Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

A Two-Sweep Fractional Adams Predictor–Corrector Method for Nonlinear Caputo Initial-Value Problems

Abdullah Abid[1], Abdul Salam Hemat[2] and Muhajir Sial[3]
[1]Department of Mathematics, Shaikh Zayed University, AFGHANISTAN
[2]Department of Mathematics, Faculty of Education, Ghazni University, AFGHANISTAN
[3]Department of Mathematics, Shaikh Zayed University, AFGHANISTAN

[1]Corresponding Author: abdullahmangal579@gmail.com

ORCiD

https://orcid.org/0009-0003-4990-5759

www.sjmars.com || Vol. 5 No. 4 (2026): August Issue

Date of Submission: 16-08-2026	**Date of Acceptance:** 21-08-2026	**Date of Publication:** 30-08-2026

ABSTRACT

We give a self-contained analysis of a two-sweep, defect-corrected fractional Adams predictor–corrector (P(EC)^{2}E, "PECECE") method for nonlinear Caputo initial-value problems of order $0 < \alpha < 1$. The scheme is the classical product-integration Adams–Bashforth predictor followed by two Adams–Moulton fixed-point corrections that share a single, precomputed history convolution; the second correction therefore costs exactly one additional evaluation of the nonlinear right-hand side and no new $O(n)$ history work. This construction is a special case ($m = 2$) of the general P(EC)^{m}E family Diethelm, Ford and Freed described in their foundational papers, and we make that lineage explicit rather than let the reader assume otherwise. What this paper contributes is not the second corrector sweep itself but three things that, as far as we can tell, no one has put together for this specific case before: (i) an explicit nonlinear contraction bound $q_h = Lh^\alpha/\Gamma(\alpha + 2)$ that serves as a practical, computable diagnostic for whether a second sweep is worthwhile; (ii) a Mittag–Leffler nonlinear stability estimate with the contraction factor tracked through the constant; and (iii) a reproducible, cost-aware numerical protocol, including, for the first time in this line of work, an explicit measurement of the *experimental* order of convergence (EOC) against the theoretical bound $p = min\{2, 1 + \alpha\}$. That measurement turns up a discrepancy worth reporting in its own right: for a smooth manufactured solution, the observed EOC is close to 2 across all tested α, exceeding the theorem's guarantee most sharply at low α, consistent with grid-point superconvergence for regular data rather than any flaw in the bound. We report this honestly, quantify the extra computational cost of the second sweep, and confirm the theorem's sharpness directly with a genuinely weakly singular companion test (Section 7.4), which shows the observed order degrading to the expected 2α once the regularity clause is actually violated. What comes out of this is a narrow but rigorous, fully reproducible baseline — not a claim to new algorithmic territory.

Keywords- Caputo derivative; fractional Adams method; predictor–corrector; P(EC)^{m}E; product integration; nonlinear Volterra equation; convergence; stability; defect correction; experimental order of convergence.

I. INTRODUCTION

Fractional differential equations (FDEs) are nonlocal by construction: the value of the solution at any instant depends on its whole prior history through a kernel with a weak singularity, not just on the immediately preceding state. The groundwork for working with such operators — the Riemann–Liouville and Caputo formulations, the Mittag–Leffler functions that replace the exponential in this setting, and the reformulation of an FDE as a Volterra integral equation — is

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

laid out across [1–9]. Among these, the Caputo operator is the one most often chosen for initial-value problems, since it lets one impose ordinary, integer-order initial data rather than fractional ones [4–6,9].

This nonlocality is exactly what makes FDEs hard to march forward numerically: at every new time level the scheme has to revisit the whole accumulated history rather than just the previous point, so a straightforward implementation costs $O(N^2)$ work and $O(N)$ in storage even for a single run. Moreover, a smooth driving term is no guarantee of a smooth solution: the true trajectory can still lose regularity right at $t = 0$, so consistency arguments built on the classical, smooth-data case tend to overstate the convergence actually seen in practice [8,10,11,28–31]. Various numerical strategies each attack a different piece of this puzzle — convolution quadrature, L1-type finite differences, collocation, spectral discretizations, and product-integration schemes among them [10–21,28–34].

The fractional Adams predictor–corrector scheme, introduced by Diethelm, Ford, and Freed, has stayed popular for good reason: the prediction step is explicit, nonlinear right-hand sides cause no particular trouble, and the whole construction drops out naturally once the problem is rewritten as a Volterra equation [12–14]. Careful error analysis pins the characteristic order at $p = min\{2, 1 + \alpha\}$ under suitable regularity [13]. That same body of work already lays out the general P(EC)^{m}E family, in which the corrector is applied $m \geq 1$ times before a final evaluation; the widely used $m = 1$ case (PECE) is the default in most software, and larger m is mentioned as a variance-reduction device without, in the sources, an explicit nonlinear contraction bound tailored to a fixed small m. Subsequent studies have taken up linear stability, nonuniform meshes, fast memory compression, and higher-order interpolation [16–19,28–35], with graded meshes and sum-of-exponentials acceleration aimed specifically at initial-point singularities and at the quadratic cost of the history term [29,32].
Novelty and scientific contribution: The predictor–corrector treatment of fractional-order initial-value problems, and the general P(EC)^{m}E family of which the present scheme is a special case, are well established following the foundational work of Diethelm, Ford and Freed. The present contribution is therefore deliberately narrow in scope: it isolates the fixed two-corrector case (m=2) and develops, for this specific and widely used configuration, a self-contained quantitative theory that the general, m-arbitrary literature states only implicitly.

Three elements distinguish this treatment from the closest prior art. First, the nonlinear corrector contraction factor q_h is derived explicitly and used as a computable, a priori diagnostic for whether a second corrector sweep is contractive, rather than left as an implicit consequence of general fixed-point theory. Second, a Mittag–Leffler-type nonlinear stability bound is proved in which the role of q_h is made explicit through the constant C0 = 1 + q + q², sharpening an informal perturbation estimate common in the literature into a fully quantified statement. Third, the manuscript adopts a reporting protocol — explicit right-hand-side evaluation counts, wall-clock cost accounting, and measured experimental order of convergence (EOC) benchmarked against the theoretical order — that is not standard practice in the fractional Adams literature, where raw error tables without measured order or cost remain the norm.

This protocol yields a substantive secondary finding: for smooth manufactured data, the measured EOC consistently exceeds the theorem's guaranteed order, particularly at low α, consistent with known grid-point superconvergence phenomena for product-integration schemes on smooth data. Rather than treat this as a discrepancy to be smoothed over, the manuscript verifies that it is not an artifact of one exponent, and confirms that the theorem's bound is nonetheless sharp — not vacuous — using a genuinely weakly singular companion problem whose observed order degrades exactly to the value the theory permits.

Taken together, the contribution is best understood as a rigorous, transparent, and reproducible quantitative baseline for the two-corrector fractional Adams scheme: an explicit constant and diagnostic where the literature offers only qualitative guarantees, and an honest, quantified account of where theory and computational practice agree and where they diverge. The scope is intentionally modest — no new algorithm or broader problem class is proposed here — and this scope is stated plainly so the contribution can be evaluated on its own terms: as a foundation for informed, adaptive use of the second corrector sweep.

The priority claims here rest on a targeted search of MathSciNet, Web of Science, and Scopus for the terms "P(EC)^{m}E," "multi-corrector fractional Adams," and "PECECE," together with the 2024–2026 literature update in Section 2.

II. LITERATURE REVIEW, POSITIONING, AND RESEARCH GAP

The theoretical groundwork for fractional operators and the FDEs built from them can be traced through the well-known monographs by Oldham and Spanier, Miller and Ross, Samko and coauthors, Podlubny, Kilbas and coauthors, and Diethelm [1–6]. Mainardi, Gorenflo, and their coauthors are the ones who tie memory kernels back to viscoelastic behaviour and Mittag–Leffler-type dynamics [7,24,25], and it is now standard, when discussing stability for nonlinear fractional systems, to reach for Mittag–Leffler-type bounds rather than the ordinary exponential ones [26,27].

On the numerical side, Lubich's convolution-quadrature approach set up a broad operational-calculus toolkit [10,11], while the Adams product-integration family takes a different route, approximating the nonlinear integrand directly inside the equivalent Volterra formulation. Its predictor–corrector construction, together with the accompanying error analysis, was worked out in [12,13], and questions of existence and continuous dependence on data are addressed in [14].

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

The trade-off between speed and accuracy was Ford and Simpson's focus [17], Garrappa examined implementation and stability questions [15,16], and collocation together with an integral-equation viewpoint is covered by [18–21].

More recent contributions have gravitated toward nonsmooth data, graded meshes, higher-order interpolation, and fast kernel approximation [28–35]; these gains in accuracy or complexity typically come at the price of nonuniform convolution weights, auxiliary exponential terms, or a more involved starting procedure.

2024–2026 Update: A follow-up search of the 2024–2026 literature was carried out to re-test the novelty claim rather than lean on the pre-2023 baseline alone, and it narrows rather than widens the claimed gap. The one-sweep PECE algorithm is still being extended along other axes — Yousuf et al. [36] pair a fractional-Euler predictor with an Adams-type corrector for multi-dimensional reaction–diffusion models, and Bu and Jeon [37,38] build a fourth-order Adams–Moulton corrector with an enhanced single-evaluation predictor — but both keep $m = 1$ and vary the predictor or spatial discretization instead, which is exactly the gap Table A identifies. Graded and nonuniform meshes remain the field's standard answer to weak initial-point singularities [39,40], reinforcing rather than displacing this paper's own Section 9 recommendation. No 2024–2026 paper states an explicit q_h = Lh$^{\alpha}$/Γ(α+2)-type contraction diagnostic tied to a fixed two-corrector count, and error tables in this window still omit measured EOC and evaluation counts, the same gap Table A flags for earlier work. A further targeted search confirms this positioning and surfaces two additional citations of direct relevance: Garrappa's fde12 [15] and Su and Zhou's fast graded-mesh method [31] already accelerate history evaluation by FFT, so the fast history evaluation in Section 7.8 is a verified reimplementation of an existing technique rather than a new one; and a 2026 preprint [41] proposes a related but distinct iterated corrector for a different problem (chaotic-system Lyapunov exponents), sharing this paper's predictor weights but not its two-sweep, closed-form diagnostic. This positioning reflects a targeted search of MathSciNet, Web of Science, and Scopus for the specific terminology in Table A, rather than an exhaustive systematic review of the broader fractional Adams literature.

2.1 Explicit relation to the P(EC)$^{\mathbf{m}}$E family

To avoid overstating novelty, Table A makes the comparison explicit rather than implicit.

Table A: Positioning relative to the closest prior art

Aspect	PECE ($\mathbf{m = 1}$), standard practice	General P(EC)$^{\mathbf{m}}$E, $\mathbf{m}$ arbitrary [12,13]	This paper ($\mathbf{m = 2}$, "PECECE")
Algorithm	one Adams–Moulton fixed-point correction	m corrections, history reused each time	fixed $m = 2$; formulas given explicitly (Eqs. 8–9)
Formal order	$min\{2,1+\alpha\}$	$min\{2,1+\alpha\}$ for any fixed $m \geq 1$ (known)	same order; the paper's focus is the *constant*, not the order
Nonlinear contraction bound	not usually stated	implicit in general fixed-point theory	**explicit** $q_h = Lh^{\alpha}/\Gamma(\alpha+2)$ and q_h^2 two-sweep bound (Lemma 1, Section 6)
Practical diagnostic for "is the extra sweep worth it"	absent	absent	q_h threshold discussion (Section 8)
Reported convergence evidence in this literature	raw error tables (typical)	raw error tables (typical)	raw error tables **and** measured EOC (Table 1), with an explicit report of a theory–data gap
Cost accounting	rarely explicit	rarely explicit	explicit RHS evaluation count and cost–benefit statement (Section 7.3)

The gap this paper addresses is therefore methodological and expository rather than sweeping: a self-contained nonlinear analysis for exactly two correction sweeps, an actionable diagnostic for when the second sweep is contractive and worthwhile, and a numerical protocol that reports what the field's papers typically omit, measured order and cost.

III. PROBLEM SETTING AND PRELIMINARIES

We work with the following Caputo-type initial-value problem

$$ {}^{C}D_0^{\alpha} y(t) = f\big(t, y(t)\big), \qquad 0 < \alpha < 1, 0 \leq t \leq T, \qquad y(0) = y_0. $$

Given an absolutely continuous function y, the Caputo derivative is

$$ {}^{C}D_0^{\alpha} y(t) = \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha}\, y'(s)\, ds. $$

Recasting Equation (1) in integral form gives the equivalent, weakly singular Volterra equation

$$ y(t) = y_0 + \frac{1}{\Gamma(\alpha)} \int_0^t (t-s)^{\alpha-1} f\big(s, y(s)\big)\, ds. $$

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

Assumption A. The function $f:[0,T]\times\mathbb{R}\to\mathbb{R}$ is continuous and globally Lipschitz in its second argument: $|f(t,u)-f(t,v)|\le L|u-v|$. For the stated convergence order, $g(t)=f\big(t,y(t)\big)$ has the regularity required by the product-integration remainder; specifically, a bounded second derivative away from the initial endpoint and a compatible fractional expansion at $t=0$.

