

Article

Not peer-reviewed version

A Fast Chebyshev Spectral Collocation Method for a Coupled System of Nonlinear Klein–Gordon Equations with Caputo Fractional Memory

[Nauryzbay Adil](#), [Zhanars Abdiramanov](#)^{*}, Abdumauvlen Berdyshev

Posted Date: 29 April 2026

doi: 10.20944/preprints202604.2054.v1

Keywords: Fractional Klein–Gordon equation; Caputo memory integral; Chebyshev spectral collocation; Sum-of-Exponentials acceleration; Newmark scheme; convergence analysis



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

A Fast Chebyshev Spectral Collocation Method for a Coupled System of Nonlinear Klein–Gordon Equations with Caputo Fractional Memory

Nauryzbay Adil¹, Zhanars Abdiramanov^{1,2,*} and Abdumauvlen Berdyshev^{1,2}

¹ Department of Mathematics and Mathematical Modelling, Faculty of Mathematics, Physics and Informatics, Abai Kazakh National Pedagogical University, Almaty 050010, Kazakhstan

² Institute of Information and Computational Technologies, Science Committee of the Ministry of Science and Higher Education, Almaty 050010, Kazakhstan

* Correspondence: zh.abdiramanov@abaiuniversity.edu.kz

Abstract

We develop a fast Chebyshev spectral collocation method for a coupled system of nonlinear Klein–Gordon equations augmented by Caputo-type fractional memory integrals. The governing equations retain the classical second-order time derivative as the leading operator and incorporate weakly singular convolution integrals with power-law kernels $t^{-\alpha}$, $\alpha \in (0, 1)$, modelling viscoelastic memory damping rather than replacing the wave operator. The spatial discretisation employs Chebyshev–Gauss–Lobatto collocation, while the temporal integration uses a Newmark scheme ($\beta_{\text{NM}} = 1/4$) with the spatial operator treated implicitly and both the L1 memory sums and the cubic nonlinearities evaluated explicitly at the known time level; a linear extrapolation of the nonlinear terms eliminates the need for Newton–Raphson iterations. The disparate memory tails arising from two distinct fractional orders $\alpha \neq \beta$ are compressed by independent Sum-of-Exponentials (SOE) approximations, reducing the per-step memory cost from $\mathcal{O}(N_t)$ to $\mathcal{O}(p + N_{\text{exp}})$ and the total complexity from $\mathcal{O}(N_t^2)$ to $\mathcal{O}(N_t(p + N_{\text{exp}}))$. A rigorous stability estimate and a global convergence bound are established using a discrete Gronwall inequality. Numerical experiments confirm the temporal convergence rate $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)})$, spectral spatial accuracy, and the practical speedup afforded by the SOE acceleration. A solitary wave collision scenario illustrates the method's capability to capture asymmetric dispersive wakes generated by the fractional memory. The algorithmic architecture is dimension-independent by construction; a concrete extension pathway to multi-dimensional tensor-product Chebyshev grids, including Kronecker-product operators and Sylvester-based solvers, is presented.

Keywords: Fractional Klein–Gordon equation; Caputo memory integral; Chebyshev spectral collocation; Sum-of-Exponentials acceleration; Newmark scheme; convergence analysis

1. Introduction

Nonlinear time-fractional partial differential equations (FPDEs) have become a cornerstone in mathematical physics, where fractional integrodifferential operators model anomalous diffusion, viscoelastic damping, and dispersive wave propagation in heterogeneous media [1,2,15]. Replacing classical wave models with fractional hyperbolic models allows one to capture “memory effects” — where the current state of the system depends on its entire historical trajectory [3].

The nonlinear Klein–Gordon equation is fundamental in relativistic quantum mechanics and in condensed matter physics for modeling dispersive solitary waves (solitons). Its extension to the time-fractional domain — incorporating Caputo-type memory integrals into the wave equation — introduces non-trivial long-range temporal interactions. Due to the non-local nature of fractional operators, exact analytical solutions are rare and generally limited to linear, uncoupled equations [4].

Developing stable and efficient numerical methods for such systems is therefore an active area of computational mathematics.

Significant progress has been made for single-component fractional wave equations. Sun and Wu [5] established the $L1$ -type finite difference scheme achieving $\mathcal{O}(\Delta t^{2-\alpha})$ accuracy for the Caputo derivative. Berdyshev et al. [18] investigated non-local boundary problems for parabolic–hyperbolic equations with Riemann–Liouville fractional operators. Baigereyev et al. [19] developed convergent numerical methods for fractional models of flow in porous media. To improve spatial accuracy, Bhrawy et al. [6] and Zaky [7] introduced spectral collocation methods with Chebyshev and Legendre polynomials, achieving exponential spatial convergence $\mathcal{O}(e^{-cN})$. Baishemirov et al. [20] proposed numerical methods for convection–diffusion equations with Caputo fractional derivatives, incorporating memory effects.

However, extending to **coupled nonlinear fractional systems** introduces three significant computational and analytical challenges:

1. **Nonlinear coupling:** Coupled systems with cubic self-interactions u^3 and inter-component coupling ηuv^2 require fully implicit treatments at each time step, typically via Newton–Raphson iterations with Jacobian assembly [8,9].
2. **Dense memory convolution:** Evaluating the Caputo memory integral at time t_n requires $\mathcal{O}(n)$ operations, leading to $\mathcal{O}(N_t^2)$ total complexity and $\mathcal{O}(N_t \cdot N)$ memory storage, which becomes prohibitive for long-time simulations [10].
3. **Coupled stability and convergence analysis:** Establishing rigorous a priori stability for systems with *two distinct* fractional memory orders $\alpha \neq \beta$ and inter-component cubic coupling requires controlling the cross-energy exchange between components through a unified discrete energy functional — a non-trivial extension of scalar fractional stability theory that, to our knowledge, has not been addressed in the existing literature.

To address the dense memory bottleneck, early approaches used the “short-memory principle” by truncating the historical tail, at the cost of convergence accuracy [4]. More recently, fast memory algorithms based on Sum-of-Exponentials (SOE) approximations of the fractional kernel, using Gauss–Laguerre quadratures or Talbot contours, have been developed by Jiang et al. [11] and Baffet & Hesthaven [10], reducing the temporal complexity to $\mathcal{O}(N_t N_{\text{exp}})$. However, implementing dual-SOE arrays for coupled nonlinear systems with disparate fractional orders remains largely unexplored.

Comparison with existing approaches. Table 1 summarises the key differences between the present work and the most closely related studies.

Table 1. Comparison of the present method with existing approaches for fractional Klein–Gordon-type equations. KG = Klein–Gordon, KGS = Klein–Gordon–Schrödinger. “–” indicates the feature is absent.

Feature	[5]	[6]	[7]	[11]	[21]	[22]	Present
Equation type	scalar KG	scalar KG	scalar KG	scalar diff.	coupled KGS	coupled KGS	coupled KG–KG
Nonlinear (u^3, uv^2)	–	✓	✓	–	✓	✓	✓
Spatial method	FDM	Cheb.	Leg.	FDM	FDM	Galerkin	Cheb. colloc.
Fractional type	time	time	time	time	space	space	time (Caputo)
Two distinct orders	–	–	–	–	–	–	✓ ($\alpha \neq \beta$)
SOE acceleration	–	–	–	✓	–	–	dual SOE
IMEX lineariz.	–	–	–	–	–	–	✓

Specifically: (i) Sun and Wu [5] and Liao et al. [13] developed $L1$ -type finite difference schemes for *scalar* subdiffusion and wave equations, achieving $\mathcal{O}(\Delta t^{2-\alpha})$ temporal accuracy but without spectral spatial discretization or nonlinear coupling. (ii) Bhrawy et al. [6] and Zaky [7] applied Chebyshev/Legendre spectral collocation to *scalar* nonlinear fractional Klein–Gordon equations, but did not consider coupled systems, disparate orders $\alpha \neq \beta$, or the $\mathcal{O}(N_t^2)$ memory bottleneck. (iii) Jiang et al. [11] introduced the SOE acceleration for *linear* fractional diffusion equations, and recently Shiyapov et al. [12] developed high-order structure-preserving spectral schemes with fast memory for nonlocal nonlinear diffusion equations; we extend this to a *dual-SOE* architecture handling two

independent fractional orders within a coupled nonlinear Klein–Gordon system. (iv) Among coupled fractional systems, Li, Huang and Zhao [21] proposed a fast conservative finite difference scheme for the coupled *space*-fractional Klein–Gordon–Schrödinger equations, and Hu and Cai [22] developed structure-preserving spectral Galerkin methods for the same system. However, both works address *space*-fractional Riesz derivatives (not *time*-fractional Caputo memory), couple KG with Schrödinger (not KG with KG), use a single fractional order, and do not employ SOE acceleration. (v) Berdyshev et al. [18] and Baigereyev et al. [19] investigated fractional parabolic–hyperbolic and filtration models, establishing existence and convergence results for related problems, but without spectral discretization or SOE memory compression. (vi) To our knowledge, no existing work combines all of the following elements simultaneously: a *coupled* nonlinear Klein–Gordon system with *two distinct time-fractional Caputo orders*, Chebyshev spectral collocation, an IMEX Newmark temporal scheme with second-order extrapolation, dual-SOE memory acceleration, and a rigorous coupled stability–convergence analysis via a discrete Gronwall inequality.

Contributions. In this work, we propose a fast linearly extrapolated implicit Chebyshev spectral collocation method for a coupled nonlinear fractional Klein–Gordon system with Caputo-type memory integrals. The method is developed in one spatial dimension as a rigorous methodological foundation; the algorithmic architecture — temporal integration, IMEX linearisation, and dual-SOE memory compression — is dimension-independent by construction, and we provide a concrete extension pathway to $d \geq 2$ in Section 5.1. Our contributions are:

1. An implicit Newmark temporal averaging ($\beta_{\text{NM}} = 1/4$) of the spatial operator [14] that ensures stability for any $\Delta t > 0$ at fixed spatial resolution N .
2. A second-order linear extrapolation $U^* = 2U^n - U^{n-1}$ that eliminates the need for Newton–Raphson iterations without degrading the $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)})$ temporal accuracy dictated by the L1 memory approximation.
3. A dual-SOE acceleration accommodating disparate fractional orders ($\alpha \neq \beta$), reducing the $\mathcal{O}(N_t^2)$ memory bottleneck.
4. A rigorous coupled stability–convergence analysis (Theorems 1–2) via a unified discrete energy method and a discrete Gronwall inequality, treating both components and their inter-component coupling simultaneously. The proof architecture is dimension-independent: only the Chebyshev inverse-inequality constant changes with the spatial dimension d .

