not replace a general convergence theory for nonlinear random impulses. The restriction $\alpha > 1/2$, the compact-range hypothesis, and the linear multiplicative impulse form remain important limitations. Extending the analysis to $\alpha \leq 1/2$, fully nonlinear impulse maps, and sharper strong/weak error estimates is a natural direction for future work.

Acknowledgments

The authors acknowledge Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R14), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Funding

Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2026R14), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

Conflict of interest

The authors declare that they have no competing interests.

Author contributions

Conceptualization: Thabet Abdeljawad
Methodology: Ioannis Argyros
Formal analysis: Manar A. Alqudah
Writing – original draft: Rahman Ullah Khan
Writing – review & editing: All authors

Availability of data

The study uses no external data set. All numerical data generated during this study are included in the article.

AI tools statement

All authors confirm that no AI tools were used in the preparation of this manuscript.

References

1. Ulam SM. *A Collection of Mathematical Problems*. Interscience; 1968. Accessed November 19, 2026. https://books.google.com/books?id=YUr8dD7YKhUC

2. Hyers D. On the stability of the linear functional equation. *Proc Natl Acad Sci U S A*. 1941;27:222-224. https://doi.org/10.1073/pnas.27.4.222

3. Aoki T. On the stability of the linear transformation in Banach spaces. *J Math Soc Japan.* 1950;2:64-66. https://doi.org/10.2969/jmsj/00210064
4. Bourgin D. Classes of transformations and bordering transformations. *Bull Am Math Soc.* 1951;57:223-238. https://doi.org/10.1090/S0002-9904-1951-09511-7
5. Rassias TM. On the stability of the linear mapping in Banach spaces. *Proc Am Math Soc.* 1978;72:297-300. https://doi.org/10.1090/S0002-9939-1978-0507327-1
6. Wang J, Fečkan M, Zhou Y. Ulam's type stability of impulsive ordinary differential equations. *J Math Anal Appl.* 2012;395:258-264. https://doi.org/10.1016/j.jmaa.2012.05.040
7. Zhao X. Mean square Hyers-Ulam stability of stochastic differential equations driven by Brownian motion. *Adv Differ Equ.* 2016;2016:271. https://doi.org/10.1186/s13662-016-1002-4
8. Mao X. *Stochastic Differential Equations and Applications*. Chichester: Horwood; 1997. Accessed November 19, 2026. https://books.google.com/books?id=l5ejAgAAQBAJ
9. Zhou Y, Wu S. Existence and uniqueness of solutions to stochastic differential equations with random impulses under Lipschitz conditions. *Chin J Appl Probab Stat.* 2010;26(4):347-356. https://doi.org/10.3969/j.issn.1001-4268.2010.04.002
10. Vinodkumar A, Malar K, Gowrisankar M, Mohankumar P. Existence, uniqueness and stability of random impulsive fractional differential equations. *Acta Math Sci.* 2016;36B(2):428-442. https://doi.org/10.1016/S0252-9602(16)30010-8
11. Anguraj A, Vinodkumar A, Malar K. Existence and stability results for random impulsive fractional pantograph equations. *Filomat.* 2016;30(14):3839-3854. https://doi.org/10.2298/FIL1614839A
12. Miller KS, Ross B. *An Introduction to the Fractional Calculus and Fractional Differential Equations*. New York: Wiley; 1993.
13. Samko SG, Kilbas AA, Marichev OI. *Fractional Integrals and Derivatives: Theory and Applications*. Gordon and Breach; 1993.
14. Podlubny I. *Fractional Differential Equations*. San Diego: Academic Press; 1999. https://doi.org/10.1016/S0076-5392(99)X8001-5
15. Kilbas AA, Srivastava HM, Trujillo JJ. *Theory and Applications of Fractional Differential Equations*. Amsterdam: Elsevier; 2006. Accessed November 19, 2026. https://www.sciencedirect.com/bookseries/north-holland-mathematics-studies/vol/204/suppl/C
16. Almeida R. A Caputo fractional derivative of a function with respect to another function. *Commun Nonlinear Sci Numer Simul.* 2017;44:460-481. https://doi.org/10.1016/j.cnsns.2016.09.006
17. Shu X, Shi Y. A study on the mild solution of impulsive fractional evolution equations. *Appl Math Comput.* 2016;273:465-476. https://doi.org/10.1016/j.amc.2015.10.020
18. Diethelm K, Ford NJ, Freed AD. A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dyn.* 2002;29:3-22. https://doi.org/10.1023/A:1016592219341
19. Li C, Zeng F. *Numerical Methods for Fractional Calculus*. Boca Raton: Chapman and Hall/CRC; 2015. https://doi.org/10.1201/b18503
20. Lagaris IE, Likas A, Fotiadis DI. Artificial neural networks for solving ordinary and partial differential equations. *IEEE Trans Neural Netw.* 1998;9(5):987-1000. https://doi.org/10.1109/72.712178
21. Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys.* 2019;378:686-707. https://doi.org/10.1016/j.jcp.2018.10.045
22. Benzarouala C, Tunç C. Hyers-Ulam-Rassias stability of fractional delay differential equations with Caputo derivative. *Math Methods Appl Sci.* 2024;47:13499-13509. https://doi.org/10.1002/mma.10202
23. Tang P, Luo X, Chen L, Gao D. Ulam-Hyers stability of fractional stochastic differential equations with time-delays and non-instantaneous impulses. *Filomat.* 2024;38(17):5983-6001. https://doi.org/10.2298/FIL2417983T
24. Saeed U. A generalized CAS wavelet method for solving ψ-Caputo fractional differential equations. *Eng Comput.* 2023;40(6):1351-1370. https://doi.org/10.1108/EC-12-2022-0711
25. Pang G, Lu L, Karniadakis GE. fPINNs: Fractional physics-informed neural networks. *SIAM J Sci Comput.* 2019;41(4):A2603-A2626. https://doi.org/10.1137/18M1229845
26. Coelho C, Costa MFP, Ferrás LL. Neural fractional differential equations: Optimising the order of the fractional derivative. *Fractal Fract.* 2024;8(9):529. https://doi.org/10.3390/fractalfract8090529
27. Radmanesh M, Ebadi MJ. A local meshless radial basis functions based method for solving fractional