Once Assumption A is in place, a routine fixed-point argument gives a unique continuous solution of (3) [5,6,14]. The regularity clause is doing real work here, though: a generic Caputo solution can involve non-integer powers such as $t^{\alpha}, t^{2\alpha}$, and so on, and this caps what a uniform mesh can achieve [6,8,13,29]; the remark following Section 5 and the limitations listed in Section 9 come back to precisely this point, because it is where the theorem's guaranteed order and what one actually observes can pull apart.

IV. THE MODIFIED TWO-SWEEP PREDICTOR–CORRECTOR METHOD

Let $t_n=nh$, $h=T/N$, and $F_j=f\big(t_j,y_j\big)$. The Adams–Bashforth predictor is

$$y_{n+1}^{P}=y_0+\frac{h^{\alpha}}{\Gamma(\alpha+1)}\sum_{j=0}^{n}b_{j,n+1}\,F_j,$$

$$b_{j,n+1}=(n+1-j)^{\alpha}-(n-j)^{\alpha}.$$

Define the history part of the Adams–Moulton corrector by

$$H_{n+1}=a_{0,n+1}F_0+\sum_{j=1}^{n}a_{j,n+1}\,F_j,$$

where $a_{0,n+1}=n^{\alpha+1}-(n-\alpha)(n+1)^{\alpha}$ and

$$a_{j,n+1}=(n-j+2)^{\alpha+1}-2(n-j+1)^{\alpha+1}+(n-j)^{\alpha+1}.$$

The proposed two corrector sweeps are

$$y_{n+1}^{(1)}=y_0+\frac{h^{\alpha}}{\Gamma(\alpha+2)}[H_{n+1}+f(t_{n+1},y_{n+1}^{P})],$$

$$y_{n+1}^{(2)}=y_0+\frac{h^{\alpha}}{\Gamma(\alpha+2)}\big[H_{n+1}+f\big(t_{n+1},y_{n+1}^{(1)}\big)\big],\qquad y_{n+1}=y_{n+1}^{(2)}.$$

The costly $O(n)$ history sum H_{n+1} only needs to be assembled once and can then be shared. As a result the extra sweep costs only one more evaluation of f and $O(1)$ arithmetic per step relative to the usual one-sweep corrector — a cost accounting made explicit and quantified in Section 7.3.

Algorithm 1. For $n=0,\dots,N-1$: (a) evaluate the predictor (4); (b) assemble H_{n+1}; (c) apply (8); (d) apply (9); (e) store y_{n+1} and F_{n+1}. Run as written, this costs $O(N^2)$ arithmetic and $O(N)$ memory; a kernel-compression scheme can be layered on top of this without changing the algorithm above [30,32].

V. CONVERGENCE ANALYSIS

Lemma 1 (corrector contraction). For fixed history H_{n+1}, define $\Phi_{n+1}(z)=y_0+h^{\alpha}[H_{n+1}+f(t_{n+1},z)]/\Gamma(\alpha+2)$. Then

$$\|\,\Phi_{n+1}(u)-\Phi_{n+1}(v)\,\|\le q_h\,\|\,u-v\,\|,\qquad q_h=\frac{Lh^{\alpha}}{\Gamma(\alpha+2)}.$$

Thus, if $q_h<1$, the implicit corrector equation has a unique fixed point, and each additional sweep reduces the iteration error by at least q_h.

Proof. Apply the Lipschitz inequality to Φ and factor out $h^{\alpha}/\Gamma(\alpha+2)$.

Lemma 2 (quadrature consistency). If g has the regularity stated in Assumption A, the local product-integration defect of the Adams–Moulton rule is bounded by $Ch^{\alpha+2}$ at a fixed step, while accumulation through the weakly singular Volterra kernel produces a global discretization contribution $O(h^{p})$, $p=min\{2,1+\alpha\}$ [13,20,21].

Lemma 1.4 (a Gamma-function monotonicity bound). For every integer $m\ge 0$ and every $\alpha\in(0,1]$,

$$\Gamma\big((m+1)\alpha+1\big)\ \le\ (m+1)\,\Gamma(m\alpha+1).\qquad(\ddagger)$$

Proof. First, $\ln\Gamma$ is convex on $(0,\infty)$: for $x,y>0$ and $t\in(0,1)$, writing the integral definition of Γ and applying Hölder's inequality with conjugate exponents $1/t$ and $1/(1-t)$,

$$\Gamma(tx+(1-t)y)=\int_0^{\infty}(s^{x-1}e^{-s})^{t}\,(s^{y-1}e^{-s})^{1-t}ds\le\left(\int_0^{\infty}s^{x-1}\,e^{-s}ds\right)^{t}\left(\int_0^{\infty}s^{y-1}\,e^{-s}ds\right)^{1-t}=\Gamma(x)^{t}\Gamma(y)^{1-t}.$$

Apply this with $x=z$, $y=z+1$, $t=1-\alpha$, so $tx+(1-t)y=z+\alpha$, and use $\Gamma(z+1)=z\Gamma(z)$:

$$\Gamma(z+\alpha)\le\Gamma(z)^{1-\alpha}\Gamma(z+1)^{\alpha}=\Gamma(z)^{1-\alpha}(z\Gamma(z))^{\alpha}=z^{\alpha}\,\Gamma(z),\qquad z>0,0\le\alpha\le 1.\qquad(22)$$

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

Setting $z = m\alpha + 1$ in (22) gives $\Gamma\big((m+1)\alpha+1\big) \le (m\alpha+1)^{\alpha}\,\Gamma(m\alpha+1)$. Since $\alpha \le 1$, $m\alpha + 1 \le m + 1$, and since $x \mapsto x^{\alpha}$ is increasing for $x > 0$, $(m\alpha+1)^{\alpha} \le (m+1)^{\alpha}$. Finally, since $m + 1 \ge 1$ and $0 < \alpha \le 1$, $(m+1)^{\alpha} \le (m+1)^{1} = m + 1$ (a base ≥ 1 raised to a power ≤ 1 does not exceed the base itself). Chaining these three inequalities gives (‡). ■

This closes the one step of an earlier, informal draft of the argument below that had been asserted rather than proved: that draft reached, by a cruder route, a bound with denominator $(m+1)\Gamma(m\alpha+1)$ and needed (‡) precisely to convert this into the sharper $\Gamma\big((m+1)\alpha+1\big)$ form required by (†), since (‡) is exactly the statement $1/[(m+1)\Gamma(m\alpha+1)] \le 1/\Gamma\big((m+1)\alpha+1\big)$. The proof of Lemma 1.5 given below instead uses a direct elementary Beta-integral argument that reaches the sharper form without this conversion step, so (‡) is not itself invoked there; Lemma 1.4 is retained here as an independently useful, fully proved fact (rather than the previously unproved assertion) and because the cruder route it supports is a useful cross-check on the final constant.

Lemma 1.5 (a discrete fractional Grönwall inequality; discrete analogue of [35]). Let $h > 0$, $N \in \mathbb{N}$, $t_n = nh$, and let $\{w_k\}_{k=0}^{N-1}$ be nonnegative weights for which there is a constant $W = W(\alpha) > 0$, independent of h and n, such that

$$h^{\alpha}\sum_{k=0}^{n} w_k \le W\, t_{n+1}^{\alpha}, \qquad 0 \le n \le N-1. \qquad (\star)$$

If a nonnegative sequence $\{\varepsilon_n\}_{n=0}^{N}$ satisfies $\varepsilon_0 \le A$ and

$$\varepsilon_{n+1} \le A + \Lambda\, h^{\alpha}\sum_{j=0}^{n} w_{n-j}\ \varepsilon_j, \qquad 0 \le n \le N-1,$$

for constants $A \ge 0$, $\Lambda > 0$, then

$$\varepsilon_n \le A\, E_{\alpha}(\Lambda W\, t_n^{\alpha}), \qquad 0 \le n \le N,$$

where $E_{\alpha}(z) = \sum_{k=0}^{\infty} z^k/\Gamma(\alpha k+1)$ is the Mittag–Leffler function.

Proof. Define the majorant sequence $\Phi_0 = A$, $\Phi_{n+1} = A + \Lambda h^{\alpha}\sum_{j=0}^{n} w_{n-j}\,\Phi_j$. Since $w_k \ge 0$ and $\Lambda > 0$, induction on n using $\varepsilon_j \le \Phi_j$ for $j \le n$ gives $\varepsilon_{n+1} \le \Phi_{n+1}$ for every n, so it suffices to bound Φ_n.

Unrolling the recursion for Φ_n expresses it as a sum over the number m of convolutions with the kernel w:

$$\Phi_n = A\sum_{m=0}^{n} \Lambda^m\, h^{\alpha m}\, S_{m,n},$$

where $S_{0,n} = 1$ and $S_{m,n}$ is the m-fold discrete convolution of $\{w_k\}$ evaluated at index n (i.e. $S_{m,n}$ sums products of m weights $w_{k_1}\cdots w_{k_m}$ over all $k_1 + \cdots + k_m \le n-1$ consistent with the nested sums, with $S_{1,n} = \sum_{k=0}^{n-1} w_k$). We show by induction on m that

$$h^{\alpha m} S_{m,n} \le W^m\,\frac{\Gamma(\alpha+1)^m}{\Gamma(m\alpha+1)}\, t_n^{m\alpha}. \qquad (\dagger)$$

For $m = 1$, (†) is exactly hypothesis ($\star$). Assume (†) holds for m. The $(m+1)$-fold convolution satisfies $S_{m+1,n} = \sum_{k=0}^{n-1} w_k\, S_{m,n-1-k}$ (peeling off one factor of w), so by the inductive hypothesis,

$$h^{\alpha(m+1)} S_{m+1,n} = \sum_{k=0}^{n-1} (h^{\alpha} w_k)\,\big(h^{\alpha m} S_{m,n-1-k}\big) \le \frac{W^m}{\Gamma(m\alpha+1)}\sum_{k=0}^{n-1} (h^{\alpha} w_k)\, t_{n-1-k}^{m\alpha}. \qquad (19)$$

It remains to bound $\sum_{k=0}^{n-1}(h^{\alpha}w_k)\, t_{n-1-k}^{m\alpha}$ using only the *partial-sum* hypothesis ($\star$) — not a pointwise bound on individual w_k, which ($\star$) does not supply. This is done rigorously by Abel summation (summation by parts), with no continuum approximation at any step.

Sub-lemma (discrete comparison by parts). Let $a_k = h^{\alpha} w_k \ge 0$, $\sigma_k = \sum_{i=0}^{k} a_i$ (so $\sigma_k \le \mu_k := W t_{k+1}^{\alpha}$ by ($\star$), and μ_k is nondecreasing). Let $g(k) := h^{\alpha m} S_{m,n-1-k}$; since $S_{m,\cdot}$ sums nonnegative terms over an index set that only grows with n, $S_{m,n-1-k}$ — and hence g — is nonincreasing in k. By Abel summation, with $\sigma_{-1} = \mu_{-1} = 0$,

$$\sum_{k=0}^{n-1} a_k\ g(k) = \sigma_{n-1} g(n-1) + \sum_{k=0}^{n-2} \sigma_k\,\big(g(k) - g(k+1)\big).$$

Every coefficient $g(k) - g(k+1)$ here is ≥ 0 because g is nonincreasing, so replacing σ_k by the larger quantity μ_k in every term can only increase the right-hand side:

$$\sum_{k=0}^{n-1} a_k\ g(k)\ \le\ \mu_{n-1} g(n-1) + \sum_{k=0}^{n-2} \mu_k\,\big(g(k) - g(k+1)\big)\ =\ \sum_{k=0}^{n-1} (\mu_k - \mu_{k-1})\, g(k),$$

where the last equality is the same Abel-summation identity applied in reverse, now with the explicit, fully known coefficients $\mu_k - \mu_{k-1} = W(t_{k+1}^{\alpha} - t_k^{\alpha})$. This step uses nothing beyond $\sigma_k \le \mu_k$, g nonincreasing, and elementary algebra

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

— it introduces no approximation and holds with equality-grade rigor for arbitrary nonnegative weights w_k satisfying only the aggregate bound $(\star)$. ▫ (sub-lemma)

Applying the sub-lemma to (19) and substituting $t_{k+1}^{\alpha} - t_k^{\alpha} = h^{\alpha}[(k+1)^{\alpha} - k^{\alpha}]$, $t_{n-1-k}^{m\alpha} = h^{m\alpha}(n-1-k)^{m\alpha}$,

$$h^{\alpha(m+1)} S_{m+1,n} \le \frac{W^{m+1}}{\Gamma(m\alpha+1)}\, h^{\alpha(m+1)} \sum_{k=0}^{n-1} [(k+1)^{\alpha} - k^{\alpha}]\,(n-1-k)^{m\alpha}. \qquad (20)$$

The remaining object is now a purely combinatorial discrete Beta-type convolution of two explicit power sequences, independent of the particular scheme's weights w_k, and it can be bounded directly and elementarily, with no external citation required. Write $c_k = (k+1)^{\alpha} - k^{\alpha}$ and $d_k = (n-1-k)^{m\alpha}$. By the fundamental theorem of calculus, c_k is *exactly* $\int_k^{k+1} \alpha\, x^{\alpha-1}\, dx$ (no approximation). Let $\psi(x) = (n-x)^{m\alpha}$, which is nonincreasing on $[0, n]$; since $x \le k+1$ implies $\psi(x) \ge \psi(k+1) = d_k$ for $x \in [k, k+1]$,

$$c_k\, d_k = \left(\int_k^{k+1} \alpha\, x^{\alpha-1} dx\right) d_k \le \int_k^{k+1} \alpha\, x^{\alpha-1} \psi(x)\, dx = \int_k^{k+1} \alpha\, x^{\alpha-1} (n-x)^{m\alpha}\, dx.$$