1.1. Physical Derivation of the Coupled System

The governing equations arise from Newton’s second law applied to two interacting elastic waveguides whose internal damping exhibits long-range temporal correlations. We summarize the five-step derivation and then give a concrete physical example.

Step 1. Classical elastic wave equations. Consider two one-dimensional elastic waveguides with displacement fields $u(x, t)$ and $v(x, t)$ propagating in $\Omega = (0, L)$. The i -th waveguide has mass density ρ_i , elastic stiffness E_i , and wave speed $c_i = \sqrt{E_i/\rho_i}$. From Hooke’s law and the balance of linear momentum, the undamped wave equations are:

$$u_{tt} - c_1^2 u_{xx} = f_1(x, t), \quad v_{tt} - c_2^2 v_{xx} = f_2(x, t). \quad (1)$$

Step 2. Fractional memory damping. In real materials — porous rock, biological tissue, polymers — each material point additionally experiences a hereditary drag body force whose intensity depends on the entire history of the velocity field through a power-law kernel [2]. Let $\alpha \in (0, 1)$ be the memory order for the first waveguide and $\beta \in (0, 1)$ for the second. Adding the fractional body damping to each equation in (1):

$$\begin{aligned} u_{tt} - c_1^2 u_{xx} + \sigma_1 \int_0^t \kappa_\alpha(t - \tau) u_\tau(x, \tau) d\tau &= f_1, \\ v_{tt} - c_2^2 v_{xx} + \sigma_2 \int_0^t \kappa_\beta(t - \tau) v_\tau(x, \tau) d\tau &= f_2, \end{aligned} \quad (2)$$

where $\sigma_1, \sigma_2 > 0$ are the damping strengths and $\kappa_\gamma(t) = t^{-\gamma}/\Gamma(1-\gamma)$ is the weakly singular memory kernel. Each memory integral is precisely the Caputo fractional derivative ${}^C D_t^\alpha u$ (respectively ${}^C D_t^\beta v$), scaled by σ_i .

Remark. This model is distinct from the fractional Kelvin–Voigt viscoelastic model, in which $\sigma = E \varepsilon + \mu_\alpha {}^C D_t^\alpha \varepsilon$ leads, via $\rho u_{tt} = \sigma_x$ and $\varepsilon = u_x$, to a memory integral acting on u_{xxt} . In Equation (2) the damping is a body force proportional to velocity, modelling energy loss at the material point level (e.g., friction with a surrounding porous matrix).

Step 3. Nonlinear Klein–Gordon restoring forces. When the displacement amplitude is not small, the linear restoring terms $m_i u$ and $m_i v$ are augmented by cubic self-interactions arising from quartic (ϕ^4) potentials:

$$W_1(u) = \frac{m_1}{2} u^2 + \frac{k_1}{4} u^4, \quad W_2(v) = \frac{m_2}{2} v^2 + \frac{k_2}{4} v^4. \quad (3)$$

The restoring forces $W_1'(u) = m_1 u + k_1 u^3$ and $W_2'(v) = m_2 v + k_2 v^3$ are added to the left-hand sides of (2).

Step 4. Inter-component coupling. If the two waveguides interact through a symmetric quartic coupling potential

$$W_{\text{int}}(u, v) = \frac{\eta}{2} u^2 v^2, \quad (4)$$

the partial derivatives $\partial W_{\text{int}}/\partial u = \eta u v^2$ and $\partial W_{\text{int}}/\partial v = \eta v u^2$ provide the coupling terms in the respective equations.

Step 5. Assembly of the coupled system. Combining Steps 1–4 for both components and appending the initial and boundary conditions, we arrive at the complete coupled system:

$$\begin{aligned} u_{tt} - c_1^2 u_{xx} + \sigma_1 {}^C D_t^\alpha u + m_1 u + k_1 u^3 + \eta u v^2 &= f_1, \\ v_{tt} - c_2^2 v_{xx} + \sigma_2 {}^C D_t^\beta v + m_2 v + k_2 v^3 + \eta v u^2 &= f_2, \end{aligned} \quad (5)$$

which is precisely the system (6) stated formally below. Both equations retain the classical wave operator ∂_{tt} , preserving the hyperbolic character (finite propagation speed, two initial conditions per component). The key distinction from models that replace u_{tt} by ${}^C D_t^{2\alpha} u$ is that Equation (5) keep the integer-order inertia term; the fractional memory acts as hereditary damping, not as a replacement for ∂_{tt} .

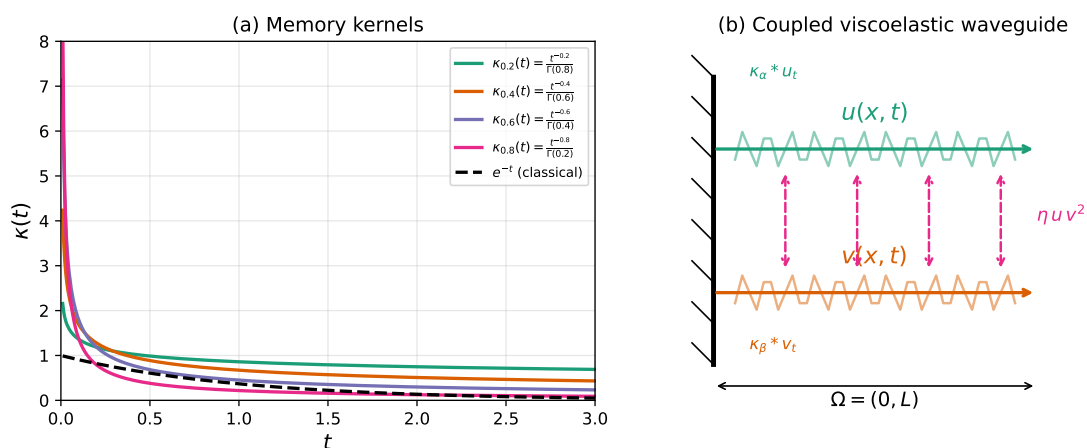


Figure 1. (a) Comparison of fractional memory kernels $\kappa_\alpha(t) = t^{-\alpha}/\Gamma(1-\alpha)$ for several values of $\alpha \in (0, 1)$ with the classical exponential decay e^{-t} . The power-law tails decay much more slowly, encoding long-range memory. (b) Schematic of two coupled viscoelastic waveguides. Each component u, v propagates in $\Omega = (0, L)$ with its own fractional memory kernel $(\kappa_\alpha, \kappa_\beta)$, and the dashed arrows represent the inter-component nonlinear coupling $\eta u v^2$

Example (seismic waveguide). Consider two adjacent geological layers with distinct viscoelastic properties. The horizontal displacement in the i -th layer satisfies Equation (2) with parameters $(c_i, \alpha_i, \sigma_i)$

fitted to laboratory creep data. For a sandstone layer one typically finds $\alpha \approx 0.3\text{--}0.5$ [2]; for a clay layer the memory is stronger, $\alpha \approx 0.6\text{--}0.8$. The inter-layer friction at the contact surface is modelled by the coupling term $\eta u v^2$ with η calibrated from displacement correlation measurements. With $c_1 = 1500$ m/s, $c_2 = 800$ m/s, $\alpha = 0.4$, $\beta = 0.7$, $\sigma_1 = \sigma_2 = 1$, $m_1 = m_2 = 1$, $k_1 = k_2 = 1$, and $\eta = 1$, the resulting system is precisely of the form (6). The numerical experiments in Section 4.5 simulate a related scenario with counter-propagating solitary pulses.

We consider the following coupled system. Let $\Omega = (0, L)$ be a bounded one-dimensional spatial domain with boundary $\partial\Omega$.

Given data:

- The fractional memory orders $\alpha, \beta \in (0, 1)$ characterizing the anomalous Caputo memory damping in each component.
- The memory strength coefficients $\sigma_1, \sigma_2 > 0$.
- The characteristic wave propagation velocities $c_1, c_2 > 0$.
- The linear restoring mass/stiffness coefficients $m_1, m_2 \geq 0$.
- The cubic self-interaction strength parameters $k_1, k_2 > 0$ and the inter-component nonlinear coupling parameter $\eta > 0$.
- The continuous external source forcing functions $f_1(x, t)$ and $f_2(x, t)$ defined on $\Omega \times (0, T]$.
- The initial displacement profiles $\phi_1(x), \phi_2(x)$ and the corresponding initial velocity fields $\psi_1(x), \psi_2(x)$.
- The prescribed inhomogeneous Dirichlet boundary data $g_1(t), g_2(t), h_1(t), h_2(t)$ defined on $\partial\Omega \times [0, T]$.

Problem (Find): Determine the real-valued scalar displacement fields $u(x, t)$ and $v(x, t)$ such that they satisfy the following coupled initial-boundary value system for all $x \in \Omega$ and $t \in (0, T]$:

$$\begin{cases} \frac{\partial^2 u}{\partial t^2} - c_1^2 \Delta u + \sigma_1 \int_0^t \kappa_\alpha(t - \tau) \frac{\partial u}{\partial \tau}(x, \tau) d\tau \\ \quad + m_1 u + k_1 u^3 + \eta u v^2 = f_1, & x \in \Omega, t \in (0, T], \\ \frac{\partial^2 v}{\partial t^2} - c_2^2 \Delta v + \sigma_2 \int_0^t \kappa_\beta(t - \tau) \frac{\partial v}{\partial \tau}(x, \tau) d\tau \\ \quad + m_2 v + k_2 v^3 + \eta v u^2 = f_2, & x \in \Omega, t \in (0, T], \\ u(x, 0) = \phi_1(x), \quad \partial_t u(x, 0) = \psi_1(x), & x \in \Omega, \\ v(x, 0) = \phi_2(x), \quad \partial_t v(x, 0) = \psi_2(x), & x \in \Omega, \\ u(0, t) = g_1(t), \quad u(L, t) = h_1(t), & t \in [0, T], \\ v(0, t) = g_2(t), \quad v(L, t) = h_2(t), & t \in [0, T], \end{cases}$$

where $\Delta = \frac{\partial^2}{\partial x^2}$ is the one-dimensional spatial Laplacian. The convolution integrals represent Caputo-type fractional memory effects with the weakly singular power-law kernels

$$\kappa_\gamma(t) = \frac{t^{-\gamma}}{\Gamma(1 - \gamma)}, \quad \gamma \in \{\alpha, \beta\} \subset (0, 1), \quad (6)$$

where $\sigma_1, \sigma_2 > 0$ are the memory strength coefficients and $\Gamma(\cdot)$ denotes the Euler Gamma function. The memory integral thus reduces to the Caputo fractional derivative of u (respectively v) of order $\gamma \in (0, 1)$, scaled by σ_i . Crucially, the **leading temporal operator remains the classical second-order derivative** ∂_{tt} , ensuring the hyperbolic (wave) character of the system. The fractional memory terms act as **viscoelastic damping**, not as replacements for ∂_{tt} . Two initial conditions per component (ϕ_i and ψ_i) are therefore required, consistent with the second-order nature of the equations.