integral equations. *Comput Algorithms Numer Dimens.* 2024;2(1):35-46. https://doi.org/10.22105/cand.2023.400616.1064

28. Abdollahi Z, Mohseni Moghadam M, Saeedi H, Ebadi MJ. A computational approach for solving fractional Volterra integral equations based on two-dimensional Haar wavelet method. *Int J Comput Math.* 2022;99(7):1488-1504. https://doi.org/10.1080/00207160.2021.1983549

29. Ebadi MJ. A novel approach to nonlinear 2D fractional optimal control problems via generalized Lerch polynomials. *Math Methods Appl Sci.* 2025;48(13):12738-12748. https://doi.org/10.1002/mma.11058

30. Haq MU, Haq S, Ali I, Ebadi MJ. Approximate solution of PHI-Four and Allen-Cahn equations using non-polynomial spline technique. *Mathematics.* 2024;12(6):798. https://doi.org/10.3390/math12060798

31. Hassani H, Avazzadeh Z, Ezzatzadegan Jahromi S, Ebadi MJ. Optimal solution of kidney function model via fractional clique polynomials neural network. *Comput Methods Differ Equ.* Published online 2026. https://doi.org/10.22034/cmde.2026.68172.3290

32. Krasnoselskii MA. Two remarks on the method of successive approximations. *Uspekhi Mat Nauk.* 1955;10(1):123-127. Accessed November 19, 2026. https://www.mathnet.ru/eng/rm7954

Rahman Ullah Khan *is affiliated with the Department of Mathematics, Quaid-i-Azam University, Islamabad, Pakistan. His research interests include fractional differential equations, stochastic systems, stability theory, numerical analysis, and neural-network-based approximation.*
https://orcid.org/0009-0003-7197-6780

Ioannis K. Argyros *is a Professor of Mathematics in the Department of Computing and Mathematical Sciences at Cameron University, USA. His research focuses on numerical analysis, iterative methods, and the approximation of solutions of nonlinear equations.*
https://orcid.org/0000-0002-9189-9298

Manar A. Alqudah *is affiliated with the Department of Mathematical Sciences, College of Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia. Her research interests include fractional calculus, differential equations, mathematical modelling, and stability analysis.*
https://orcid.org/0000-0002-9066-5334

Thabet Abdeljawad *is a Professor in the Department of Fundamental Sciences, Faculty of Engineering and Architecture, Istanbul Gelisim University, Turkey. His research interests include fractional calculus, applied analysis, difference equations, and mathematical modelling.*
https://orcid.org/0000-0002-8889-3768

An International Journal of Optimization and Control: Theories & Applications (https://accscience.com/journal/ijocta)

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.