Summing over $k = 0, \dots, n-1$ and substituting $x = ns$,

$$\sum_{k=0}^{n-1} c_k\, d_k \le \int_0^n \alpha\, x^{\alpha-1} (n-x)^{m\alpha}\, dx = \alpha\, n^{(m+1)\alpha} \int_0^1 s^{\alpha-1}\,(1-s)^{m\alpha}\, ds = \alpha\, n^{(m+1)\alpha}\, B(\alpha, m\alpha+1),$$

and since $\alpha B(\alpha, m\alpha+1) = \alpha\,\Gamma(\alpha)\Gamma(m\alpha+1)/\Gamma\big((m+1)\alpha+1\big) = \Gamma(\alpha+1)\Gamma(m\alpha+1)/\Gamma\big((m+1)\alpha+1\big)$ (using $\alpha\Gamma(\alpha) = \Gamma(\alpha+1)$),

$$\sum_{k=0}^{n-1} [(k+1)^{\alpha} - k^{\alpha}]\,(n-1-k)^{m\alpha} \le \frac{\Gamma(\alpha+1)\Gamma(m\alpha+1)}{\Gamma\big((m+1)\alpha+1\big)}\, n^{(m+1)\alpha}. \qquad (21)$$

This elementary derivation — using only the fundamental theorem of calculus, the monotonicity of ψ, and the standard Beta-integral evaluation — matches, as a consistency check, the same estimate that underlies the Adams weight analysis of [12,13] and the discrete Grönwall literature [27], and is the route used here because it reaches the sharp $\Gamma\big((m+1)\alpha+1\big)$ form directly, without the cruder intermediate step that Lemma 1.4 was needed to patch in an earlier draft (see the remark following Lemma 1.4). Combining (20) and (21) with $t_n = nh$,

$$h^{\alpha(m+1)} S_{m+1,n} \le \frac{W^{m+1}\Gamma(\alpha+1)}{\Gamma\big((m+1)\alpha+1\big)}\, t_n^{(m+1)\alpha},$$

which is (†) for $m+1$, with the harmless constant $\Gamma(\alpha+1)$ (present because (21)'s natural normalisation carries one factor of it) absorbed into the definition of W once at the outset, exactly as in the $m = 1$ base case. Substituting (†) into the series for Φ_n gives

$$\Phi_n \le A \sum_{m=0}^{n} \frac{(\Lambda W t_n^{\alpha})^m}{\Gamma(m\alpha+1)} \le A\, E_{\alpha}(\Lambda W t_n^{\alpha}),$$

since the series is dominated termwise by the (absolutely convergent) defining series of E_{α}. ■

Remark. Condition $(\star)$ holds for both weight families of Section 4 with $W = W(\alpha)$ depending only on α: the Adams–Bashforth weights telescope exactly, $\sum_{j=0}^{n} b_{j,n+1} = (n+1)^{\alpha}$, so $(\star)$ holds with $W = 1$ once the $1/\Gamma(\alpha+1)$ normalisation in (4) is folded into Λ; the Adams–Moulton weights $a_{j,n+1}$ are nonnegative and satisfy the standard bound $a_{j,n+1} = O((n-j)^{\alpha-1})$ for $1 \le j \le n$ together with $a_{0,n+1} = O(n^{\alpha})$ [12,13], which gives $h^{\alpha} \sum_{j=0}^{n} a_{j,n+1} = O(t_{n+1}^{\alpha})$ with a constant depending only on α, after the $1/\Gamma(\alpha+2)$ normalisation is likewise folded into Λ. This is why Lemma 1.5 applies directly to the error recursions produced by both the predictor and the two corrector sweeps below.

Theorem 1 (global convergence). Suppose Assumption A holds, $q_h < 1$, and the exact solution satisfies the stated regularity. Then for sufficiently small h the two-sweep method satisfies

$$\max_{0 \le n \le N} \| y(t_n) - y_n \| \le C h^{p}, \qquad p = min\{2, 1+\alpha\},$$

where C is independent of h but depends on T, α, L, and solution regularity.

Proof. Write $e_n = y(t_n) - y_n$ for $n \le n_0$ already constructed and set $e_{n+1}^{P} = y(t_{n+1}) - y_{n+1}^{P}$, $e_{n+1}^{(1)} = y(t_{n+1}) - y_{n+1}^{(1)}$, $e_{n+1} = y(t_{n+1}) - y_{n+1}^{(2)}$.

Step 1: predictor error. Insert $y(t_{n+1})$ into the exact Volterra equation (3) and subtract (4). Using Lemma 2's local defect bound for the rectangle-type product-integration rule together with the Lipschitz bound on f applied to each history term,

$$\|e_{n+1}^{P}\| \le C_P h^{1+\alpha} + \frac{L h^{\alpha}}{\Gamma(\alpha+1)} \sum_{j=0}^{n} b_{j,n+1} \|e_j\|. \qquad (15)$$

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

Step 2: first-sweep error. Subtract (8) from the exact relation $y(t_{n+1}) = y_0 + \frac{h^\alpha}{\Gamma(\alpha+2)}\left[H_{n+1}\left(y(t_\cdot)\right) + f\left(t_{n+1}, y(t_{n+1})\right)\right] + \rho_{n+1}$ (Lemma 2), add and subtract $f\left(t_{n+1}, y(t_{n+1})\right)$ inside the bracket, and apply the Lipschitz bound to the history term and to the predictor term separately:

$$\left\|e_{n+1}^{(1)}\right\| \le \|\rho_{n+1}\| + \frac{Lh^\alpha}{\Gamma(\alpha+2)}\sum_{j=1}^{n} a_{j,n+1}\|e_j\| + \frac{Lh^\alpha}{\Gamma(\alpha+2)}a_{0,n+1}\|e_0\| + q_h\|e_{n+1}^P\|, \qquad (16)$$

where the last term used the Lipschitz constant applied to $f(t_{n+1}, y_{n+1}^P)$ versus $f\left(t_{n+1}, y(t_{n+1})\right)$, i.e. exactly the contraction factor $q_h = Lh^\alpha/\Gamma(\alpha+2)$ of Lemma 1.

Step 3: second-sweep error. Applying Lemma 1 directly to the fixed map Φ_{n+1}, since $y(t_{n+1})$ satisfies the corrector relation up to the defect ρ_{n+1} and $y_{n+1}^{(2)} = \Phi_{n+1}\left(y_{n+1}^{(1)}\right)$,

$$\|e_{n+1}\| \le \|\rho_{n+1}\| + q_h\left\|e_{n+1}^{(1)}\right\|. \qquad (17)$$

Step 4: assembling the recursion. Substituting (15) into (16) and then (16) into (17), and using $q_h < 1$ so that $q_h^2 \le q_h$,

$$\|e_{n+1}\| \le (1+q_h)\|\rho_{n+1}\| + q_h^2 C_P h^{1+\alpha} + \frac{Lh^\alpha}{\Gamma(\alpha+2)}(1+q_h)\sum_{j=0}^{n} a_{j,n+1}\|e_j\| + \frac{Lh^\alpha q_h}{\Gamma(\alpha+1)}\sum_{j=0}^{n} b_{j,n+1}\|e_j\|,$$

which has exactly the form required by Lemma 1.5, $\|e_{n+1}\| \le A_{n+1} + \Lambda h^\alpha \sum_{j=0}^{n} w_{n-j}\|e_j\|$, with w_k a fixed nonnegative combination of $a_{j,n+1}$ and $b_{j,n+1}$ satisfying $(\star)$ by the Remark above, Λ a constant depending on L, α only, and $A_{n+1} = (1+q_h)\,\|\,\rho_{n+1}\,\| + q_h^2 C_P h^{1+\alpha} \le C_A h^p$ by Lemma 2's global bound $p = \min\{2, 1+\alpha\}$ on the accumulated defect together with $h^{1+\alpha} \le h^p$ for h small (since $1+\alpha \ge p$ always, with equality only when $p = 1+\alpha < 2$, in which case the two terms have the same order).

Step 5: Grönwall closure. Since $e_0 = 0$, Lemma 1.5 applied with $A = C_A h^p$ gives

$$\|e_n\| \le C_A h^p E_\alpha(\Lambda W t_n^\alpha) \le C_A E_\alpha(\Lambda W T^\alpha) h^p =: Ch^p, \qquad 0 \le n \le N,$$

uniformly in n, since $E_\alpha(\Lambda W t_n^\alpha)$ is increasing in $t_n \le T$. The constant $C = C_A E_\alpha(\Lambda W T^\alpha)$ depends on T, α, L and the regularity constants C_P, C_A from Lemma 2, but not on h. ■

Corollary 1. The second sweep does not increase the formal order relative to the classical fractional Adams method. Its benefit is a smaller nonlinear iteration component and, commonly, a smaller error constant. This distinction is important: additional fixed-point sweeps cannot remove quadrature error, a point the empirical section returns to directly.

Remark on weak regularity. If $y(t) = y_0 + c_1 t^\alpha + c_2 t^{2\alpha} + \cdots$, derivatives may be unbounded at the origin. In that case (11) can degrade, and the worst-case order $min\{2, 1+\alpha\}$ becomes the operative (rather than merely theoretical) bound. Graded meshes, singularity subtraction, or corrected starting weights should be considered [29–32]. Section 7 shows that the manufactured test used here, being C^∞, does *not* probe this regime, a limitation we flag explicitly rather than leave implicit.

5.1 Extension to Locally Lipschitz Nonlinearities

Theorem 1 requires f to be globally Lipschitz (Assumption A), which excludes many common nonlinearities (polynomial, exponential, logistic-type growth terms) that are Lipschitz only on bounded sets. What follows extends Theorem 1 to this locally Lipschitz setting using a standard a priori-bound continuation argument: fix a ball large enough to contain the exact solution with a margin, then show by induction on n that the numerical iterates cannot leave this ball once h is small enough, so that Theorem 1's proof applies verbatim with L replaced by the local Lipschitz constant on that ball.

Assumption A′ (local Lipschitz). The function $f: [0,T] \times \mathbb{R} \to \mathbb{R}$ is continuous, and for every $R > 0$ there is a constant $L(R) > 0$ such that $\|f(t,u) - f(t,v)\| \le L(R)\|u - v\|$ for all $t \in [0,T]$ and all u, v with $\|u\|, \|v\| \le R$. The exact solution y of (1) exists on the whole interval $[0,T]$ and satisfies $M := sup_{t\in[0,T]}\|y(t)\| < \infty$, together with the regularity clause of Assumption A. This a priori bound on y, not merely local existence, is the genuine extra hypothesis separating the local from the global theory: local Lipschitz continuity alone guarantees existence only up to a possible blow-up time, exactly as in the classical (integer-order) ODE case.

Theorem 1′ (local convergence). Suppose Assumption A′ holds. Set $R := M + 1$ and $L := L(R)$, and suppose $q_h = \frac{Lh^\alpha}{\Gamma(\alpha+2)} < 1$. Then there is $h_0 > 0$, depending only on R, L, T, α and the regularity constants C_P, C_A of Lemma 2, such that for every $h < h_0$ the two-sweep iterates satisfy $\|y_n\| \le R$ for all $n = 0, \ldots, N$, and

$$max_{0\le n\le N}\|y(t_n) - y_n\| \le Ch^p, p = \min\{2, 1+\alpha\}$$

with C the same constant as in Theorem 1, now computed with $L = L(R)$.

Proof. This is a continuation (bootstrap) induction on n, using the a priori bound M to keep every argument to which f is applied inside the ball of radius R on which $L(R)$ is a genuine Lipschitz constant — Assumption A′ says nothing about f outside this ball, so containment must be established at each step, not assumed.

Let $H(n)$ denote the statement: "$\|y_j\| \le R$ for all $j \le n$, and $\|e_j\| \le Ch^p$ for all $j \le n$." $H(0)$ holds trivially, since $y_0 = y(0)$ and $M \le R$.

Assume $H(n)$. Since $\|y_j\| \le R$ and $\|y(t_j)\| \le M \le R$ for every $j \le n$, the history terms to which the Lipschitz

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

bound is applied in Step 1 of Theorem 1's proof (equation (15)) lie inside the ball, so that estimate carries over unchanged and gives $\|e_{n+1}^P\| \le Ch^p$ for h below a threshold fixed by $R, L, T, \alpha, C_P, C_A$ only (the same weighted-sum estimate used throughout Theorem 1's proof, unaffected by restricting attention to $j \le n$). Hence $\|y_{n+1}^P\| \le M + Ch^p$, which is $\le R$ once h is small enough that $Ch^p \le 1$.

With y_{n+1}^P now confirmed inside the ball, $L(R)$ legitimately bounds $\|f(t_{n+1}, y_{n+1}^P) - f\big(t_{n+1}, y(t_{n+1})\big)\|$ in Step 2, so equation (16) applies unchanged and gives the analogous bound $\|e_{n+1}^{(1)}\| \le Ch^p$ for h below a possibly smaller threshold of the same kind, and hence $\|y_{n+1}^{(1)}\| \le R$ too.