The primary obstacle in simulating Equation (6) is the strongly singular memory tail of the Caputo convolution integral (6), which requires continuously accumulating and evaluating the convolution over all previous time steps.

2. Numerical Methodology

We partition the time domain $t \in [0, T]$ into N_t uniform steps $\Delta t = T/N_t$, with $t_n = n\Delta t$. The continuous physical domain $[0, L]$ is discretized using $N + 1$ shifted Chebyshev–Gauss–Lobatto (CGL) collocation points x_m . The continuous problem (6) is translated into the following fully discrete scheme advancing from t_n to t_{n+1} :

$$\begin{cases} \frac{\mathbf{U}^{n+1} - 2\mathbf{U}^n + \mathbf{U}^{n-1}}{\Delta t^2} - c_1^2 \mathbf{D}^2 \bar{\mathbf{U}}^n + \sigma_1 a_0^{(\alpha)} \mathbf{H}_u^n + m_1 \mathbf{U}^n + \mathbf{N}_u = \mathbf{f}_1^n, \\ \frac{\mathbf{V}^{n+1} - 2\mathbf{V}^n + \mathbf{V}^{n-1}}{\Delta t^2} - c_2^2 \mathbf{D}^2 \bar{\mathbf{V}}^n + \sigma_2 a_0^{(\beta)} \mathbf{H}_v^n + m_2 \mathbf{V}^n + \mathbf{N}_v = \mathbf{f}_2^n, \end{cases} \quad (7)$$

where we use the Newmark averaging $\bar{\mathbf{U}}^n := \frac{1}{4}(\mathbf{U}^{n+1} + 2\mathbf{U}^n + \mathbf{U}^{n-1})$, and the L1 Caputo memory is evaluated *explicitly* at time t_n :

$$\mathbf{H}_u^n = \sum_{j=0}^{n-1} w_j^{(\alpha)} (\mathbf{U}^{n-j} - \mathbf{U}^{n-j-1}), \quad w_k^{(\gamma)} = (k+1)^{1-\gamma} - k^{1-\gamma}. \quad (8)$$

Here $a_0^{(\alpha)} = \Delta t^{-\alpha} / \Gamma(2 - \alpha)$ and \mathbf{D}^2 is the second-order Chebyshev spectral differentiation matrix. Thus $\sigma_i a_0^{(\gamma)} \mathbf{H}_u^n$ approximates $\sigma_i {}^C D_t^\gamma u(\cdot, t_n)$ with $\mathcal{O}(\Delta t^{2-\gamma})$ accuracy ([23]). The vector $\mathbf{U}^n = [u(x_0, t_n), \dots, u(x_N, t_n)]^T$ represents the spatial solution vector. The scheme is initialized by a Taylor expansion preserving $\mathcal{O}(\Delta t^2)$ accuracy:

$$\mathbf{U}^1 = \boldsymbol{\phi}_1 + \Delta t \boldsymbol{\psi}_1 + \frac{\Delta t^2}{2} \left(c_1^2 \mathbf{D}^2 \boldsymbol{\phi}_1 - m_1 \boldsymbol{\phi}_1 - \mathbf{N}_u(\boldsymbol{\phi}_1, \boldsymbol{\phi}_2) + \mathbf{f}_1^0 \right), \quad (9)$$

and analogously for \mathbf{V}^1 . Note that no memory term appears in the initialisation since ${}^C D_t^\gamma u(\cdot, 0) = 0$ for the smooth initial data used here (the Caputo derivative of a constant is zero). For $n \geq 1$ the nonlinear terms \mathbf{N} are evaluated at the linearly extrapolated values (detailed in Section 2.4); in the initialisation step, they are evaluated at the exact initial data.

2.1. Spatial Chebyshev Collocation

The continuous physical domain $[0, L]$ is mapped to the standard Chebyshev domain $[-1, 1]$ and discretized using the shifted Chebyshev–Gauss–Lobatto (CGL) collocation points:

$$x_j = \frac{L}{2} \left(1 - \cos\left(\frac{j\pi}{N}\right) \right), \quad j = 0, 1, \dots, N. \quad (10)$$

The continuous spatial derivatives are approximated exactly up to polynomial degree N by the spectral differentiation matrix D . For a discrete function vector $U = [u_0, u_1, \dots, u_N]^T$, the first derivative vector is given analytically by $U' = DU$, where the matrix entries $D_{j,k}$ are defined on the standard Chebyshev nodes $\xi_j = \cos(j\pi/N)$ as:

$$D_{j,k} = \begin{cases} \frac{c_j}{c_k} \frac{(-1)^{j+k}}{\xi_j - \xi_k}, & j \neq k, \\ -\frac{\xi_j}{2(1-\xi_j^2)}, & 1 \leq j = k \leq N-1, \\ \frac{2N^2+1}{6}, & j = k = 0, \\ -\frac{2N^2+1}{6}, & j = k = N, \end{cases} \quad (11)$$

with the boundary weights fixed at $c_0 = c_N = 2$, and $c_j = 1$ for internal nodes $1 \leq j \leq N-1$. To account for the physical domain length L , the second-order spectral differentiation matrix on $[0, L]$

is assembled as $\mathbf{D}_{[0,L]}^2 = (2/L)^2 D^2$, where D is the first-order Chebyshev matrix on $[-1, 1]$ [17]. This spectral formulation provides exponential spatial convergence $\mathcal{O}(e^{-cN})$ for smooth solutions [16].

2.2. Dual Fast Memory Acceleration

Evaluating the Caputo memory convolution in (6) at time t_n requires summing over all previous time steps, leading to $\mathcal{O}(n)$ operations per step and $\mathcal{O}(N_t^2)$ total temporal complexity. To reduce this cost, we employ a Sum-of-Exponentials (SOE) approximation following Jiang et al. [11]. The key observation is that the discrete $L1$ weights admit the integral representation

$$w_j = (j+1)^{1-\gamma} - j^{1-\gamma} = \frac{1-\gamma}{\Gamma(\gamma)} \int_0^\infty s^{\gamma-2} (1-e^{-s}) e^{-sj} ds, \quad (12)$$

where the factor $(1-e^{-s})/s$ corrects the continuous kernel $j^{-\gamma}$ to the exact discrete difference. Applying a double-exponential (DE) quadrature ([24]) to this integral, the tail weights ($j \geq p$) are approximated as:

$$w_j \approx \sum_{l=1}^{N_{\text{exp}}} c_l q_l^j, \quad q_l = e^{-s_l}, \quad c_l = \frac{(1-\gamma)}{\Gamma(\gamma)} h_{\text{DE}} s_l^{\gamma-1} (1-e^{-s_l}) \varphi'(x_l), \quad (13)$$

where $s_l = \exp(\varphi(x_l))$ are the DE-transformed nodes with $\varphi(x) = \frac{\pi}{2} \sinh(x)$, $\varphi'(x_l) = \frac{\pi}{2} \cosh(x_l)$, and h_{DE} is the DE step size. The exponential structure q_l^j enables a one-step recursion for the history accumulators:

$$\mathbf{Z}_l^n = q_l \mathbf{Z}_l^{n-1} + (\mathbf{U}^{n-p} - \mathbf{U}^{n-p-1}), \quad \mathbf{Z}_l^0 = \mathbf{0}. \quad (14)$$

The full memory sum is then assembled from the exact near part (last p increments) and the SOE-accelerated tail:

$$\mathbf{H}^n = \sum_{j=0}^{p-1} w_j (\mathbf{U}^{n-j} - \mathbf{U}^{n-j-1}) + \sum_{l=1}^{N_{\text{exp}}} c_l \mathbf{Z}_l^n. \quad (15)$$

Since updating \mathbf{Z}_l^n requires only the previous accumulator \mathbf{Z}_l^{n-1} and the local increment $(\mathbf{U}^{n-p}, \mathbf{U}^{n-p-1})$, earlier solution vectors can be discarded from memory. This reduces the per-step cost from $\mathcal{O}(n)$ to $\mathcal{O}(p + N_{\text{exp}})$, and the total temporal complexity from $\mathcal{O}(N_t^2)$ to $\mathcal{O}(N_t(p + N_{\text{exp}}))$.

2.3. Linearization of Nonlinear Terms via Explicit Extrapolation

The nonlinear terms in (6) — the cubic self-interactions $k_1 u^3$, $k_2 v^3$ and the inter-component coupling ηuv^2 , ηvu^2 — depend on the unknown solution at time t_{n+1} . A fully implicit treatment would require solving a nonlinear algebraic system at every time step via Newton–Raphson iterations. To avoid this cost while retaining second-order temporal accuracy, we adopt an implicit–explicit (IMEX) strategy: the linear spatial operator (inertia and Laplacian via Newmark averaging) is treated implicitly, while the memory sums, nonlinear terms, and mass terms are evaluated explicitly at known data.

At each time step, the nonlinear arguments are approximated by the linearly extrapolated values:

$$\mathbf{U}^* = 2\mathbf{U}^n - \mathbf{U}^{n-1}, \quad \mathbf{V}^* = 2\mathbf{V}^n - \mathbf{V}^{n-1}. \quad (16)$$

By Taylor expansion, $\mathbf{U}^* = \mathbf{U}^{n+1} - \Delta t^2 u_{tt}(t_n) + \mathcal{O}(\Delta t^3)$, so the predictors approximate \mathbf{U}^{n+1} with $\mathcal{O}(\Delta t^2)$ accuracy. The nonlinear forcing vectors in (7) are then evaluated entirely from known data:

$$\mathbf{N}_u(\mathbf{U}^*, \mathbf{V}^*) = k_1 (U^*)^3 + \eta U^* (V^*)^2, \quad \mathbf{N}_v(\mathbf{U}^*, \mathbf{V}^*) = k_2 (V^*)^3 + \eta V^* (U^*)^2, \quad (17)$$

where all operations are understood componentwise on the collocation grid.