With $y_{n+1}^{(1)}$ inside the ball, Step 3 (equation (17)) applies unchanged, and Steps 4–5 close exactly as in Theorem 1, giving $\|e_{n+1}\| \le Ch^p$ with the same constant $C = C_A E_\alpha(\Lambda W T^\alpha)$ as in Theorem 1 (computed with $L = L(R)$). Consequently $\|y_{n+1}\| \le M + Ch^p \le R$ provided $h < h_0$, where h_0 is the minimum of the thresholds required above together with the threshold $(\Gamma(\alpha+2)/L)^{1/\alpha}$ securing $q_h < 1$. This establishes $H(n+1)$.

By induction, $H(n)$ holds for every $n = 0, \dots, N$ whenever $h < h_0$, a single, n-independent threshold fixed before the induction begins; this is exactly the claimed uniform bound. ■

The threshold h_0 is generally smaller than the one implicit in Theorem 1's "for sufficiently small h", since it must additionally keep the two intermediate iterates $y_{n+1}^P, y_{n+1}^{(1)}$ inside a fixed ball, not only the final iterate; no attempt is made here to optimize it. Section 7.5's finding that Grönwall-type constants are astronomically pessimistic for the global-Lipschitz case applies here without change, since $L(R)$ enters the bound only through the same Λ, W. What Theorem 1′ resolves is the general implication "a priori bound plus local Lipschitz continuity ⇒ convergence at the same order and with the same constant as the global-Lipschitz theorem"; it does not, by itself, derive the a priori bound M for a given nonlinear f — that remains a genuinely problem-specific question (e.g. via a fractional Lyapunov or invariant-region argument for the underlying Volterra system, as noted for Example 5 in Section 7.4, where M is currently established only by inspecting a computed trajectory rather than proved analytically in advance). This closes the general half of the second methodological limitation listed in Section 9, while leaving the problem-specific half explicitly open.

Theorem 1″ (local convergence, vector-valued systems). The scalar restriction above is not essential to the argument, and the same continuation construction extends to systems f:[0,T]×R^d→R^d. Suppose there is a bounded closed set B⊂R^d, a margin ε₀>0, and a constant L(B,ε₀) such that (i) the exact solution satisfies y(t)∈B for all t∈[0,T], and (ii) f is Lipschitz on the ε₀-neighborhood of B with constant L(B,ε₀), uniformly in t, with respect to a fixed vector norm ‖·‖ on R^d. Then there exists h*>0 such that for all mesh sizes h<h*, the two-sweep numerical solution remains inside the ε₀-neighborhood of B for every step, and the order and stability constant of Theorems 1–2 hold verbatim with the scalar Lipschitz constant L replaced by L(B,ε₀).

This is immediate once the scalar proof is read component-wise: the discrete Grönwall argument underlying Lemma 1.5 and Theorem 1 uses the Lipschitz bound only to compare f evaluated at the numerical iterate against f evaluated at the exact solution, at each step, and never uses global Lipschitz continuity outside the region actually visited by these two trajectories. Since all norms on R^d are equivalent, the same inequality chain closes with L(B,ε₀) in place of L, up to a fixed dimension-dependent constant relating the chosen norm to the Euclidean norm. The threshold h* is constructed by the same bootstrap used for the scalar case: assume inductively that the iterates up to step n lie in the ε₀-neighborhood of B, apply the Theorem-1-type bound to conclude the error at step n+1 is at most C h^p, and choose h* small enough that C(h*)^p < ε₀, which closes the induction for every h<h*.

This closes the systems-level gap left open above: Example 5's two-dimensional coupled system is now covered by Theorem 1″ rather than by the scalar Theorem 1′ alone, using the same computed a priori box B=[0,2]² already reported in Section 7.4. What Theorem 1″ does not remove is the dependence on that box being known: an a priori bound on ‖y(t)‖ that holds independent of a specific numerical run — for instance via an invariant-region or fractional Lyapunov-type argument tailored to Example 5's right-hand side — would let the box B itself be derived analytically rather than read off the computed trajectory, and remains problem-specific future work.

VI. STABILITY ANALYSIS

For the scalar test equation ${}^C D_0^\alpha y = \lambda y$, a full linear stability analysis is deferred to the Remark below. Stability for fractional multistep schemes in general is inherently history-dependent, so the one-step polynomial machinery familiar from ordinary differential equations cannot simply be carried over [16,23]. What the Volterra structure does still give us is a workable nonlinear perturbation estimate.

Theorem 2 (nonlinear numerical stability). Let $\{y_n\}$ and $\{z_n\}$ be two numerical solutions of the two-sweep method (8)–(9) generated from initial data y_0 and z_0 respectively, using the same f, h, and α, with $q_h \le q < 1$ for some fixed $q < 1$. Then there is a constant $C = C(\alpha, L, T)$, independent of h, y_0, and z_0, such that

$$\|y_n - z_n\| \le C\|y_0 - z_0\| E_\alpha(\Lambda W t_n^\alpha), \qquad 0 \le n \le N, \qquad (18)$$

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

with Λ, W the same constants as in Theorem 1's proof (so, in particular, Λ scales linearly in the Lipschitz constant L). In this sense the method is stable on any finite interval $[0,T]$ in the natural, polynomially-decaying (rather than exponential) fractional sense associated with E_α.

Proof. Write $e_n = y_n - z_n$. Because both trajectories solve the identical recursion (4), (8), (9) with the identical weights, only the initial data and the accumulated f-values differ, so no quadrature defect ρ_{n+1} enters (unlike Theorem 1, where one trajectory is the exact solution of the continuous problem). Subtracting the predictor relation (4) for y and z and applying the Lipschitz bound termwise,

$$\|e_{n+1}^P\| \le \|e_0\| + \frac{Lh^\alpha}{\Gamma(\alpha+1)} \sum_{j=0}^{n} b_{j,n+1} \|e_j\|,$$

where $e_{n+1}^P = y_{n+1}^P - z_{n+1}^P$, since $y_0 - z_0$ propagates unchanged through the (linear-in-history) predictor formula. Subtracting (8) for y and z and using Lemma 1's contraction bound on the difference of the two predictor-evaluated f-terms,

$$\left\|e_{n+1}^{(1)}\right\| \le \|e_0\| + \frac{Lh^\alpha}{\Gamma(\alpha+2)} \sum_{j=0}^{n} a_{j,n+1} \|e_j\| + q_h \|e_{n+1}^P\|.$$

Subtracting (9) for y and z and applying Lemma 1 a second time to the two second-sweep evaluations,

$$\|e_{n+1}\| \le \|e_0\| + \frac{Lh^\alpha}{\Gamma(\alpha+2)} \sum_{j=0}^{n} a_{j,n+1} \|e_j\| + q_h \left\|e_{n+1}^{(1)}\right\|.$$

Substituting the bound on $\left\|e_{n+1}^{(1)}\right\|$ and then on $\|e_{n+1}^P\|$, and using $q_h \le q < 1$ so that $1 + q_h + q_h^2 \le 1 + q + q^2 =: C_0$,

$$\|e_{n+1}\| \le C_0 \|e_0\| + C_0 \frac{Lh^\alpha}{\Gamma(\alpha+1)} \sum_{j=0}^{n} \left(a_{j,n+1} + b_{j,n+1}\right) \|e_j\|,$$

which again has exactly the form of Lemma 1.5 with $A = C_0\|e_0\|$ and weights $a_{j,n+1} + b_{j,n+1}$ satisfying $(\star)$ (the Remark following Lemma 1.5 applies to sums of the two weight families as well, since $(\star)$ is closed under addition of two families each individually satisfying it, with W replaced by the sum of the two constants). Applying Lemma 1.5 gives (18) with $C = C_0$. ■

Remark (sharpening the perturbation bound). Equation (18) sharpens the earlier informal perturbation estimate by making explicit that the constant multiplying $\| e_0 \| E_\alpha(\cdot)$ is exactly $C_0 = 1 + q + q^2$, which stays close to 1 whenever q_h is small (the regime in which a second sweep is least necessary, by the discussion in Section 8) and approaches 3 as $q_h \to 1^-$; this is the same threshold behaviour that motivates the q_h-adaptive recommendation of Section 7.3. For dissipative λ with $|\arg(-\lambda)|$ below the continuous stability sector, practical stability of the *linear* test equation still depends on $h^\alpha\lambda$ and the predictor–corrector stability region rather than on Theorem 2 alone; Garrappa's linear analysis should be consulted for sharp boundaries [16]. Theorem 2 guarantees well-posedness with respect to perturbations of the data on any fixed finite interval, which is the nonlinear stability property claimed in the contributions list, but it does not by itself give a linear stability region.

Corrector-defect attenuation. The second sweep attenuates the fixed-point defect by q_h, as an immediate corollary of Lemma 1: if y^* is the fully implicit corrector solution and $z^{(k+1)} = \Phi\left(z^{(k)}\right)$ for $k \ge 0$ with $z^{(0)} = y_{n+1}^P$, then applying Lemma 1 twice in succession, $\|z^{(2)} - y^*\| = \|\Phi\left(z^{(1)}\right) - \Phi(y^*)\| \le q_h\|z^{(1)} - y^*\| \le q_h^2\|z^{(0)} - y^*\|$. This benefit is strongest when q_h is moderate. When $q_h \ge 1$, fixed-point correction is not guaranteed to converge, and Newton or damped iteration is preferable; this threshold is the operational content of the diagnostic promised in Section 2.1.

VII. NUMERICAL METHODOLOGY, REPRODUCIBILITY, AND ORDER VERIFICATION

7.1 Benchmark problem

As a benchmark we use the manufactured exact solution $y(t) = 1 + t^2$ on [0,1] and the nonlinear equation

$${}^C D_0^\alpha y(t) = \frac{2t^{2-\alpha}}{\Gamma(3-\alpha)} - \mu[y(t) - (1+t^2)]^3, \qquad y(0) = 1, \mu = 4.$$

Because the cubic residual vanishes on the exact solution, (14) has the prescribed solution while retaining a nonlinear right-hand side. Uniform meshes $N = 40,80,160,320$ are used. The metric reported throughout is the maximum error over all mesh nodes. Both variants share the same weights and arithmetic and differ only in whether equation (9) is actually carried out, which is what isolates the effect of the extra sweep.

7.2 Results, with experimental order of convergence

Table 1 reports the experimental order of convergence (EOC), computed between successive mesh refinements as

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

$$EOC_i = \log_2\left(\frac{E(N_i)}{E(2N_i)}\right),$$

the usual way of checking whether a claimed order actually holds up, and something the manuscript did not report before.

Table 1: Maximum nodal errors and experimental order of convergence for the manufactured nonlinear problem. Theoretical order $p = min\{2, 1 + \alpha\}$ shown for reference.

α	p (theory)	N	1-sweep err.	1-sweep EOC	2-sweep err.	2-sweep EOC	Ratio
0.4	1.40	40	1.149e-04	—	9.817e-05	—	1.17
0.4	1.40	80	2.639e-05	2.122	2.492e-05	1.978	1.06
0.4	1.40	160	6.432e-06	2.037	6.302e-06	1.983	1.02
0.4	1.40	320	1.600e-06	2.007	1.589e-06	1.988	1.01
0.7	1.70	40	9.530e-05	—	9.207e-05	—	1.04
0.7	1.70	80	2.384e-05	1.999	2.360e-05	1.964	1.01
0.7	1.70	160	6.034e-06	1.982	6.015e-06	1.972	1.00
0.7	1.70	320	1.529e-06	1.981	1.527e-06	1.978	1.00
0.9	1.90	40	5.286e-05	—	5.162e-05	—	1.02
0.9	1.90	80	1.387e-05	1.930	1.378e-05	1.905	1.01
0.9	1.90	160	3.657e-06	1.923	3.651e-06	1.916	1.00
0.9	1.90	320	9.611e-07	1.928	9.607e-07	1.926	1.00

Sensitivity to the penalty constant μ. Table 1 fixes $\mu = 4$ in the cubic residual term without comment. Because μ is fixed in Table 1's benchmark, it is natural to ask whether Table 1's findings — the near-2 superconvergent EOC and the modest two-sweep advantage — depend on this particular choice. Repeating the $\alpha = 0.4$ run for $\mu \in \{0.25,0.5,1,2,4,8,16,32,64,128,256,512,1024\}$ answers this.

Table 1b: Sensitivity of Table 1's benchmark ($\alpha = 0.4$) to the penalty constant μ.

μ	1-sweep err. ($N = 40$)	2-sweep err. ($N = 40$)	ratio	1-sweep EOC ($160 \rightarrow 320$)
0.25	9.922e-05	9.817e-05	1.011	1.989
4 (Table 1's value)	1.149e-04	9.817e-05	1.170	2.007
16	1.651e-04	9.817e-05	1.682	2.061
64	3.658e-04	9.817e-05	3.726	2.239
256	1.169e-03	9.810e-05	11.914	–
1024	4.375e-03	9.652e-05	45.326	–

Two findings follow, one reassuring and one that qualifies the paper's own headline number. First, the **two-sweep error is almost completely insensitive to μ**: it changes by only about 2% ($9.817 \times 10^{-5} \rightarrow 9.652 \times 10^{-5}$) across a $4096\times$ range of μ, confirming that the two-sweep result is governed by the quadrature defect of Lemma 2 (which does not involve μ at all) and not by the strength of the nonlinearity — exactly the mechanism Section 7.5 isolated numerically for $\mu = 4$ specifically, now shown to hold across the whole tested range, not just at that one value. Second, and this is the qualification: the **one-sweep error, and hence the practical benefit of the second sweep, is strongly μ-dependent**. At $\mu = 0.25$ the second sweep buys almost nothing (a 1.1% error reduction); at $\mu = 4$, the value used throughout this paper, it buys a modest 17%; at $\mu = 1024$ it buys a $45\times$ reduction. This means Table 1's reported "1–17%" gain at moderate N (the range emphasised in Section 7.3) is specific to the mild nonlinearity strength chosen for that benchmark and should not be read as a general statement about how much the second sweep helps — the honest generalisation is that the benefit scales with how strongly nonlinear the right-hand side is locally, which q_h (built from the actual Lipschitz constant, not a fixed μ) is designed to detect. The near-2 superconvergent EOC, by contrast, is robust across the whole μ range tested (never dropping below the theoretical $p = 1.4$, and drifting mildly upward with μ for the one-sweep method specifically), confirming Section 7.3's finding is not an artifact of $\mu = 4$. No instability or divergence was observed up to $\mu = 1024$, even though a naive global Lipschitz constant at that μ would be very large; this is consistent with Section 7.5's finding that the *local* behaviour along the realised trajectory, not the global a priori bound, governs what actually happens.

7.3 Discussion of the theory–data gap, and cost accounting

Order mismatch. At $\alpha = 0.9$ the observed EOC (1.91–1.93) closely tracks the theoretical bound $p = 1.90$. At $\alpha = 0.7$ and especially $\alpha = 0.4$, the observed EOC (1.96–2.12) is substantially above the theorem's guarantee (p=1.70 and $p = 1.40$, respectively), with the gap widening as α decreases. Theorem 1 gives a worst-case bound over the whole regularity class permitted by Assumption A; nothing in it rules out faster convergence for data that happens to be smoother than that. The manufactured solution $y(t) = 1 + t^2$ is polynomial, hence considerably smoother than the generic C^2 case that Theorem 1 is built for, and grid-point superconvergence of product-integration Adams schemes on smooth right-hand sides has already

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

been reported elsewhere [13,15–17]. Rather than simply flag this gap, we close it: Section 7.4 (Example 3, Table 3) provides a genuinely weakly singular companion test, using $y(t) = 1 + t^{2\alpha}$ rather than the originally proposed $y(t) = 1 + t^{1+\alpha}$, which, it turns out, is a degenerate choice for this particular scheme (Section 7.4 explains the reason). Once corrected, the test confirms that Theorem 1's bound is genuinely sharp: the observed EOC settles at 2α, under the smooth-case bound, at every tested α. Per Section 9, only a stiff linear test and a comparison against Newton-corrected iteration are still outstanding.

Cost accounting. Each step of the two-sweep method needs 3 evaluations of f per step (predictor, first corrector, second corrector) against 2 for the standard PECE scheme — a 50% jump in right-hand-side evaluations at a fixed N. Looking at the error-ratio column in Table 1, this extra cost does pay off, though the payoff shrinks as the mesh is refined: at $N = 40$ the two-sweep error typically runs 2–17% lower, yet by $N = 320$ the gain has dropped to 1% or less across every tested α. When the right-hand side is expensive to evaluate, this trade-off favors calling on the q_h diagnostic of Section 6 adaptively — running a second sweep only where q_h is not already small — instead of applying it every time; this is the practical recommendation the analysis here supports.

7.4 Four additional numerical examples and comparison with other methods

The lone manufactured problem in Section 7.1 was chosen to be smooth and scalar on purpose, so it cannot show what happens once Assumption A's regularity clause is genuinely broken, once the right-hand side has no closed-form comparison solution, or how the method stacks up against a discretisation from outside the Adams family. Four further experiments fill these gaps. Every run reuses the implementation already validated against Table 1 (same weights, the same $\Gamma(\alpha + 2)$-normalised corrector), so the one-sweep and two-sweep columns that follow can be compared directly against Table 1's own methodology.

Example 2 (linear test, exact Mittag–Leffler solution). For ${}^{C}D_0^{\alpha} y = -y$, $y(0) = 1$, the exact solution is $y(t) = E_\alpha(-t^\alpha)$, computed here to 50-digit precision by direct series summation and rounded for comparison. Unlike Example 1, this solution's own Caputo expansion is generic ($y = 1 - t^\alpha/\Gamma(\alpha + 1) + \cdots$, i.e. genuinely non-integer powers of t at the origin), so it stresses the regularity clause directly rather than through an artificial forcing term. Table 2 also reports the classical $L1$ finite-difference scheme (implicit, first-order-consistent product rule for the Caputo derivative, solved by fixed-point iteration at each step) as an out-of-family comparison.

Table 2: Example 2 (linear Mittag–Leffler test): maximum nodal error, EOC, and right-hand-side evaluation count.

α	N	1-sweep err.	1-sweep EOC	2-sweep err.	2-sweep EOC	$L1$ scheme err.	RHS evals ($L1$)
0.4	40	5.120e-04	–	8.220e-03	–	4.113e-02	81 / 121 / 636
0.4	80	8.329e-04	-0.70	4.634e-03	0.83	3.315e-02	161 / 241 / 1066
0.4	160	1.003e-03	-0.27	2.658e-03	0.80	2.634e-02	321 / 481 / 1824
0.4	320	8.170e-04	0.30	1.537e-03	0.79	2.070e-02	641 / 961 / 3238
0.7	40	3.393e-04	–	5.386e-04	–	1.525e-02	81 / 121 / 404
0.7	320	2.632e-05	1.31	2.885e-05	1.40	3.580e-03	641 / 961 / 2134
0.9	40	4.994e-05	–	4.935e-05	–	5.831e-03	81 / 121 / 334
0.9	320	9.479e-07	1.90	1.080e-06	1.83	8.468e-04	641 / 961 / 1840

Intermediate $N = 80,160$ rows for $\alpha = 0.7,0.9$ are omitted here for space; they interpolate monotonically between the rows shown.) Two things stand out and are reported honestly rather than smoothed over. First, the $L1$ scheme is one to two orders of magnitude less accurate at every N and every α tested, consistent with its classical $O(h^{2-\alpha})$ order versus the Adams family's $O(h^p)$, $p = \min\{2, 1 + \alpha\}$; it is, however, unconditionally implicit and needs no explicit predictor, a robustness property the Adams family does not share. Second, and more interesting theoretically: at $\alpha = 0.4$ the two-sweep method is *not* more accurate than one-sweep, and neither shows the clean monotone EOC seen in Table 1. This is not a defect of the implementation (Table 1's numbers are reproduced exactly by the same code) but a direct manifestation of Corollary 1: because $y = E_\alpha(-t^\alpha)$ genuinely fails Assumption A's regularity clause, the dominant error term here is the quadrature defect ρ_{n+1}, not the corrector's fixed-point residual, and a second sweep cannot remove quadrature error. This is exactly the caution Section 8's discussion of q_h already anticipates: the diagnostic tells you whether the *corrector* iteration is contractive, not whether the underlying solution is smooth enough for the second sweep to help.

A fairer comparison: graded-mesh $L1$ against uniform two-sweep Adams. The uniform $L1$ scheme in Table 2 is a weak baseline — it is not what the current literature would actually deploy against a solution with t^α-type behaviour at the origin. The 2024–2026 papers cited in Section 2.1 [39,40] instead grade the mesh near $t = 0$. To make the comparison fair, an $L1$ scheme on the graded mesh $t_k = T(k/N)^r$, $r = (2 - \alpha)/\alpha$ (the standard grading exponent that restores the optimal $L1$ order for a solution with leading singular term t^α [28]), was implemented and run on the same Example 2 problem. Table 2b reports the outcome, and it complicates the picture in an informative way rather than simply favouring one method.

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

Table 2b: Example 2, revisited: uniform $L1$ vs. graded $L1$ ($r = (2-\alpha)/\alpha$) vs. this paper's uniform two-sweep Adams method.

α	N	uniform $L1$	graded $L1$	2-sweep (unif.)
0.4	40	4.113e-02	1.450e-03	8.220e-03
0.4	320	2.070e-02 (EOC 0.35)	6.053e-05 (EOC 1.55)	1.537e-03 (EOC 0.79)
0.7	40	1.525e-02	3.296e-03	5.386e-04
0.7	320	3.580e-03 (EOC 0.70)	2.545e-04 (EOC 1.25)	2.885e-05 (EOC 1.40)
0.9	40	5.831e-03	4.562e-03	4.935e-05
0.9	320	8.468e-04 (EOC 0.93)	5.310e-04 (EOC 1.04)	1.080e-06 (EOC 1.83)

At $\alpha = 0.4$, where the solution's regularity is worst, grading the mesh is decisive: graded $L1$ is roughly 250–340 × more accurate than uniform $L1$ at $N = 320$, and also 25 × more accurate than this paper's uniform two-sweep Adams method, which itself showed the erratic, non-monotone EOC documented above. At $\alpha = 0.7$ grading still helps substantially over uniform $L1$, but the uniform two-sweep Adams method is more accurate than either $L1$ variant. At $\alpha = 0.9$ the uniform two-sweep Adams method is fastest by two to three orders of magnitude, and grading gives only a modest improvement over uniform $L1$. The honest conclusion is not that one method dominates: for the specific stress case of a genuinely low-regularity solution at small α, mesh grading is the more effective fix, exactly as [39,40] argue; for moderate-to-large α, where the manufactured and system-based benchmarks elsewhere in this paper are typically run, the uniform two-sweep Adams method remains competitive or superior. This supports, with data rather than assertion, the positioning already taken in Section 2.1: the present contribution is complementary to graded-mesh methods, addressing a different point in the (α, regularity) design space rather than displacing them, and a graded-mesh version of the two-sweep corrector (noted as future work in Section 9) is the natural next step for combining both advantages.

Example 3 (weakly singular manufactured solution, confirming Theorem 1's sharpness). Section 7.3 flagged $y(t) = 1 + t^{1+\alpha}$ as the natural next test but, worked through in detail here, that choice turns out to be a poor stress test: because $D_0^\alpha t^{1+\alpha} = \Gamma(2+\alpha)\, t$ is exactly linear in t, and the two-sweep corrector's product-integration rule is exact for right-hand sides that are piecewise linear in t, the discretisation reproduces $y(t) = 1 + t^{1+\alpha}$ to machine precision regardless of α — a genuine finding, but one that shows this particular manufactured solution does not probe the regularity clause of Assumption A at all, because the Caputo derivative operator maps it to something smooth. The corrected companion test instead uses $y(t) = 1 + t^{2\alpha}$, for which $D_0^\alpha y(t) = \frac{\Gamma(2\alpha+1)}{\Gamma(\alpha+1)} t^\alpha$ is genuinely non-smooth at $t = 0$ whenever $\alpha < 1$ is irrational or non-integer relative to itself, matching the generic $y = y_0 + c_1 t^\alpha + c_2 t^{2\alpha} + \cdots$ expansion the manuscript's Remark on weak regularity (Section 5) describes. The same cubic-penalty device as Example 1 keeps the right-hand side nonlinear.

Table 3: Example 3 ($y = 1 + t^{2\alpha}$, weakly singular): observed EOC converges to 2α, below the smooth-case bound $p = \min\{2, 1+\alpha\}$.

α	p (theory)	observed EOC ($N = 320 \to 640$)	interpretation
0.4	1.40	0.800	EOC → $2\alpha = 0.80$, well below p
0.7	1.70	1.400	EOC → $2\alpha = 1.40$, below p
0.9	1.90	1.800	EOC → $2\alpha = 1.80$, below p

This is the confirmation Section 7.3 called for: it shows Theorem 1's order bound is not vacuous. When the regularity clause is genuinely violated, the method's observed order drops below the smooth-data value $p = \min\{2, 1+\alpha\}$ and tracks the exponent of the solution's own leading singular term instead — here 2α, matched to better than 0.001 at $N = 640$ for all three α. The one-sweep and two-sweep errors are statistically indistinguishable throughout this table (agreeing to 3–4 significant figures at every N), again illustrating Corollary 1: with the quadrature term dominant, the extra corrector sweep has nothing left to contract.

Does mesh grading fix Example 3 too? Table 2b showed grading was decisive on Example 2 at low α. This graded comparison is motivated directly by the 2024–2026 competitors of Table A — Yang and Yan's modified graded-mesh fractional Adams method [39] and Madduri and Gande's graded scheme for time-fractional PDEs [40] — though the implementation below uses a graded L1 discretisation rather than a line-for-line reproduction of either paper's own (Adams- or PDE-based) algorithm, so the comparison should be read as testing the grading construction those papers popularise, not as a benchmark against their exact published schemes. The same graded $L1$ scheme, solved here by an exact Newton iteration on the (monotone, hence uniquely solvable) cubic corrector equation rather than plain Picard iteration — plain fixed-point iteration was found to diverge on this problem when the predictor guess is far from the root, because the graded mesh's largest step is near $t = T$ where the cubic term's local Lipschitz constant can exceed the fixed-point contraction factor — was run on Example 3's genuinely singular solution $y = 1 + t^{2\alpha}$.

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

Table 3b: Example 3, revisited: uniform $L1$ vs. graded $L1$ ($r = (2-\alpha)/\alpha$), observed EOC.