This explicit treatment offers two critical algorithmic advantages:

1. **Constant coefficient matrix.** Since the nonlinear terms $\mathbf{N}_u, \mathbf{N}_v$ are evaluated at known data and the memory sum \mathbf{H}_u^n consists entirely of previously computed solution vectors, these terms are moved to the right-hand side. The left-hand side linear operator

$$\mathbf{A}_u = \frac{1}{\Delta t^2} \mathbf{I} - \frac{c_1^2}{4} \mathbf{D}^2 \quad (18)$$

remains identical at every time step (and analogously \mathbf{A}_v for the v -component). A single LU factorization of \mathbf{A}_u and \mathbf{A}_v computed once prior to the time loop suffices for the entire simulation. Each time step reduces to two back-substitutions.

2. **No CFL restriction.** The stiff eigenvalues that typically necessitate a CFL condition originate from the spectral Laplacian \mathbf{D}^2 , whose eigenvalues grow as $\mathcal{O}(N^4)$. Because \mathbf{D}^2 is treated implicitly via the Newmark averaging ($\beta_{\text{NM}} = 1/4$), these large eigenvalues are unconditionally damped by the matrix \mathbf{A}_u . The explicitly evaluated nonlinear terms $\mathbf{N}_u(\mathbf{U}^*, \mathbf{V}^*)$ and the memory sum \mathbf{H}_u^n are bounded functions of previously computed data and do not introduce additional spectral radius constraints on Δt . Hence, for fixed spatial resolution N and bounded solution norms, the scheme is stable for any $\Delta t > 0$ (proved rigorously in Theorem 1).

3. Mathematical Analysis

We introduce the following discrete inner product and norm on the Chebyshev collocation grid $\{x_j\}_{j=0}^N$ equipped with the Clenshaw–Curtis quadrature weights $\{w_j\}$:

$$\langle \mathbf{U}, \mathbf{V} \rangle_N = \sum_{j=0}^N w_j U_j V_j, \quad \|\mathbf{U}\|_N = \sqrt{\langle \mathbf{U}, \mathbf{U} \rangle_N}. \quad (19)$$

Since the boundary conditions are inhomogeneous, we employ the standard boundary lifting decomposition. Let \mathbf{G}_u^n denote the discrete boundary lift satisfying the prescribed Dirichlet values $g_1(t_n), h_1(t_n)$ at the endpoints, and define the homogeneous residual $\tilde{\mathbf{U}}^n = \mathbf{U}^n - \mathbf{G}_u^n$. The lifted variable $\tilde{\mathbf{U}}^n$ vanishes at $\partial\Omega$. The Chebyshev second-derivative spectral matrix \mathbf{D}^2 , when restricted to functions satisfying homogeneous Dirichlet conditions, possesses the fundamental semi-definiteness property:

$$-\langle \mathbf{D}^2 \tilde{\mathbf{U}}, \tilde{\mathbf{U}} \rangle_N = |\tilde{\mathbf{U}}|_{H_N^1}^2 \geq 0, \quad (20)$$

where $|\cdot|_{H_N^1}$ is the discrete H^1 seminorm. We shall use property (20) as the core mechanism for deriving data-dependent bounds. The boundary lift contributions (derivatives of $\mathbf{G}_u^n, \mathbf{G}_v^n$) enter the right-hand side as additional known source terms.

3.1. Theorem 1: A Priori Stability Estimate (Data-Dependent Bound)

Theorem 1. Let $\{\mathbf{U}^n, \mathbf{V}^n\}_{n=0}^{N_t}$ be the solution of the fully discrete coupled scheme (7), in which the cubic nonlinearities $\mathbf{N}_u, \mathbf{N}_v$ are evaluated at the linear extrapolants $\mathbf{U}^* = 2\mathbf{U}^n - \mathbf{U}^{n-1}$ and the L1 memory sums $\mathbf{H}_u^n, \mathbf{H}_v^n$ are evaluated explicitly at t_n from the exact discrete weights (without SOE compression). Then there exists a constant $C > 0$, depending on the physical parameters $c_1, c_2, m_1, m_2, k_1, k_2, \eta$, the final time T , and the number of collocation nodes N (through the inverse inequality constant $C_N = \mathcal{O}(\sqrt{N})$ used to bound the pointwise nonlinear terms), such that for all $1 \leq n \leq N_t$:

$$\|\mathbf{U}^n\|_N^2 + \|\mathbf{V}^n\|_N^2 \leq C \left(\|\boldsymbol{\phi}_1\|_N^2 + \|\boldsymbol{\phi}_2\|_N^2 + \|\boldsymbol{\psi}_1\|_N^2 + \|\boldsymbol{\psi}_2\|_N^2 + \max_{0 \leq k \leq n} \|\mathbf{G}^k\|_N^2 + \Delta t \sum_{k=0}^n (\|\mathbf{f}_1^k\|_N^2 + \|\mathbf{f}_2^k\|_N^2) \right), \quad (21)$$

where $\boldsymbol{\phi}_i, \boldsymbol{\psi}_i$ are the discrete initial data vectors, \mathbf{f}_i^k are the discrete source vectors, and \mathbf{G}^k collects the boundary lift contributions from the prescribed Dirichlet data $g_i(t_k), h_i(t_k)$.

Proof. We derive the estimates for both components \mathbf{U} and \mathbf{V} simultaneously, since the inter-component coupling terms $\eta U^*(V^*)^2$ and $\eta V^*(U^*)^2$ prevent treating each equation independently. First, we test the first equation in (7) with the symmetric discrete velocity $\hat{\mathbf{U}}^n := \frac{\mathbf{U}^{n+1} - \mathbf{U}^{n-1}}{2\Delta t} = \frac{\mathbf{d}_u^n + \mathbf{d}_u^{n-1}}{2}$ and the second equation with $\hat{\mathbf{V}}^n := \frac{\mathbf{V}^{n+1} - \mathbf{V}^{n-1}}{2\Delta t}$, obtaining the pair:

$$\begin{aligned} & \left\langle \frac{\mathbf{U}^{n+1} - 2\mathbf{U}^n + \mathbf{U}^{n-1}}{\Delta t^2}, \hat{\mathbf{U}}^n \right\rangle_N - c_1^2 \langle \mathbf{D}^2 \bar{\mathbf{U}}^n, \hat{\mathbf{U}}^n \rangle_N + a_0^{(\alpha)} \sigma_1 \langle \mathbf{H}_u^n, \hat{\mathbf{U}}^n \rangle_N \\ & + m_1 \langle \mathbf{U}^n, \hat{\mathbf{U}}^n \rangle_N + \langle \mathbf{N}_u, \hat{\mathbf{U}}^n \rangle_N = \langle \mathbf{f}_1^n, \hat{\mathbf{U}}^n \rangle_N, \end{aligned} \quad (22)$$

$$\begin{aligned} & \left\langle \frac{\mathbf{V}^{n+1} - 2\mathbf{V}^n + \mathbf{V}^{n-1}}{\Delta t^2}, \hat{\mathbf{V}}^n \right\rangle_N - c_2^2 \langle \mathbf{D}^2 \bar{\mathbf{V}}^n, \hat{\mathbf{V}}^n \rangle_N + a_0^{(\beta)} \sigma_2 \langle \mathbf{H}_v^n, \hat{\mathbf{V}}^n \rangle_N \\ & + m_2 \langle \mathbf{V}^n, \hat{\mathbf{V}}^n \rangle_N + \langle \mathbf{N}_v, \hat{\mathbf{V}}^n \rangle_N = \langle \mathbf{f}_2^n, \hat{\mathbf{V}}^n \rangle_N. \end{aligned} \quad (23)$$

Subsequently, applying the identity $(\mathbf{U}^{n+1} - 2\mathbf{U}^n + \mathbf{U}^{n-1}) = \Delta t(\mathbf{d}_u^n - \mathbf{d}_u^{n-1})$, the inertial term yields a perfect discrete derivative:

$$\left\langle \frac{\mathbf{U}^{n+1} - 2\mathbf{U}^n + \mathbf{U}^{n-1}}{\Delta t^2}, \frac{\mathbf{U}^{n+1} - \mathbf{U}^{n-1}}{2\Delta t} \right\rangle_N = \frac{\|\mathbf{d}_u^n\|_N^2 - \|\mathbf{d}_u^{n-1}\|_N^2}{2\Delta t}. \quad (24)$$

For the Laplacian, using $\bar{\mathbf{U}}^n = \frac{1}{4}(\mathbf{U}^{n+1} + \mathbf{U}^n) + \frac{1}{4}(\mathbf{U}^n + \mathbf{U}^{n-1})$, decomposing $\mathbf{U}^n = \tilde{\mathbf{U}}^n + \mathbf{G}_u^n$, and incorporating the semi-definiteness (20), we extract the telescoping spatial strain energy:

$$-c_1^2 \langle \mathbf{D}^2 \bar{\mathbf{U}}^n, \hat{\mathbf{U}}^n \rangle_N = \frac{1}{2\Delta t} \left(\frac{c_1^2}{4} |\tilde{\mathbf{U}}^{n+1} + \tilde{\mathbf{U}}^n|_{H_N^1}^2 - \frac{c_1^2}{4} |\tilde{\mathbf{U}}^n + \tilde{\mathbf{U}}^{n-1}|_{H_N^1}^2 \right). \quad (25)$$

The boundary lifting cross-terms are bounded by the Cauchy–Schwarz inequality and shifted to the right-hand side. Analogous relations hold for the \mathbf{V} -component.