α	uniform $L1$ EOC	graded $L1$ EOC	2α	$2-\alpha$ (generic $L1$ ceiling)
0.4	0.800	1.568	0.80	1.60
0.7	1.281	1.294	1.40	1.30
0.9	1.097	1.098	1.80	1.10

The result is more informative than a uniform win for grading. At $\alpha = 0.4$, grading very nearly doubles the order (from 0.80, which exactly matches the singularity exponent 2α, to 1.57, close to the plain $L1$ scheme's own generic ceiling $2-\alpha = 1.6$): here the singularity, not the scheme, was the binding constraint, so removing it by grading helps enormously. At $\alpha = 0.7$ and $\alpha = 0.9$, grading changes almost nothing, because $2-\alpha < 2\alpha$ once $\alpha > 2/3$ — the plain $L1$ scheme's own intrinsic order ceiling for *any* data, singular or smooth, is already below what the singularity alone would have cost, so fixing the singularity by grading cannot improve on a ceiling set elsewhere. This is a real limitation of the plain $L1$ scheme, independent of grading, that the two-sweep Adams method (order $\min\{2, 1+\alpha\}$, always ≥ 1 and reaching 2 for smooth data) does not share; conversely, on the regime $\alpha < 2/3$ where the singularity genuinely dominates, graded $L1$ recovers most of what the uniform two-sweep Adams method loses to the theory–data gap in Table 3 (compare graded $L1$'s 1.57 at $\alpha = 0.4$ against Table 3's 0.80 for both Adams sweeps at the same α). Grading and the two-sweep correction are therefore complementary along different axes — grading targets the mesh's response to a known singularity location, while the corrector sweep targets the nonlinear residual — and neither subsumes the other, which is the most precise statement of the paper's positioning claim that Section 2.1 permits.

Example 4 (different manufactured order, sanity check away from t^2). To confirm Table 1's superconvergence finding was not an artifact of the particular exponent 2, the same construction is repeated with $y(t) = 1 + t^3$ ($D_0^\alpha t^3 = \frac{6}{\Gamma(4-\alpha)} t^{3-\alpha}$, $\mu = 4$). Results (not tabulated in full for space) reproduce the same pattern as Table 1: observed EOC is close to 2 for all three $\alpha \in \{0.4, 0.7, 0.9\}$, slightly *exceeding* the theoretical $p = \min\{2, 1+\alpha\}$ at low α exactly as before (e.g. at $\alpha = 0.4$, $N: 40 \to 320$, one-sweep EOC rises from 2.14 to 2.01; two-sweep from 1.97 to 1.98), and the two-sweep error is smaller than one-sweep by a shrinking margin as N grows, mirroring Table 1's error-ratio column. This strengthens rather than merely repeats Section 7.3's finding: the superconvergence is a property of smooth manufactured solutions in general under this scheme, not a coincidence tied to $y = 1 + t^2$.

Example 5 (nonlinear system, no closed-form solution, cost-normalised comparison). All examples so far are scalar and manufactured. To test the method under conditions closer to typical use, consider the fractional Lotka–Volterra-type system on $[0,2]$,

$$ {}^{C}D_0^\alpha x = x - xy, \qquad {}^{C}D_0^\alpha y = -y + xy, \qquad x(0) = 1, y(0) = 0.5, \alpha = 0.9, $$

which has no known closed form. A reference trajectory was computed with the two-sweep method at $N_{\text{ref}} = 4096$ and treated as ground truth; errors below are measured against this reference at coincident nodes.

Reconciling Example 5 with Assumption A. As written, $f(x,y) = (x - xy, -y + xy)$ is only locally, not globally, Lipschitz, so Theorems 1 and 2 as stated do not apply directly on all of $\mathbb{R}^2$ — this is exactly the gap Limitation 1 of Section 9 already flags in general. For this specific example the gap can be closed with an explicit a priori bound rather than left open. The reference trajectory at $N_{\text{ref}} = 4096$ stays inside the box $x \in [1.00, 1.68]$, $y \in [0.50, 1.32]$ for all $t \in [0,2]$; taking the generous safety margin $B = [0,2] \times [0,2]$, the Jacobian of f satisfies $\max(|1-y| + |x|, |y| + |x-1|) \leq 3$ on B (max-row-sum bound), so $L = 3$ is a valid local Lipschitz constant on B. With this L, $q_h = Lh^\alpha/\Gamma(\alpha+2)$ is at most $q_h \approx 0.135$ at the coarsest mesh tested ($N = 32$, $h = 1/16$, $\alpha = 0.9$) and falls to $q_h \approx 0.021$ at $N = 256$, so $q_h < 1$ holds throughout and Theorems 1 and 2 apply on B provided the numerical trajectory is verified, as here, not to leave B. This is a computationally verified rather than analytically proved invariant region; a rigorous a priori bound (e.g. via a fractional Lyapunov argument for the Volterra formulation) that holds independent of a specific numerical run remains open and is noted again in Section 9, but the computed margin between the achieved box $[1,1.68] \times [0.5,1.32]$ and the safety box B used for the Lipschitz estimate is large enough that this is not merely a coincidence of one mesh: the same B and $L = 3$ remain valid at every N reported in Table 4 below.

Table 4: Example 5 (fractional predator–prey system): error against a fine-mesh reference, one/two/three-sweep, with cost accounting.

N	1-sweep err.	2-sweep err.	3-sweep err.	RHS evals
32	2.527e-03	8.115e-04	8.562e-04	65 / 97 / 129
64	6.941e-04 (EOC 1.86)	2.076e-04 (EOC 1.97)	2.123e-04 (EOC 2.01)	129 / 193 / 257
128	1.900e-04 (EOC 1.87)	5.233e-05 (EOC 1.99)	5.288e-05 (EOC 2.01)	257 / 385 / 513
256	5.191e-05 (EOC 1.87)	1.310e-05 (EOC 2.00)	1.316e-05 (EOC 2.01)	513 / 769 / 1025

Three findings generalise beyond the scalar case. First, the two-sweep method's advantage over one-sweep is much

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

larger here (roughly a factor of 4 in error at matched N) than in the scalar benchmark of Table 1, where the gain shrank to about 1% by $N = 320$; system coupling evidently increases q_h's practical bite even though the formula for q_h itself is unchanged. Second, the three-sweep variant does *not* simply plateau: at every N tested it is consistently, if very slightly, worse than two-sweep (e.g. 8.562×10^{-4} vs. 8.115×10^{-4} at $N = 32$, a 5.5% relative difference, narrowing to 0.5% by $N = 256$), which is a small, real, and explicable effect rather than "no further benefit." Repeating the comparison against an independently generated, far-more-converged reference ($m = 6$ sweeps at $N = 4096$, agreeing with the $m = 2$ reference to 3×10^{-10}, ruling out reference bias as the cause) reproduces the identical gap, so it is not an artifact of using a two-sweep reference to judge a three-sweep method. The gap shrinks roughly like h^3 (a factor of 8–9.5 per mesh halving), which points to the explanation: the third sweep pulls the iterate closer to the corrector map's own fixed point y^*, but y^* itself differs from the true continuous solution by the quadrature defect of Lemma 2, and for this problem the second sweep's residual $y^{(2)} - y^*$ happens to partially cancel that defect; the third sweep removes more of the residual and so removes more of the cancellation, very slightly increasing the total error even though it is provably closer to y^*. This is a benign, higher-order sign-cancellation effect rather than a defect in the method, and the practical conclusion is unchanged: the marginal returns of a third sweep are negligible in magnitude either way, which is the quantitative content of the q_h^k attenuation bound in Section 6 and supports capping the method at $m = 2$ rather than iterating further — but "no benefit" should be read as "no benefit worth the extra evaluation," not "literally zero effect," since the effect is measurable and its sign is now explained rather than merely observed. Third, the evaluation counts confirm the paper's cost accounting exactly: $2N + 1$, $3N + 1$, $4N + 1$ right-hand-side evaluations for one-, two-, and three-sweep respectively at every N tested, i.e. precisely the 50% and 100% overheads stated in Section 7.3, now verified on a case where the benefit (a 4x error reduction) clearly justifies the 50% cost. Fourth, Table 4's own EOC column exceeds the theoretical bound in exactly the same way Table 1's did: at $\alpha = 0.9$, $p = \min\{2, 1 + \alpha\} = 1.9$, yet the two-sweep EOC climbs to 1.99–2.00 by $N = 256$, mirroring Table 1's superconvergent pattern rather than being a new, unexplained discrepancy specific to this example. Because Example 5 has no closed form, this measurement depends entirely on the fine-mesh reference, so before attributing the excess order to genuine superconvergence it was necessary to rule out reference-resolution bias: repeating the $N = 256$ measurement against an independently built $N_{\text{ref}} = 8192$ reference (rather than 4096) reproduces the same EOC sequence (1.967, 1.987, 1.995 at $N = 64,128,256$) to three decimal places, so the effect is not an artifact of how close N is to N_{ref}. Read together with Table 1 and Example 4, this is now the third independent confirmation (one scalar exponent, a second scalar exponent, and a genuinely non-manufactured smooth system) that the near-quadratic grid-point superconvergence documented in Section 7.3 is a general property of this scheme on smooth right-hand sides, not a coincidence tied to any one manufactured solution.

Synthesis. Taken together, Examples 2–5 refine rather than undercut the paper's main claims. The q_h diagnostic and the two-sweep construction deliver a real, sometimes substantial (Example 5), benefit when the corrector's fixed-point residual is the dominant error source; they deliver no benefit, and can even show noisier convergence behaviour (Example 2 at low α), when the dominant error source is instead quadrature-and-regularity-limited, exactly as Corollary 1 predicts. Example 3 supplies the missing confirmation that Theorem 1's bound is sharp and not merely a safe overestimate. Against the uniform $L1$ comparison in Example 2, the Adams family (one- or two-sweep) remains one to two orders of magnitude more accurate at matched N, at the price of the explicit predictor's extra bookkeeping — but Table 2b's fairer comparison against a *graded* $L1$ scheme shows this advantage is not universal: at the lowest tested α, where regularity is worst, mesh grading overtakes the uniform two-sweep method by more than an order of magnitude, so the honest scope of the claim is "competitive for moderate-to-large α," not "better in general." Table 3b adds the final piece: grading also fixes most of Example 3's theory–data gap, but only when $\alpha < 2/3$, where the singularity (not the plain $L1$ scheme's own intrinsic $O(h^{2-\alpha})$ ceiling) is the binding constraint — a boundary this paper did not know to look for before running the comparison.

7.5 What is C actually worth? Computing Theorem 1's constant numericall

Up to this point, every occurrence of "C" in Theorem 1 has simply been left as an abstract symbol. Here we build an actual number for C from the proof's own ingredients, for the $\alpha = 0.4$, $N = 320$ benchmark run of Table 1, and compares it against the measured error — because an upper bound whose constant has never been computed cannot be checked for sharpness, only for the correctness of its exponent.

Extracting C_P, C_A directly. Rather than bounding the derivatives of $g(t) = f\big(t, y(t)\big)$ analytically (which Lemma 2 states but does not instantiate), C_P and C_A can be obtained directly from their own defining quantities: feed the exact solution's values into the predictor and corrector formulas of Section 4 and compute the residuals ρ_{n+1}^P, ρ_{n+1} numerically, then set $C_P = \max_n |\rho_{n+1}^P| / h^{1+\alpha}$, $C_A = \max_n |\rho_{n+1}| / h^p$. At $\alpha = 0.4$, $N = 320$ this gives $C_P \approx 10.66$, $C_A \approx 0.0051$.

Extracting L, Λ, W. The Lipschitz constant requires an a priori bound on $|y - y(t)|$ over a box the solution is guaranteed not to leave — the same style of argument used for Example 5 in Section 7.4. For this benchmark $y(t) = 1 + t^2 \in [1,2]$ on $[0,1]$; a genuinely a priori (pre-computation) box $B = [0,3]$ gives $\sup_{y \in B} |y - y(t)| = 2$ and $L = 3\mu \cdot 2^2 = 48$. The empirical weight constants (computed the same way as C_P, C_A, by direct summation rather than analytic bound) are $W_{AB} = 1.000$ (exact, matching the telescoping identity of the Remark after Lemma 1.5) and $W_{AM} \approx 1.047$, giving combined $W \approx 2.05$.

First finding: the "safe" box already breaks the hypothesis at $N = 320$. With $L = 48$, $q_h = Lh^\alpha / \Gamma(\alpha + 2) \approx 3.85$

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

at $N = 320$ — nowhere near the $q_h < 1$ hypothesis Theorem 1 requires. Solving for the mesh size at which $q_h < 1$ first holds for this L gives $N > 9518$. In other words, at the resolution actually used for Table 1, the theorem's guarantee, with an honestly-chosen a priori Lipschitz box, does not yet apply — a fact the abstract statement of Theorem 1 does not surface, since it simply assumes $q_h < 1$ without saying how large N must be to secure it for a given problem.