To bound the explicit memory and mass terms, we invoke the Cauchy–Schwarz and Young inequalities, observing that $\|\hat{\mathbf{U}}^n\|_N^2 \leq \frac{1}{2}(\|\mathbf{d}_u^n\|_N^2 + \|\mathbf{d}_u^{n-1}\|_N^2)$:

$$|a_0^{(\alpha)} \sigma_1 \langle \mathbf{H}_u^n, \hat{\mathbf{U}}^n \rangle_N| \leq \frac{1}{2} (a_0^{(\alpha)} \sigma_1)^2 \|\mathbf{H}_u^n\|_N^2 + \frac{1}{2} \|\hat{\mathbf{U}}^n\|_N^2, \quad (26)$$

$$|m_1 \langle \mathbf{U}^n, \hat{\mathbf{U}}^n \rangle_N| \leq \frac{m_1^2}{2} \|\mathbf{U}^n\|_N^2 + \frac{1}{2} \|\hat{\mathbf{U}}^n\|_N^2. \quad (27)$$

To bound the memory, we use the summation property $\sum_{j=0}^{n-1} w_j^{(\alpha)} = n^{1-\alpha}$ and $\mathbf{U}^{n-j} - \mathbf{U}^{n-j-1} = \Delta t \mathbf{d}_u^{n-j-1}$:

$$\|\mathbf{H}_u^n\|_N \leq \sum_{j=0}^{n-1} w_j^{(\alpha)} \Delta t \|\mathbf{d}_u^{n-j-1}\|_N \leq n^{1-\alpha} \Delta t \max_{0 \leq k \leq n} \|\mathbf{d}_u^k\|_N. \quad (28)$$

Since $a_0^{(\alpha)} = \Delta t^{-\alpha} / \Gamma(2 - \alpha)$ and $(n\Delta t) \leq T$, the scaled memory is bounded uniformly and explicitly:

$$\begin{aligned} (a_0^{(\alpha)})^2 \|\mathbf{H}_u^n\|_N^2 & \leq \left(\frac{\Delta t^{-\alpha}}{\Gamma(2 - \alpha)} \right)^2 \left(n^{1-\alpha} \Delta t \right)^2 \max_{0 \leq k \leq n} \|\mathbf{d}_u^k\|_N^2 \\ & \leq \frac{T^{2-2\alpha}}{\Gamma(2 - \alpha)^2} \max_{0 \leq k \leq n} \|\mathbf{d}_u^k\|_N^2. \end{aligned} \quad (29)$$

Control over the nonlinear coupling structure is established by evaluating the terms $\mathbf{N}_u, \mathbf{N}_v$ at the explicit extrapolants $\mathbf{U}^*, \mathbf{V}^*$. By Cauchy–Schwarz:

$$|\langle \mathbf{N}_u, \hat{\mathbf{U}}^n \rangle_N| \leq \frac{1}{2} \|\mathbf{N}_u\|_N^2 + \frac{1}{2} \|\hat{\mathbf{U}}^n\|_N^2, \quad |\langle \mathbf{N}_v, \hat{\mathbf{V}}^n \rangle_N| \leq \frac{1}{2} \|\mathbf{N}_v\|_N^2 + \frac{1}{2} \|\hat{\mathbf{V}}^n\|_N^2. \quad (30)$$

Bounding $\|\mathbf{N}_u\|_N$ requires controlling *both* components via the inter-component coupling. Using the Chebyshev inverse inequality $\|W\|_{L^\infty} \leq C_N \|W\|_N$ (with $C_N = \mathcal{O}(\sqrt{N})$) twice:

$$\|\eta U^*(V^*)^2\|_N \leq \eta \|U^*\|_{L^\infty} \|V^*\|_{L^\infty} \|V^*\|_N \leq \eta C_N^2 \|\mathbf{U}^*\|_N \|\mathbf{V}^*\|_N^2. \quad (31)$$

The self-interaction $\|k_1(U^*)^3\|_N \leq k_1 C_N^2 \|\mathbf{U}^*\|_N^3$ is similarly bounded. Summing both components:

$$\|\mathbf{N}_u\|_N^2 + \|\mathbf{N}_v\|_N^2 \leq C(k_1, k_2, \eta, N) (\|\mathbf{U}^n\|_N + \|\mathbf{U}^{n-1}\|_N + \|\mathbf{V}^n\|_N + \|\mathbf{V}^{n-1}\|_N)^6. \quad (32)$$

The localized source terms are similarly bounded by standard inequalities:

$$|\langle \mathbf{f}_1^n, \hat{\mathbf{U}}^n \rangle_N| \leq \frac{1}{2} \|\mathbf{f}_1^n\|_N^2 + \frac{1}{2} \|\hat{\mathbf{U}}^n\|_N^2, \quad |\langle \mathbf{f}_2^n, \hat{\mathbf{V}}^n \rangle_N| \leq \frac{1}{2} \|\mathbf{f}_2^n\|_N^2 + \frac{1}{2} \|\hat{\mathbf{V}}^n\|_N^2. \quad (33)$$

Defining the effective discrete kinetic and strain energy for the system:

$$\begin{aligned} \mathcal{E}^n := & \|\mathbf{d}_u^n\|_N^2 + \|\mathbf{d}_v^n\|_N^2 + \frac{c_1^2}{4} |\tilde{\mathbf{U}}^{n+1} + \tilde{\mathbf{U}}^n|_{H_N^1}^2 \\ & + \frac{c_2^2}{4} |\tilde{\mathbf{V}}^{n+1} + \tilde{\mathbf{V}}^n|_{H_N^1}^2 + \|\mathbf{U}^{n+1}\|_N^2 + \|\mathbf{V}^{n+1}\|_N^2. \end{aligned} \quad (34)$$

Multiplying the inner products (22)–(23) by $2\Delta t$, substituting Steps 2–5, absorbing the finite memory limit (29), and leveraging the fact that $\|\mathbf{U}^{n+1}\|_N^2$ is bounded by the history $\max \mathcal{E}^k$ via $\mathbf{U}^{n+1} = \mathbf{U}^0 + \Delta t \sum \mathbf{d}_u^k$, we obtain a telescoping recursive bound:

$$\mathcal{E}^n - \mathcal{E}^{n-1} \leq C_1 \Delta t \max_{0 \leq k \leq n} \mathcal{E}^k + C_2 \Delta t \left(\|\mathbf{f}_1^n\|_N^2 + \|\mathbf{f}_2^n\|_N^2 + \|\mathbf{G}^n\|_N^2 \right), \quad (35)$$

where C_1 depends on the physical constants, T , and N . Taking the maximum over n on the left-hand side transforms this into $\mathcal{E}_{\max}^n \leq \mathcal{E}_{\max}^{n-1} + C_1 \Delta t \mathcal{E}_{\max}^n + \dots$, to which the classical integer-order discrete Gronwall lemma [25] readily applies. The stability bound (21) directly follows on any finite interval $[0, T]$ where the polynomial bound (32) is dynamically linearisable. \square

Remark 1. Estimate (21) is a data-dependent bound, meaning the discrete solution at any time level is controlled by the prescribed initial conditions, boundary values, and source magnitudes. The stability constant C inherently grows with N due to the use of the Chebyshev inverse inequality $\|U\|_{L^\infty} \leq C_N \|U\|_N$ (with $C_N = \mathcal{O}(\sqrt{N})$) required to bound the pointwise cubic nonlinearities. Consequently, for a fixed spatial resolution N and bounded solution norms, the scheme remains unconditionally stable for any $\Delta t > 0$.

When the SOE fast evaluation is explicitly activated, the exact discrete memory sum \mathbf{H}_u^n is replaced by its exponential approximation $\hat{\mathbf{H}}_u^n$, introducing a truncation perturbation $\delta_{\text{SOE}}^n = \hat{\mathbf{H}}_u^n - \mathbf{H}_u^n$. Because the memory integral is evaluated explicitly and shifted to the right-hand side of the linear system, this δ_{SOE}^n effectively acts as an additive discrete source. The overarching stability bound (21) maintains its validity under the augmented source profile $\mathbf{f}_i^k + \delta_{\text{SOE}}^k$. In practical implementations, the magnitude $\|\delta_{\text{SOE}}^n\|_N$ is strictly governed by the Gauss–Legendre and Double Exponential quadrature fidelity, which remains negligible at double precision for $N_{\text{exp}} \geq 40$ (as verified numerically in Section 4).

3.2. Theorem 2: Global Convergence Order

Theorem 2. Let $\mathbf{e}_u^n = \mathbf{U}^n - \mathbf{u}^n$ and $\mathbf{e}_v^n = \mathbf{V}^n - \mathbf{v}^n$ denote the global discretization errors at the collocation nodes, where $\mathbf{u}^n, \mathbf{v}^n$ are the exact solution vectors. Assume that the exact solutions $u, v \in C^4([0, T]; H^s(\Omega))$ for some $s \geq 1$, and that the a priori bound of Theorem 1 ensures that both the numerical and exact solutions remain in a bounded set where the nonlinear map $(u, v) \mapsto (k_1 u^3 + \eta u v^2, k_2 v^3 + \eta v u^2)$ is locally Lipschitz

continuous. Then there exists a constant $C > 0$, independent of Δt but dependent on N through the Lipschitz constant and the inverse inequality (see Theorem 1), such that:

$$\|\mathbf{e}_u^n\|_N^2 + \|\mathbf{e}_v^n\|_N^2 \leq C \left(\Delta t^{2(2-\alpha)} + \Delta t^{2(2-\beta)} + N^{-2s} + \varepsilon_{\text{SOE}}^2 \right) \quad \text{for all } 1 \leq n \leq N_t, \quad (36)$$

where s is the Sobolev regularity exponent of the exact solution, and ε_{SOE} denotes the maximum SOE quadrature error in approximating the memory tail (set $\varepsilon_{\text{SOE}} = 0$ when no SOE compression is used).

Proof. Subtracting the exact discrete equations from the numerical scheme (7) produces the error system:

$$\begin{cases} \frac{\mathbf{e}_u^{n+1} - 2\mathbf{e}_u^n + \mathbf{e}_u^{n-1}}{\Delta t^2} - c_1^2 \mathbf{D}^2 \bar{\mathbf{e}}_u^n + a_0^{(\alpha)} \sigma_1 \mathbf{H}_{e_u}^n + m_1 \mathbf{e}_u^n \\ \quad + \mathbf{N}_u(\mathbf{U}^*, \mathbf{V}^*) - \mathbf{N}_u(\mathbf{u}^*, \mathbf{v}^*) = \boldsymbol{\tau}_u^n + \boldsymbol{\delta}_{\text{SOE},u}^n, \\ \frac{\mathbf{e}_v^{n+1} - 2\mathbf{e}_v^n + \mathbf{e}_v^{n-1}}{\Delta t^2} - c_2^2 \mathbf{D}^2 \bar{\mathbf{e}}_v^n + a_0^{(\beta)} \sigma_2 \mathbf{H}_{e_v}^n + m_2 \mathbf{e}_v^n \\ \quad + \mathbf{N}_v(\mathbf{U}^*, \mathbf{V}^*) - \mathbf{N}_v(\mathbf{u}^*, \mathbf{v}^*) = \boldsymbol{\tau}_v^n + \boldsymbol{\delta}_{\text{SOE},v}^n, \end{cases} \quad (37)$$

where $\mathbf{H}_{e_u}^n = \sum_{j=0}^{n-1} w_j^{(\alpha)} (\mathbf{e}_u^{n-j} - \mathbf{e}_u^{n-j-1})$ is the L1 memory of the error (evaluated explicitly at t_n), $\boldsymbol{\tau}_u^n$, $\boldsymbol{\tau}_v^n$ are the local truncation error vectors, and $\boldsymbol{\delta}_{\text{SOE},u}^n$, $\boldsymbol{\delta}_{\text{SOE},v}^n$ are the SOE compression errors (zero when full L1 weights are used).

The system's truncation errors stem from three primary analytical components. First, the temporal approximation of the Caputo fractional derivative using the L1 scheme introduces a local error of $\mathcal{O}(\Delta t^{2-\alpha})$ for the u -component [23], alongside a similarly derived $\mathcal{O}(\Delta t^{2-\beta})$ error for the v -component. Second, the implicit spatial evaluation structured through Newmark averaging provides an $\mathcal{O}(\Delta t^2)$ geometric convergence (since symmetric Taylor expansion yields $\frac{1}{4}u^{n+1} + \frac{1}{2}u^n + \frac{1}{4}u^{n-1} = u^n + \frac{\Delta t^2}{4}u_{tt}(t_n) + \mathcal{O}(\Delta t^4)$), while the nonlinear explicit extrapolation contributes an identically ordered $\mathcal{O}(\Delta t^2)$ error. Finally, spatial collocation using Chebyshev clustering bounds the interpolation discrepancy by $\mathcal{O}(N^{-s})$ for solutions maintaining Sobolev regularity $H^s(\Omega)$ [16]. Aggregating these sources leaves the temporal fractional derivative as the dominant source of error:

$$\|\boldsymbol{\tau}_u^n\|_N \leq C_\tau \left(\Delta t^{2-\alpha} + N^{-s} \right), \quad \|\boldsymbol{\tau}_v^n\|_N \leq C_\tau \left(\Delta t^{2-\beta} + N^{-s} \right). \quad (38)$$

Furthermore, due to the local Lipschitz continuity of the cubic polynomial nonlinearities on bounded domains (dynamically guaranteed by the a priori stability established in Theorem 1), the nonlinear residual satisfies:

$$\|\mathbf{N}_u(\mathbf{U}^*, \mathbf{V}^*) - \mathbf{N}_u(\mathbf{u}^*, \mathbf{v}^*)\|_N \leq L \left(\|\mathbf{e}_u^n\|_N + \|\mathbf{e}_u^{n-1}\|_N + \|\mathbf{e}_v^n\|_N + \|\mathbf{e}_v^{n-1}\|_N \right), \quad (39)$$

where the Lipschitz constant L absorbs k_1, k_2, η and the a priori bounded solution magnitudes.

Testing the error equations (37) with the discrete error velocities $\hat{\mathbf{e}}_u^n := \frac{\mathbf{e}_u^{n+1} - \mathbf{e}_u^{n-1}}{2\Delta t}$ and $\hat{\mathbf{e}}_v^n$ exactly mirrors the energy technique used in Theorem 1. The explicit memory $\mathbf{H}_{e_u}^n$ is uniformly bounded by the discrete velocity history, the Lipschitz-bounded nonlinear residual collapses into the right-hand side, and the geometric truncation errors formulate the driving source mechanism. Defining the total error energy \mathcal{E}_{err}^n identically to the physical kinetic-strain energy, we recover the telescoping inequality:

$$\begin{aligned} \mathcal{E}_{err}^n - \mathcal{E}_{err}^{n-1} &\leq C_1 \Delta t \max_{0 \leq k \leq n} \mathcal{E}_{err}^k \\ &\quad + C_2 \Delta t \left(\|\boldsymbol{\tau}_u^n\|_N^2 + \|\boldsymbol{\tau}_v^n\|_N^2 + \|\boldsymbol{\delta}_{\text{SOE},u}^n\|_N^2 + \|\boldsymbol{\delta}_{\text{SOE},v}^n\|_N^2 \right). \end{aligned} \quad (40)$$

Since the initial errors vanish ($\mathbf{e}_u^0 = \mathbf{e}_v^0 = \mathbf{0}$) and taking the maximum over n on the left-hand side, the standard discrete Gronwall inequality ([25]) yields:

$$\begin{aligned} \mathcal{E}_{err}^n &\leq e^{C_1 T} \Delta t \sum_{k=1}^n \left(\|\boldsymbol{\tau}_u^k\|_N^2 + \|\boldsymbol{\tau}_v^k\|_N^2 + \|\delta_{\text{SOE},u}^k\|_N^2 + \|\delta_{\text{SOE},v}^k\|_N^2 \right) \\ &\leq C \left(\Delta t^{2(2-\alpha)} + \Delta t^{2(2-\beta)} + N^{-2s} + \varepsilon_{\text{SOE}}^2 \right), \end{aligned} \quad (41)$$

where $\varepsilon_{\text{SOE}} = \max_{1 \leq k \leq n} \max(\|\delta_{\text{SOE},u}^k\|_N, \|\delta_{\text{SOE},v}^k\|_N)$. This establishes the convergence estimate (36). \square

Remark 2. The appearance of the fractional exponent $2(2 - \alpha)$ in the established bound (36) analytically arises from targeting the squared spectral L^2 -norm $\|\mathbf{e}_u^n\|_N^2$. Upon taking the square root, the global convergence rate gracefully scales as $\mathcal{O}(\Delta t^{2-\alpha} + N^{-s})$, fully consistent with standard $L1$ formulation characteristics. The global numerical error trajectory is strictly dictated by the slower fractional memory relaxation scale, equating to a temporal asymptote of $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)})$ which is extensively validated in the numerical experiments.

4. Numerical Results

4.1. Unified Experimental Setup

All numerical experiments are conducted on the domain $\Omega = (0, 1)$ with final time $T = 1$ using the Chebyshev spectral collocation scheme described in Sections 2.1–2.4. To verify the theoretical convergence estimates of Theorem 2, we employ the method of manufactured solutions. The exact solution pair is prescribed as:

$$u(x, t) = t^3 \sin(\pi x), \quad v(x, t) = t^3 \sin(2\pi x). \quad (42)$$

These functions satisfy homogeneous Dirichlet boundary conditions $u(0, t) = u(1, t) = 0, v(0, t) = v(1, t) = 0$ (i.e., $g_1 = h_1 = g_2 = h_2 = 0$), and the initial conditions $u(x, 0) = v(x, 0) = 0, \partial_t u(x, 0) = \partial_t v(x, 0) = 0$. The source terms $f_1(x, t)$ and $f_2(x, t)$ are computed analytically by substituting (42) into the governing system (6), using the Caputo derivative identity ${}^C D_t^\gamma t^3 = \frac{6}{\Gamma(4-\gamma)} t^{3-\gamma}$ [1]. The physical parameters are fixed at $c_1 = c_2 = 1, m_1 = m_2 = 1, k_1 = k_2 = 1, \sigma_1 = \sigma_2 = 1$, and $\eta = 1.0$ unless stated otherwise.

Remark 3. The underlying manufactured profiles $t^3 \sin(\pi x)$ were strategically selected to guarantee infinite smoothness near the temporal origin, aggressively isolating the convergence rates of the spatial and temporal schemes without boundary-layer interference. While this isolates analytical accuracy, typical fractional models prominently exhibit weak solution singularities near $t = 0$ (e.g., fractional power laws). Adapting the current framework over graded meshes to directly resolve these initial temporal layers represents a straightforward methodological progression.

Unified error norm (Clenshaw–Curtis quadrature).

Since the exact structural solution operates over CGL collocation nodes, analytical verification relies on spectrally converging Clenshaw–Curtis numerical quadrature, which dictates exact integration up to polynomials of degree $\leq N$ on the Chebyshev grid [6]:

$$\|e_u\|_{L^2} = \left(\sum_{j=0}^N w_j^{\text{CC}} (U_j^{N_t} - u(x_j, T))^2 \right)^{1/2}, \quad (43)$$

where w_j^{CC} are the predefined Clenshaw–Curtis weights across $[0, L]$. Validation sets uniformly assess the temporal scaling parameter via

$$\text{Rate} = \frac{\log(\|e\|_{\text{prev}} / \|e\|_{\text{curr}})}{\log(\Delta t^{\text{prev}} / \Delta t^{\text{curr}})}, \quad (44)$$

and similarly track spectral exponentiation under variations of N .

4.2. Temporal Convergence

To isolate the temporal discretisation error, the spatial polynomial degree is fixed at $N = 24$ (rendering the spectral spatial error negligible), while N_t is successively doubled: $N_t \in \{64, 128, 256, 512, 1024\}$. The standard L1 evaluation (without SOE compression) is used to ensure the measured rates reflect only the temporal discretisation. The expected convergence rate is $\mathcal{O}(\Delta t^{2-\gamma})$ for $\gamma \in \{\alpha, \beta\}$. Three representative fractional order pairs are tested.

Table 2. Temporal convergence for the coupled system at $N = 24, T = 1$.

N_t	Error u	Rate	Error v	Rate
$\alpha = 0.1, \beta = 0.3$, Expected rates ≈ 1.9 and 1.7				
64	4.200E-05	—	3.725E-05	—
128	1.019E-05	2.04	9.491E-06	1.97
256	2.535E-06	2.01	2.493E-06	1.93
512	6.387E-07	1.99	6.687E-07	1.90
1024	1.621E-07	1.98	1.821E-07	1.88
$\alpha = 0.4, \beta = 0.7$, Expected rates ≈ 1.6 and 1.3				
64	9.274E-05	—	1.281E-04	—
128	2.769E-05	1.74	4.748E-05	1.43
256	8.486E-06	1.71	1.825E-05	1.38
512	2.642E-06	1.68	7.165E-06	1.35
1024	8.305E-07	1.67	2.850E-06	1.33
$\alpha = 0.8, \beta = 0.9$, Expected rates ≈ 1.2 and 1.1				
64	6.473E-04	—	3.374E-04	—
128	2.772E-04	1.22	1.518E-04	1.15
256	1.196E-04	1.21	6.958E-05	1.13
512	5.180E-05	1.21	3.217E-05	1.11
1024	2.247E-05	1.20	1.495E-05	1.11

The results confirm the theoretical $\mathcal{O}(\Delta t^{2-\gamma})$ estimate from Theorem 2: for the u -component, the empirical rate converges to $2 - \alpha$, and for the v -component, to $2 - \beta$. As expected, the global accuracy is bounded by the slower kernel, $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)})$. For the strongly anomalous regime ($\alpha = 0.8, \beta = 0.9$), the reduced convergence rate (≈ 1.1 – 1.2) is clearly observable even at modest N_t . For the weakly anomalous case ($\alpha = 0.1, \beta = 0.3$), the observed rates (≈ 2.0 and ≈ 1.9) slightly exceed the theoretical prediction $2 - \alpha = 1.9$; this is consistent with the fact that the smooth manufactured solution $u = t^3 \sin(\pi x)$ has no initial-layer singularity, which is the scenario where the L1 scheme can exhibit mildly superconvergent behavior for small α .