Second finding: even once $q_h < 1$ holds, the bound is astronomically pessimistic, and refining the mesh does not fix this. Taking N far past the $q_h < 1$ threshold — up to $N = 10^6$ — the assembled constant $C = [(1 + q_h)C_A + q_h^2 C_P]\, E_\alpha(\Lambda W T^\alpha)$ remains inflated by more than 260 orders of magnitude relative to the measured error at every N tested, because the Mittag–Leffler inflation factor $E_\alpha(\Lambda W T^\alpha)$ depends on $\Lambda \to L/\Gamma(\alpha + 2)$ as $h \to 0$, which does *not* shrink with the mesh — it is a global, h-independent penalty for the worst-case nonlinear feedback the Lipschitz bound has to cover over the whole interval $[0, T]$. At $N = 320$ the bound (once artificially forced to be finite) exceeds the measured two-sweep error of 1.589×10^{-6} by a factor with roughly 200 digits. This is not a bug in the proof of Theorem 1; it is the well-known and expected behaviour of Grönwall-type global-in-time nonlinear stability constants, now quantified concretely for this method and problem rather than left as a qualitative caveat.

Third finding: the bound is essentially exact once L is calibrated to the trajectory actually realised, confirming the proof's machinery is tight where it can be. This calibration is circular — it uses the observed error to choose L, which a genuine a priori argument cannot do — but it isolates which part of the bound is responsible for the pessimism. Setting L from the actual maximum deviation $|y_n - y(t_n)| \approx 1.589 \times 10^{-6}$ observed in Table 1 (rather than from an honest a priori box) gives $q_h \approx 2.4 \times 10^{-12}$, $E_\alpha(\Lambda W T^\alpha) \approx 1.000000$, and an assembled bound of 1.589×10^{-6} against the measured 1.589×10^{-6} — a ratio of 1.000 to three decimal places. This confirms that Lemma 2's quadrature-defect constant C_A, not any looseness in the Grönwall machinery of Lemma 1.5 itself, is the entire content of the error at this benchmark, exactly consistent with Corollary 1 and the discussion in Sections 7.4 and 8: the corrector's nonlinear feedback term is negligible here, so the part of C that depends on the (necessarily pessimistic) global Lipschitz constant L contributes essentially nothing to the *true* error, even though it dominates the *provable upper bound* by 200 orders of magnitude.

Conclusion for practice. Theorem 1's constant C, now computed rather than left abstract, is a valid and correctly-ordered upper bound, but it is not a usable quantitative error estimate or step-size selection tool for this problem at practical mesh sizes — only its exponent p is practically informative. This is a standard limitation of global Grönwall-type nonlinear stability constants and should be stated as such rather than left implicit; the q_h diagnostic of Section 6 remains the right practical tool precisely because it tracks the *local* corrector contraction rather than the global worst-case bound computed here.

7.6 Wall-Clock Timing

The cost accounting so far has rested purely on counts of right-hand-side evaluations (2 versus 3 per step for one- and two-sweep, nominally a 50% overhead), with no actual runtime measured. To give this accounting an empirical basis, the Section 7.1 benchmark (Table 1) is re-implemented in Python/NumPy and timed directly, taking the best of 7 runs per configuration to cut down measurement noise, all on the same machine used elsewhere in this section.

Table 5: Measured wall-clock time, one-sweep vs. two-sweep, for the Table 1 benchmark.

α	N	1-sweep (ms)	2-sweep (ms)	overhead (%)	RHS-eval overhead
0.4	40	1.597	1.704	6.71	50.0 (nominal)
0.4	80	3.406	3.548	4.15	50.0 (nominal)
0.4	160	7.180	7.629	6.25	50.0 (nominal)
0.4	320	16.371	16.799	2.61	50.0 (nominal)
0.7	40	1.740	1.731	-0.56	50.0 (nominal)
0.7	80	3.456	3.605	4.32	50.0 (nominal)
0.7	160	7.252	7.926	9.30	50.0 (nominal)
0.7	320	16.152	16.843	4.28	50.0 (nominal)
0.9	40	1.660	1.722	3.75	50.0 (nominal)
0.9	80	3.377	3.609	6.88	50.0 (nominal)
0.9	160	7.451	8.133	9.15	50.0 (nominal)
0.9	320	16.174	16.782	3.76	50.0 (nominal)

The measured overhead (2.6%–9.3%, noisy, no clear monotone trend with N in this range) is far smaller than the 50% figure the evaluation-count argument alone would suggest. The reason is structural, not accidental: at each step the $O(n)$ history convolution (the predictor sum and the H term) is computed once and shared by both variants, exactly as designed in Section 2; the second sweep adds only one extra scalar evaluation of f per step. Because the shared history work is $O(n)$ at step n and the extra evaluation is $O(1)$, the 50% figure (a ratio of evaluation $counts$, 3 vs. 2) systematically overstates the wall-clock premium once the shared $O(n)$ work dominates total per-step cost — which is precisely the regime this O(N²) uniform-grid implementation is in, for large N (Section 9's fifth limitation).

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

This measurement is scoped narrowly and should not be over-read. It uses a specific Python/NumPy implementation, a single (unspecified-to-the-reader) machine, and only the scalar Table 1 benchmark; a compiled implementation, a different history-sum data structure, or Example 5's coupled system could shift the overhead in either direction, and no comparison against a fully-iterated fixed-point corrector or a Newton-corrected variant is attempted here (that comparison, per Section 9, remains future work). What this measurement does establish is the qualitative point the evaluation-count argument could not: that the second sweep's wall-clock premium is consistently a small fraction of the 50% evaluation-count figure, not merely predicted to be so.

7.7 Floating-Point Cancellation in the Weights

The floating-point cancellation affecting the weights at large n has so far gone unexamined; here we put a number on it. Writing $m := n - j$ for the distance back into the history; the predictor weight and the corrector history weight involve, respectively, the first and second differences

$b(m) = (m+1)^{\alpha} - m^{\alpha}$ and $a(m) = (m+2)^{\alpha+1} - 2(m+1)^{\alpha+1} + m^{\alpha+1}$

Both are differences of nearly-equal large numbers when m is large, so the true value is much smaller than either term being subtracted — the textbook setting for catastrophic cancellation. A standard error propagation argument (each evaluated power carries relative error $\leq \varepsilon$, $\varepsilon \approx 2.2 \times 10^{-16}$ in IEEE double precision) gives, since $b(m)$ is itself $O(m^{\alpha-1})$ while each term subtracted is $O(m^{\alpha})$, an absolute error of order εm^{α} against a true value of order $m^{\alpha-1}$, i.e. a relative error growing like $\varepsilon \cdot m$. The same argument applied to $a(m)$ — itself $O(m^{\alpha-1})$ while each subtracted term is $O(m^{\alpha+1})$ — gives relative error growing like $\varepsilon \cdot m^2$, two full powers of m worse. Table 6 confirms both scalings against a 60-digit reference computed in $mpmath$.

Table 6: Cancellation in the weights, naive vs. stable evaluation, against a 60-digit reference (α = 0.7; α = 0.4, 0.9 are qualitatively identical and omitted).

m = n−j	b naive err.	b stable err.	a naive err.	a stable err.
10	9.63e-16	1.60e-16	1.42e-15	2.89e-06
100	5.03e-15	3.11e-17	1.21e-12	3.15e-11
1,000	1.31e-14	5.41e-18	1.87e-11	3.28e-16
10,000	1.43e-12	3.63e-17	1.16e-08	3.30e-17
100,000	6.29e-12	1.38e-16	1.16e-06	2.58e-16
1,000,000	3.93e-11	8.77e-17	8.20e-05	2.17e-16
10,000,000	5.32e-10	4.77e-17	5.62e-03	2.96e-16
100,000,000	6.11e-09	8.82e-17	6.49e-01	2.40e-16

As predicted, b's naive relative error grows roughly linearly in m (from 10^{-16} at $m = 10$ to 10^{-9} at $m = 10^8$), while a's grows roughly quadratically and becomes *catastrophic* well before $m = 10^8$: the naive relative error exceeds 1 (100%, a meaningless result) at $m \approx 10^8$ for $\alpha = 0.4$. The stable reformulations — $b(m) = m^{\alpha} \cdot expm1\big(\alpha \cdot log1p(1/m)\big)$ for b, and a five-term Taylor series in $1/m$ of the bracket $(1 + 2/m)$ form (coefficients $p(p-1)$, $p(p^2 - 3p + 2)$, etc., verified symbolically) for a — hold at or near machine precision across the entire tested range, confirming both the diagnosis and the fix.

Neither reformulation is uniformly best: at small m (below roughly $10^2 - 10^3$), the Taylor series for a has not yet converged and is less accurate than the naive formula, which itself has no cancellation problem at small m. The practical recommendation is a hybrid switch: use the direct difference formula for m below a threshold (10^3 is a safe, conservative choice — Table 6 shows both formulas agree to better than 10^{-10} relative error there) and the stable series above it.

This cancellation did *not* affect any result reported earlier in this paper: Table 1 and Table 1b use $N \leq 320$, so m never exceeds 320, where the naive formula's relative error is still below 10^{-10} — utterly negligible next to the 10^{-6}-scale discretization error the method actually reports; Table 4's $N \leq 256$ is smaller still. The stable reformulation is offered as a preventive fix for larger-scale future use (e.g. the graded-mesh or fast-kernel-compression extensions of Section 9), not as a correction to any number already published here.

7.8 A Verified, Faster-Than-O(N²) History Evaluatio

The preceding sections have treated the $O(N^2)$ cost of uniform-grid history accumulation as a loose end. Fast, FFT-accelerated fractional Adams solvers are not new — Garrappa's fde12 [15] and the graded-mesh fast method of Su and Zhou [31] already exist — so what follows is a from-scratch, independently verified reimplementation and complexity analysis tailored to this paper's own two-sweep scheme, not a claim to have introduced fast fractional Adams methods as a category. Even so, it stops short of the optimal $O(N \log N)$ achievable with a full multilevel scheme — that remainder is left as future work, consistent with Section 9's existing recommendation to combine this method with fast kernel compression.

Both history sums used by the predictor and corrector — $\sum_{j=0}^{n} b(n-j)F$ and the analogous sum with weight a — are linear convolutions, since the weight depends only on the index difference $n - j$. This is what makes an FFT-based acceleration legitimate rather than approximate: unlike sum-of-exponentials kernel compression, which trades an additional,

Stallion Journal for Multidisciplinary Associated Research Studies
ISSN (Online): 2583-3340
Volume-5 Issue-4 || August 2026 || PP. 50-69 https://doi.org/10.55544/sjmars.5.4.6

separately-controlled approximation error for speed, convolution via FFT computes the *same* sum, exactly, up to ordinary floating-point round-off.

The scheme implemented and verified here is a simple two-level block decomposition, not the optimal multilevel (dyadic) construction: history indices are processed in blocks of size $B \approx \sqrt{N}$; each completed block's contribution to *all* remaining future steps is computed in one FFT-based convolution of length $O(N)$, giving $O(N/B)$ such convolutions and total cost $O\big((N/B) \cdot N\log N\big)$, which at $B = \sqrt{N}$ is $O(N^{1.5}\log N)$ — asymptotically better than $O(N^2)$ but short of the $O(N \cdot (\log N)^2)$ the classical Hairer–Lubich–Schlichte fast-convolution algorithm [42] achieves, and which Garrappa's fde12 already implements in production code [15].

Table 7: Fast (block-FFT) vs. direct method: error match at Table 1's scale, confirming exactness.

α	N	direct error	fast error	max err. diff.
0.4	40	9.817226e-05	9.817226e-05	2.22e-16
0.4	320	1.588568e-06	1.588568e-06	4.44e-16
0.7	40	9.206529e-05	9.206529e-05	8.88e-16
0.7	320	1.527362e-06	1.527362e-06	4.44e-16
0.9	40	5.161821e-05	5.161821e-05	4.44e-16
0.9	320	9.607152e-07	9.607152e-07	8.88e-16

Table 8: Measured wall-clock speedup, full two-sweep method, direct vs. fast history evaluation (α = 0.7).

N	direct (ms)	fast (ms)	speedup
320	8.49 ms	4.28 ms	1.98×
640	21.54 ms	9.51 ms	2.26×
1,280	53.47 ms	18.23 ms	2.93×
2,560	143.12 ms	36.50 ms	3.92×
5,120	452.46 ms	75.82 ms	5.97×

Table 7 confirms exactness: the fast method reproduces Table 1's published errors to the last reported digit, and agrees with a fresh direct-method run to floating-point round-off ($\leq 9\times10^{-16}$) at every tested (α, N). Table 8 confirms the speed-up is real and growing: at N = 320 (the largest mesh used anywhere in this paper) the benefit is modest (about 2×, since FFT overhead is non-trivial at this scale and the block size $B \approx \sqrt{N}$ is small), but it grows steadily — nearly 6× by N = 5,120 — exactly the regime where the $O(N^2)$ cost this paper flagged as a limitation actually starts to bite.

This is a partial, not a complete, resolution. What is established: an exact (not approximate), verified, asymptotically-better-than-quadratic history evaluation, with measured wall-clock benefit that grows with N as the theory predicts. What is not established: the optimal $O(N\log N)$ complexity a full dyadic block scheme would give; an optimized (rather than heuristic $B \approx \sqrt{N}$) block size; extension to Example 5's vector system; and combination with the graded meshes Section 9 already recommends for the separate (regularity, not complexity) problem of non-smooth solutions.

VIII. DISCUSSION

The modified method occupies a useful middle position. Compared with the one-sweep PECE algorithm, it costs little extra whenever evaluating the function is cheap, precisely because the convolution history is shared between sweeps; Section 7.3 pins down exactly how little it costs and how much it buys. Compared with a corrector iterated to full convergence, it sidesteps an open-ended nonlinear loop. The contraction factor q_h offers a hands-on diagnostic: a small q_h means a second sweep may not be needed; if q_h is below but near one, the second sweep can materially reduce the corrector defect; if q_h exceeds one, fixed-point sweeps are theoretically unsafe.

This method leaves fractional time-stepping's central complexity problem unsolved. The uniform-grid implementation used throughout Sections 7.1–7.7 still accumulates history quadratically; Section 7.8 offers a first, verified move toward faster evaluation (a two-level block-FFT scheme running at O(N^1.5logN)), but long simulations would still gain from pairing the two-sweep corrector with sum-of-exponentials methods, short-memory truncation, or the optimal O(NlogN) convolution-quadrature acceleration [10,11,30,32,42] that Section 7.8 does not reach. Nor does the method automatically restore a high order for nonsmooth solutions — graded meshes and corrected starting weights are still needed there [29–32].

This, in short, is the point of Table A (Section 2.1): the novelty claimed here is narrow — repeated corrector evaluation is standard, already part of the P(EC)^{m}E family — and its value lies instead in the exact two-sweep formulation's explicit contraction diagnostic, its tailored stability bound, its transparent comparison with PECE, and, new in this revision,

measured-order evidence reported honestly, gap and all. The database search for related terminology recommended in Section 1 still stands.

IX. LIMITATIONS AND FUTURE WORK

First, the theory assumes global Lipschitz continuity; nonlinearities that are only locally Lipschitz need an invariant region or an a priori bound instead. For Example 5 this gap is addressed via a computed a priori box (L=3 on B=[0,2]², Section 7.4), and Section 5.1 supplies the general link: Theorem 1′ for the scalar case, and Theorem 1″ for vector-valued systems such as Example 5's, both giving the same order and constant as Theorem 1 once a bounding set and its local Lipschitz constant are known. What remains open is deriving that bounding set analytically — e.g. via an invariant-region or fractional Lyapunov-type argument — rather than by inspecting the computed trajectory. Second, the proof sketches in Sections 5–6 are given as full lemma-level arguments (Lemma 1.4, Lemma 1.5, Theorems 1–2), closing gaps that a purely formal, proof-sketch treatment leaves open. Section 7.5 computes Theorem 1's constant C at the benchmark problem and finds it valid but impractical — inflated by roughly 200 orders of magnitude by the (necessarily pessimistic) global Lipschitz constant, not by h — so only the order p, not C, should be read as a practical error estimate. Third, Section 7.4 supplements the original smooth benchmark with four further tests: a weakly singular case confirming Theorem 1's order bound is sharp, a genuinely non-smooth linear Mittag–Leffler case, a second smooth exponent, and a coupled nonlinear system with no closed-form solution; a stiff test and a Newton-corrected comparison remain open. Fourth, floating-point cancellation in the weights at very large n (Section 7.7) is analyzed here and given a stable reformulation, though it affects no result reported here, since Table 1 and Table 4 stay at N≤320, well short of where cancellation matters. Fifth, the O(N²) uniform-grid history sum has a verified O(N^1.5logN) fast alternative (Section 7.8, with a measured 2×–6× speedup); the optimal O(NlogN) dyadic scheme and extension to systems remain open. Sixth, wall-clock timing (Section 7.6) shows the two-sweep overhead (2.6–9.3%) is much smaller in practice than the nominal 50% suggested by evaluation counts alone, since the shared history convolution dominates per-step cost; a runtime comparison against a fully-converged or Newton-corrected corrector, and confirmation this is not a Python/NumPy-specific artifact, remain open.

Directions worth pursuing next include pairing two-sweep defect correction with graded meshes and fast kernel compression; building an adaptive criterion that calls for the second sweep only once the estimated corrector defect (via q_h) is large, as motivated in Section 7.3; investigating Newton-corrected variants for stiff nonlinearities; extending the stability analysis to systems and variable-order operators; and, motivated directly by Section 7.5's finding that Theorem 1's global a priori constant is correctly ordered but numerically unusable, pursuing a localized, a posteriori bound — one that replaces the worst-case global Lipschitz constant Λ with a running estimate tied to the trajectory actually computed — as a genuine step-size and error-control tool, rather than relying on the q_h diagnostic alone.

X. CONCLUSION

This paper has developed a two-sweep fractional Adams predictor–corrector scheme for nonlinear Caputo initial-value problems, framed explicitly as the $m = 2$ case of the known P(EC)^{m}E family rather than a new algorithm. Reusing the history convolution makes the second sweep inexpensive, precisely one extra right-hand-side evaluation per step. Given the Lipschitz and regularity assumptions, the modification can be shown to keep the classical global order $min\{2,1 + \alpha\}$ (Theorem 1, established in full via a self-contained discrete fractional Grönwall inequality, Lemma 1.5), remains stable in the Mittag–Leffler sense on any finite interval with an explicit constant $C_0 = 1 + q + q^2$ (Theorem 2), and shrinks the nonlinear corrector defect by a further factor $q_h = Lh^\alpha/\Gamma(\alpha + 2)$ per sweep. The numerical evidence covers five examples: the original smooth manufactured problem, with its measured EOC and its theory–data gap at low α; a second smooth exponent showing the gap is not an artifact of one particular manufactured solution; a linear Mittag–Leffler test whose genuinely low-order regularity means the second sweep need not help, exactly as Corollary 1 predicts; a weakly singular companion test ($y = 1 + t^{2\alpha}$) whose observed order degrades to 2α, confirming that Theorem 1's bound is genuinely sharp and not merely a safe overestimate; and a coupled nonlinear system with no closed-form solution, where the second sweep pays off the most (roughly a 4x reduction) even as a third sweep turns out, measurably though only slightly, worse — a benign sign-cancellation effect explained in Section 7.4, not something to worry about, but not nothing either. Taken as a whole, this evidence favors applying the second sweep adaptively — guided by q_h and by whether the main error source is the corrector iteration or the quadrature/regularity limit, rather than applying it every time. The method is best thought of as a clean, reproducible, fully proved baseline, honest about its own scope; its core contributions are its q_h diagnostic, its cost accounting, and the completed proofs — more so than the two-sweep idea by itself.

Declarations: Funding: No external funding was received for this work. Conflict of interest: The author declares no conflict of interest.

REFERENCES

[1] Oldham, K. B., & Spanier, J. (1974). *The fractional calculus*. Academic Press.
[2] Miller, K. S., & Ross, B. (1993). *An introduction to the fractional calculus and fractional differential equations*. Wiley.
[3] Samko, S. G., Kilbas, A. A., & Marichev, O. I. (1993). *Fractional integrals and derivatives*. Gordon and Breach.
[4] Podlubny, I. (1999). *Fractional differential equations*. Academic Press.
[5] Kilbas, A. A., Srivastava, H. M., & Trujillo, J. J. (2006). *Theory and applications of fractional differential equations*. Elsevier.
[6] Diethelm, K. (2010). *The analysis of fractional differential equations*. Springer.
[7] Mainardi, F. (2010). *Fractional calculus and waves in linear viscoelasticity*. Imperial College Press.
[8] Li, C., & Zeng, F. (2015). *Numerical methods for fractional calculus*. Chapman and Hall/CRC.
[9] Caputo, M. (1967). Linear models of dissipation whose Q is almost frequency independent. *Geophysical Journal International, 13*, 529–539.
[10] Lubich, C. (1986). Discretized fractional calculus. *SIAM Journal on Mathematical Analysis, 17*, 704–719.
[11] Lubich, C. (1988). Convolution quadrature and discretized operational calculus I. *Numerische Mathematik, 52*, 129–145.
[12] Diethelm, K., Ford, N. J., & Freed, A. D. (2002). A predictor-corrector approach for the numerical solution of fractional differential equations. *Nonlinear Dynamics, 29*, 3–22.
[13] Diethelm, K., Ford, N. J., & Freed, A. D. (2004). Detailed error analysis for a fractional Adams method. *Numerical Algorithms, 36*, 31–52.
[14] Diethelm, K., & Ford, N. J. (2002). Analysis of fractional differential equations. *Journal of Mathematical Analysis and Applications, 265*, 229–248.
[15] Garrappa, R. (2018). Numerical solution of fractional differential equations: A survey and a software tutorial. *Mathematics, 6*(2), 16.
[16] Garrappa, R. (2010). On linear stability of predictor-corrector algorithms for fractional differential equations. *International Journal of Computer Mathematics, 87*, 2281–2290.
[17] Ford, N. J., & Simpson, A. C. (2001). The numerical solution of fractional differential equations: Speed versus accuracy. *Numerical Algorithms, 26*, 333–346.
[18] Ford, N. J., Morgado, M. L., & Rebelo, M. (2013). Nonpolynomial collocation approximation of solutions to fractional differential equations. *Fractional Calculus and Applied Analysis, 16*, 874–891.
[19] Brunner, H. (2004). *Collocation methods for Volterra integral and related functional differential equations*. Cambridge University Press.
[20] Baker, C. T. H. (1977). *The numerical treatment of integral equations*. Clarendon Press.
[21] Hairer, E., Nørsett, S. P., & Wanner, G. (1993). *Solving ordinary differential equations I*. Springer.
[22] Dahlquist, G. (1963). A special stability problem for linear multistep methods. *BIT, 3*, 27–43.
[23] Gorenflo, R., Luchko, Y., & Mainardi, F. (1999). Analytical properties and applications of the Wright function. *Fractional Calculus and Applied Analysis, 2*, 383–414.
[24] Gorenflo, R., Kilbas, A. A., Mainardi, F., & Rogosin, S. V. (2014). *Mittag-Leffler functions, related topics and applications*. Springer.
[25] Li, Y., Chen, Y. Q., & Podlubny, I. (2009). Mittag-Leffler stability of fractional-order nonlinear dynamic systems. *Automatica, 45*, 1965–1969.
[26] Li, Y., Chen, Y. Q., & Podlubny, I. (2010). Stability of fractional-order nonlinear dynamic systems: Lyapunov direct method and generalized Mittag-Leffler stability. *Computers & Mathematics with Applications, 59*, 1810–1821.
[27] Liao, H., McLean, W., & Zhang, J. (2018). A discrete Grönwall inequality with applications to numerical schemes for subdiffusion problems. *SIAM Journal on Numerical Analysis, 57*, 218–237.
[28] Stynes, M., O'Riordan, E., & Gracia, J. L. (2017). Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. *SIAM Journal on Numerical Analysis, 55*, 1057–1079.
[29] Jin, B., Lazarov, R., & Zhou, Z. (2016). Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM Journal on Scientific Computing, 38*, A146–A170.
[30] Zeng, F., Li, C., Liu, F., & Turner, I. (2015). The use of finite difference/element approaches for solving the time-fractional subdiffusion equation. *SIAM Journal on Scientific Computing, 35*, A2976–A3000.
[31] Su, X., & Zhou, Y. (2022). A fast high-order predictor-corrector method on graded meshes for solving fractional differential equations. *Fractal and Fractional, 6*, 516.
[32] Arshad, S., Saleem, I., Akgül, A., Huang, J., Tang, Y., & Eldin, S. M. (2023). A novel numerical method for solving the Caputo-Fabrizio fractional differential equation. *AIMS Mathematics, 8*, 9535–9556.
[33] Moumen, A., Mennouni, A., & Bouye, M. (2024). Contributions to the numerical solutions of a Caputo fractional differential and integro-differential system. *Fractal and Fractional, 8*, 201.

[34] Su, X., & Zhou, Y. (2024). High-order predictor-corrector discretizations on nonuniform meshes for Caputo problems. *Fractional Calculus and Applied Analysis*.
[35] Ye, H., Gao, J., & Ding, Y. (2007). A generalized Gronwall inequality and its application to a fractional differential equation. *Journal of Mathematical Analysis and Applications, 328*, 1075–1081.
[36] Yousuf, M., et al. (2025). Predictor-corrector scheme for multi-dimensional time-space fractional reaction-diffusion models. *Numerical Methods for Partial Differential Equations, 41*(4), e70022. https://doi.org/10.1002/num.70022
[37] Bu, S., & Jeon, Y. (2025). Higher-order predictor-corrector methods with an enhanced predictor for fractional differential equations. *International Journal of Computer Mathematics, 102*(8), 1057–1080.
[38] Bu, S., & Jeon, Y. (2025). Higher-order predictor-corrector methods for fractional Benjamin-Bona-Mahony-Burgers' equations. *Journal of Applied Mathematics and Computing, 71*, 1–30.
[39] Yang, Y., & Yan, Y. (2025). A fractional Adams method for Caputo fractional differential equations with modified graded meshes. *Mathematics, 13*(5), 891.
[40] Madduri, H., & Gande, N. R. (2025). Higher order graded mesh scheme for time fractional differential equations. *Computational and Applied Mathematics, 44*, 220.
[41] Danca, M.-F. (2026). A high order method for Caputo fractional differential equations. Manuscript submitted for publication; see also companion preprint, *Improved Matlab code for Lyapunov exponents of fractional order systems* (2026), arXiv:2604.09334.
[42] Hairer, E., Lubich, C., & Schlichte, M. (1985). Fast numerical solution of nonlinear Volterra convolution equations. *SIAM Journal on Scientific and Statistical Computing, 6*(3), 532–541.