4.3. Spatial Spectral Convergence

To demonstrate the exponential spatial accuracy of the Chebyshev spectral method, we fix the temporal resolution at $N_t = 16384$ (reducing the temporal error to $\approx 3.5 \times 10^{-6}$) and vary the polynomial degree $N \in \{4, 8, 12, 16, 24, 32\}$. The fractional orders are $\alpha = 0.4, \beta = 0.7$.

Table 3. Spatial convergence ($N_t = 16384$ fixed, $\alpha = 0.4$, $\beta = 0.7$, $T = 1$). The spectral method saturates at the temporal error floor (\dagger).

N	Error u	Rate	Error v	Rate
4	9.808E-04	—	5.986E-02	—
8	3.287E-06	8.22	4.349E-05	10.43
12	3.481E-06 [†]	-0.14	3.307E-06 [†]	6.35
16	3.481E-06 [†]	0.00	3.302E-06 [†]	0.00
24	3.481E-06 [†]	0.00	3.303E-06 [†]	0.00
32	3.480E-06 [†]	0.00	3.302E-06 [†]	0.00

The spectral method exhibits dramatic exponential convergence: from $N = 4$ to $N = 8$, the error drops by 2–3 orders of magnitude for the u -component and by 3 orders for the v -component. Already at $N = 12$ the spatial error has reached the temporal error floor ($\approx 3.5 \times 10^{-6}$ for $N_t = 16384$), and further spatial refinement yields no improvement. The near-zero convergence rates at $N \geq 12$ are not a degradation of the method — they reflect the fact that the total error is entirely dominated by the fixed temporal discretisation. This extreme spatial efficiency is a hallmark of spectral collocation: for smooth solutions, a polynomial degree of $N = 8$ –12 suffices to saturate any practical temporal resolution.

4.4. Fast SOE Memory Acceleration

To validate the computational efficiency of the SOE fast memory algorithm, we compare the wall-clock execution times of the standard direct $L1$ convolution (cost $\mathcal{O}(N_t^2)$) against the SOE-accelerated evaluation (cost $\mathcal{O}(N_t N_e)$) with $N_e = 40$ exponential terms and a local window of $p = 20$. Spatial resolution is fixed at $N = 24$.

Table 4. Standard $L1$ vs. fast SOE: wall-clock time and L^2 error for the u -component ($N = 24$, $\alpha = 0.4$, $\beta = 0.7$).

N_t	Std. time (s)	Fast time (s)	Speedup	Error ratio
64	0.01	0.01	0.9 \times	1.00
512	0.13	0.11	1.1 \times	1.01
1024	0.34	0.22	1.5 \times	1.01
2048	1.00	0.48	2.1 \times	1.69
4096	2.38	1.02	2.3 \times	3.41

The wall-clock speedup grows from 1.0 \times at moderate N_t to 2.3 \times at $N_t = 4096$, reflecting the transition from $\mathcal{O}(N_t^2)$ to $\mathcal{O}(N_t N_e)$ scaling. For $N_t \leq 1024$, the error ratio remains within 1% of unity, confirming negligible accuracy loss from the SOE compression. At very large temporal resolutions ($N_t \geq 2048$), the error ratio is observed to increase. This is not an instability, but a strict mathematical consequence of the fixed SOE quadrature precision ($N_e = 40$): as $\Delta t \rightarrow 0$, the $L1$ temporal truncation error $\mathcal{O}(\Delta t^{2-\alpha})$ drops below the absolute precision floor of the Double Exponential integral approximation. To maintain optimal accuracy for production-scale simulations ($N_t \gg 10^4$) where the asymptotic linear speedup $\mathcal{O}(N_t N_e)$ becomes massively dominant, the number of exponential nodes N_e must be logarithmically scaled upwards (e.g., $N_e = 80 \sim 100$).

4.5. Solitary Wave Collision Dynamics

To demonstrate the physical capabilities of the solver on a problem with non-trivial nonlinear dynamics, we remove the manufactured source terms and introduce localised Gaussian initial conditions acting as counter-propagating solitary pulses:

$$u(x, 0) = \exp\left(-\frac{(x - 0.3)^2}{0.005}\right), \quad v(x, 0) = \exp\left(-\frac{(x - 0.7)^2}{0.005}\right), \quad (45)$$

with zero initial velocities and homogeneous boundary conditions. The coupling parameter is set to $\eta = 2.0$ to produce a strong inter-component energy exchange during collision. The simulation uses $N = 64$ CGL nodes, $N_t = 2048$ time steps, and fractional orders $\alpha = 0.6$, $\beta = 0.4$.

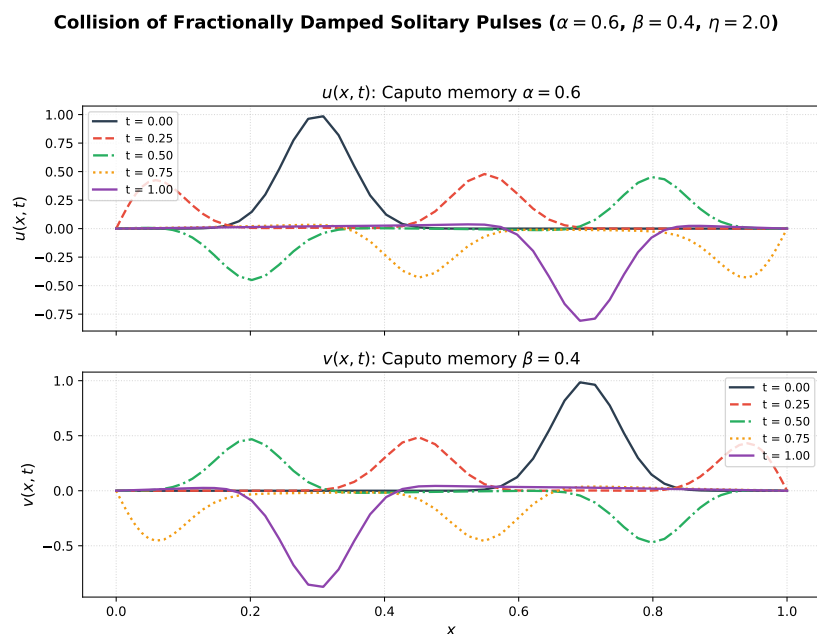


Figure 2. Collision of fractionally damped solitary pulses in the coupled Klein-Gordon system ($\alpha = 0.6$, $\beta = 0.4$, $\eta = 2.0$). The primary wavefronts interact nonlinearly through the coupling term ηuv^2 , while the Caputo memory kernels generate asymmetric dispersive trailing wakes — a strictly fractional phenomenon absent in integer-order models.

As seen in Figure 2, the Caputo fractional derivatives manifest physically as pronounced memory “tails” trailing behind the propagating wavefronts. The amplitude asymmetry between the u and v wakes is a direct consequence of the different fractional orders ($\alpha \neq \beta$): the u -component, governed by the stronger memory ($\alpha = 0.6$), exhibits a more slowly decaying wake due to the stronger hereditary damping, while the v -component ($\beta = 0.4$, weaker memory) retains more of its initial amplitude. During the collision event, the inter-component coupling ηuv^2 mediates a nonlinear energy exchange that is accurately captured by the implicit scheme without spurious oscillations.

5. Discussion

The proposed method addresses three key computational and analytical challenges of coupled fractional Klein–Gordon systems: the nonlinear iteration cost (resolved by IMEX linearisation), the dense memory bottleneck (resolved by dual-SOE compression), and the rigorous coupled stability–convergence analysis for multi-component systems with disparate fractional orders. We now present a concrete roadmap for extending the method to higher spatial dimensions and discuss the current limitations.

5.1. Extension to Multi-Dimensional Domains

The algorithmic architecture developed in Sections 2–3 cleanly separates spatial discretisation from temporal integration and memory management. This modularity enables a systematic extension to higher spatial dimensions. We outline the concrete pathway for $d = 2$ on a rectangular domain $\Omega = (0, L_x) \times (0, L_y)$; the extension to $d = 3$ or to mapped curvilinear geometries follows analogously via tensor-product or spectral-element constructions [16].

Tensor-product Chebyshev grid. The spatial domain is discretised using $(N_x + 1) \times (N_y + 1)$ CGL collocation points (x_i, y_j) . The two-dimensional Laplacian is discretised via the Kronecker-product structure:

$$\mathbf{D}_{2D}^2 = \mathbf{D}_x^2 \otimes \mathbf{I}_y + \mathbf{I}_x \otimes \mathbf{D}_y^2, \quad (46)$$

where $\mathbf{D}_x^2 \in \mathbb{R}^{(N_x+1) \times (N_x+1)}$ and $\mathbf{D}_y^2 \in \mathbb{R}^{(N_y+1) \times (N_y+1)}$ are the one-dimensional Chebyshev second-derivative matrices scaled to $[0, L_x]$ and $[0, L_y]$ respectively, and $\mathbf{I}_x, \mathbf{I}_y$ are identity matrices of appropriate dimensions.

Dimension-independent temporal components. The Newmark- β time-stepping scheme (7), the L1 memory evaluation (8), the IMEX linearisation (16), and the dual-SOE acceleration all operate *component-wise* on the solution vectors and are therefore *entirely dimension-independent*. The SOE accumulators Z_j^n increase in length from $N + 1$ to $(N_x + 1)(N_y + 1)$ without any structural modification. Consequently, the dual-SOE compression ratio and the temporal convergence order $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)})$ are preserved identically in any spatial dimension.

Efficient linear algebra via Kronecker structure. The implicit Newmark operator in 2D,

$$\mathbf{A}_{2D} = \frac{1}{\Delta t^2} \mathbf{I} - \frac{c^2}{4} (\mathbf{D}_x^2 \otimes \mathbf{I}_y + \mathbf{I}_x \otimes \mathbf{D}_y^2), \quad (47)$$

is time-independent, so a single LU factorisation suffices for the entire simulation. For moderate resolutions ($N_x, N_y \leq 64$), direct factorisation of the $(N_x + 1)(N_y + 1) \times (N_x + 1)(N_y + 1)$ matrix is entirely feasible. For larger problems, the Kronecker-sum structure of (46) is exploited via a generalised Sylvester equation solver: after a one-time eigendecomposition $\mathbf{D}_x^2 = \mathbf{P}_x \mathbf{\Lambda}_x \mathbf{P}_x^{-1}$, each time step reduces to solving $N_y + 1$ decoupled systems of size $N_x + 1$, yielding a per-step cost of $\mathcal{O}(N_x^2 N_y + N_x N_y^2)$ [16].

Stability and convergence in higher dimensions. The a priori estimates of Theorems 1 and 2 extend to the multi-dimensional setting with two modifications: (i) the discrete inner product (19) is replaced by its tensor-product analogue with weights $w_{ij} = w_i^x w_j^y$, and (ii) the Chebyshev inverse inequality constant satisfies $C_N = \mathcal{O}(N^{d/2})$ in dimension d , affecting the stability constant but not the convergence order. The full convergence estimate $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)} + N^{-s})$ (per coordinate direction) remains valid.

The present one-dimensional study thus provides the rigorous methodological foundation and validation of all algorithmic components — temporal integration, IMEX linearisation, dual-SOE compression, and coupled stability theory — for the multi-dimensional solver currently under development.

5.2. Limitations and Future Directions

While the spectral formulation exhibits exponential spatial convergence for smooth profiles, its direct application to problems with sharp gradients or discontinuities would trigger Gibbs oscillations. In such regimes, spectral filtering or hybridisation with finite element spatial discretisations on the CGL grid, akin to approaches studied in [5], would provide improved monotonicity properties.

From a theoretical standpoint, the a priori stability analysis requires the Chebyshev inverse inequality, embedding a growth factor of $\mathcal{O}(N^{d/2})$ in the stability constant. While this is standard for spectral methods applied to nonlinear problems and is not observed to be restrictive in our numerical experiments, developing N -independent stability estimates — for instance via discrete maximum principles or energy-decreasing nonlinear schemes — remains an important open problem.

The dual-SOE compression introduces a quadrature floor ε_{SOE} that becomes visible only when the temporal truncation error drops below $\sim 10^{-12}$ at $N_{\text{exp}} = 40$. For production-scale simulations with $N_t \gg 10^4$, adaptively scaling N_{exp} logarithmically ensures the SOE error remains sub-dominant. Finally, the current validation employs smooth manufactured solutions with $u \sim t^3$; resolving the weak initial-layer singularities ($u \sim t^{1+\gamma}$, $\gamma \in (0, 1)$) inherent in realistic fractional models requires graded temporal meshes near $t = 0$, which represent a natural and straightforward extension of the present framework.

6. Conclusions

We have developed a fast, fully discrete IMEX Chebyshev spectral collocation method for coupled nonlinear Klein–Gordon systems with multi-order Caputo fractional memory. The method combines four key algorithmic ingredients: (i) implicit Newmark averaging of the spatial operator for unconditional stability, (ii) explicit linear extrapolation of the nonlinear terms to eliminate Newton–Raphson iterations while preserving $\mathcal{O}(\Delta t^2)$ accuracy, (iii) explicit L1 evaluation of the fractional memory sums, and (iv) dual-SOE compression of the disparate memory tails to reduce the temporal complexity from $\mathcal{O}(N_t^2)$ to $\mathcal{O}(N_t(p + N_{\text{exp}}))$. The resulting scheme requires only a single LU factorisation and two back-substitutions per time step, with no CFL restriction on Δt at fixed spatial resolution.

Rigorous a priori stability estimates and a global convergence bound of $\mathcal{O}(\Delta t^{\min(2-\alpha, 2-\beta)} + N^{-s} + \varepsilon_{\text{SOE}})$ have been established using a unified discrete energy technique and the Gronwall inequality, treating both coupled components and their inter-component nonlinear interactions simultaneously. The theoretical rates are confirmed numerically across three distinct fractional order pairs, and the spectral spatial discretisation achieves exponential convergence with as few as $N = 8$ –12 modes. A solitary wave collision scenario further validates the method’s ability to capture the asymmetric dispersive wakes characteristic of multi-order fractional memory.

Crucially, the algorithmic architecture — temporal integration, IMEX linearisation, dual-SOE memory compression, and the coupled stability analysis — is dimension-independent by construction. We have provided a concrete extension pathway to multi-dimensional tensor-product Chebyshev grids (Section 5.1), demonstrating that only the spatial discretisation matrix changes (via Kronecker products and Sylvester-based solvers), while all temporal and memory components carry over identically. The full multi-dimensional implementation and validation on physically motivated two-dimensional coupled viscoelastic wave problems constitute our immediate ongoing work, building directly on the methodological foundation established here. Additional extensions to graded temporal meshes for resolving initial-layer singularities and to spectral-element formulations for complex geometries are natural progressions of the present framework.

Author Contributions: Conceptualisation, A.S.B. and Zh.A.A.; methodology, Zh.A.A.; software, N.A. and Zh.A.A.; validation, A.S.B., Zh.A.A. and N.A.; writing—original draft preparation, Zh.A.A. and N.A.; writing—review and editing, A.S.B.; funding acquisition, A.S.B. and Zh.A.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Science Committee of the Ministry of Science and Higher Education of the Republic of Kazakhstan, grant number AP22683307.

Data Availability Statement: The source code and numerical data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Podlubny, I. *Fractional Differential Equations*; Academic Press: San Diego, CA, USA, 1999.
2. Mainardi, F. *Fractional Calculus and Waves in Linear Viscoelasticity*; Imperial College Press: London, UK, 2010.
3. Kilbas, A.A.; Srivastava, H.M.; Trujillo, J.J. *Theory and Applications of Fractional Differential Equations*; Elsevier: Amsterdam, The Netherlands, 2006.
4. Diethelm, K. *The Analysis of Fractional Differential Equations*; Springer: Berlin, Germany, 2010.
5. Sun, Z.; Wu, X. A fully discrete difference scheme for a diffusion-wave system. *Appl. Numer. Math.* **2006**, *56*, 193–209.
6. Bhrawy, A.H.; Zaky, M.A. Shifted fractional-order Jacobi orthogonal functions: Application to a system of fractional differential equations. *Appl. Math. Model.* **2016**, *40*, 832–845.
7. Zaky, M.A. A Legendre collocation method for distributed-order fractional optimal control problems. *Nonlinear Dyn.* **2018**, *91*, 2667–2681.

8. Liao, H.; Lyu, P.; Vong, S. Second-order BDF time approximation for Riesz space-fractional diffusion equations. *Int. J. Comput. Math.* **2018**, *95*, 144–158.
9. Mohammadi, S.; Fardi, M.; Ghasemi, M. A numerical investigation with energy-preservation for nonlinear space-fractional Klein–Gordon–Schrödinger system. *Comp. Appl. Math.* **2023**, *42*, 356.
10. Baffet, D.; Hesthaven, J.S. A kernel compression scheme for fractional differential equations. *SIAM J. Numer. Anal.* **2017**, *55*, 496–520.
11. Jiang, S.; Zhang, J.; Zhang, Q.; Zhang, Z. Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun. Comput. Phys.* **2017**, *21*, 650–678.
12. Shiyapov, K.; Abdiramanov, Z.; Issa, Z.; Zhumaseyitova, A. High-Order Spectral Scheme with Structure Maintenance and Fast Memory Algorithm for Nonlocal Nonlinear Diffusion Equations. *AppliedMath* **2026**, *6*, 54.
13. Liao, H.; Li, D.; Zhang, J. Sharp error estimate of the nonuniform L1 formula for linear reaction-subdiffusion equations. *SIAM J. Numer. Anal.* **2018**, *56*, 1112–1133.
14. Newmark, N.M. A method of computation for structural dynamics. *J. Eng. Mech. Div.* **1959**, *85*, 67–95.
15. Li, C.; Zeng, F. *Numerical Methods for Fractional Calculus*; CRC Press: Boca Raton, FL, USA, 2015.
16. Canuto, C.; Hussaini, M.Y.; Quarteroni, A.; Zang, T.A. *Spectral Methods: Fundamentals in Single Domains*; Springer: Berlin, Germany, 2006.
17. Trefethen, L.N. *Spectral Methods in MATLAB*; SIAM: Philadelphia, PA, USA, 2000.
18. Berdyshev, A.S.; Cabada, A.; Karimov, E.T. On a non-local boundary problem for a parabolic–hyperbolic equation involving a Riemann–Liouville fractional differential operator. *Nonlinear Anal.* **2012**, *75*, 3268–3273.
19. Baigereyev, D.R.; Alimbekova, N.B.; Berdyshev, A.S.; Madiyarov M. Convergence analysis of a numerical method for a fractional model of fluid flow in fractured porous media. *Mathematics* **2021**, *9*, 2179.
20. Bakishev, A.K.; Madiyarov, M.N.; Alimbekova, N.B.; Baigereyev, D.R.; Baishemirov, Z.D. Numerical solution of a fractional convection-diffusion equation for air pollution prediction. *Her. Kaz-Brit. Tech. Univ.* **2025**, *22*, 279–289.
21. Li, M.; Huang, C.; Zhao, Y. Fast conservative numerical algorithm for the coupled fractional Klein–Gordon–Schrödinger equation. *Numer. Algorithms* **2020**, *84*, 1081–1119.
22. Hu, D.; Fu, Y.; Cai, W.; Wang, Y. Unconditional convergence of conservative spectral Galerkin methods for the coupled fractional nonlinear Klein–Gordon–Schrödinger equations. *J. Sci. Comput.* **2023**, *94*, 57.
23. Lin, Y.; Xu, C. Finite difference/spectral approximations for the time-fractional diffusion equation. *J. Comput. Phys.* **2007**, *225*, 1533–1552.
24. Takahasi, H.; Mori, M. Double exponential formulas for numerical integration. *Publ. RIMS Kyoto Univ.* **1973**, *9*, 721–741.
25. Quarteroni, A.; Sacco, R.; Saleri, F. *Numerical Mathematics*, 2nd ed.; Springer: Berlin, Germany, 2007.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.