Asymptotically Compatible Fractional Grönwall Inequality and its Applications

Daopeng Yin · Liquan Mei

Abstract In this work, we will give proper estimates for the discrete convolution complementary (DCC) kernels, which leads to the asymptotically compatible fractional Grönwall inequality. The consequence can be applied in analysis of the stability and pointwise-in-time error of difference-type schemes on a non-uniform mesh. The pointwise error have explicit bound when a non-uniform time grid is given by specific scale function e.g. graded mesh, can be given directly. Numerical experiments towards the conclusion of this work validate the error analysis.

1 Introduction

In this paper, we provide asymptotically compatible fractional Gröwanll inequality as a useful tool for time-fractional evolution equations. The discrete complementary convolution (DCC) kernels were first introduced in the literature [15,17] to prove a class of fractional order inequalities. Due to the lack of proper estimation for DCC kernels, there are certain flaws in their asymptotic compatibility and optimal error estimation. Its asymptotic expression is an open problem up to now [28] which will be rigorously proved in this work. Consequently, we provide pointwise-in-time error estimates of the non-uniform time mesh L1-scheme for solving time-fractional equations involving non-dissipation term under general regularity assumptions, based on the mentioned above results coupled with the discrete convolution summation (DCS) estimation. In such an analysis framework,

D. Yin
School of Mathematics and Statistics. Xi'an Jiaotong University. No.28, West Xianning Road, Xi'an, Shaanxi, 710049, P.R. China

L. Mei
School of Mathematics and Statistics. Xi'an Jiaotong University. No.28, West Xianning Road, Xi'an, Shaanxi, 710049, P.R. China *Corresponding author*: E-mail: lqmei@mail.xjtu.edu.cn

this class of methods can be extended to other time difference-type schemes as well. Compared to the construction and analysis of barrier function, this route is considerably less expensive to analyze, yet our theoretical results concur with Kopteva's results [13]. Our main works are listed below

- The continuous counterpart of DCC will be given, and their coefficient depends of mesh scale and the ratio of adjacent grids was rigorously analyzed.
- The asymptotically compatible fractional Grönwall inequality will be proved which will be applied to the analysis of stability and error for time-fractional reaction-diffusion equations.
- In providing grid functions, such as graded mesh, the explicit expression for pointwise-in-time error estimation can be strictly written based on the sharp estimate for DCS.

As an extension of the parabolic problems, Caputo time-fractional subdiffusion equations have a wide range of applications. In this article, the following time-fractional order reaction-diffusion equation is taken into consideration. For $\kappa > 0$,

$$
\begin{align}
\partial_t^\alpha u - \Delta u =& \kappa u + f(\mathbf{x}, t), & \mathbf{x} \in \Omega \times (0, T], & \tag{1a} \\
u(\mathbf{x}, 0) =& u_0(\mathbf{x}), & \mathbf{x} \in \Omega, & \tag{1b} \\
u(\mathbf{x}, t) =& \phi(\mathbf{x}, t), & \mathbf{x} \in \partial\Omega \times [0, T], & \tag{1c}
\end{align}
$$

where the $\Omega \subset \mathbb{R}^n$, which is a n-dimensional bounded Lipschitz domain and the time-fractional derivative utilizes the definition of Caputo-sense

$$
\partial_t^\alpha u(\mathbf{x}, t) := \frac{1}{\Gamma(1-\alpha)} \int_0^t \frac{\partial u(\mathbf{x}, s)}{\partial s} \frac{1}{(t-s)^\alpha} ds, \ 0 < \alpha < 1. \tag{2}
$$

When $\alpha \to 1$, the above-mentioned problem will be degenerated into a classical parabolic equation. Typically denote $\omega_\alpha(s) := s^{\alpha-1}/\Gamma(\alpha)$ for $s \geq 0$ which also be called Gelfand-Shilov function, and the above definition has an alternative convolution form

$$
\partial_t^\alpha u(\mathbf{x}, t) = \int_0^t \omega_{1-\alpha}(t-s) \partial_s u(\mathbf{x}, s) ds.
$$

In time discretization, the t_k for $1 \leq k \leq N$ represent the time discrete nodes on general non-uniform meshes and its k-th time step $\tau_k = t_k - t_{k-1}$. The temporal semi-discrete scheme of continuous (1) is governed as

$$
D_\tau^\alpha u^{n-\sigma} = \Delta u^{n-\sigma} + \kappa u^{n-\sigma} + f(\mathbf{x}, t_{n-\sigma}). \tag{3}
$$

The discrete Caputo operator $D_\tau^\alpha u^{n-\sigma} := \sum_{k=1}^n a_{n-k}^{(n)} \nabla_\tau u^k$, where the difference operator $\nabla_\tau u^k := u^k - u^{k-1}$. Here the expression of discrete convolution (DC) kernels $a_{n-k}^{(n)}$ derived from specific discretization strategy e.g.

- the DC kernels via L1 scheme is

$$
a_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} \omega_{1-\alpha}(t_n - s) ds, \quad \text{for } 1 \leq k \leq n.
$$

– the DC kernels via fast L1 scheme is

$$a_{n-k}^{(n)} := \begin{cases} \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \omega_{1-\alpha}(t_n - s)\mathrm{d}s, & \text{for } k = n, \\ \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} \sum_{l=1}^{N_q} \mathrm{d}s, & \text{for } 1 \le k \le n-1. \end{cases}$$

– the DC kernels via L21-σ scheme is

$$a_{n-k}^{(n)} := \mathsf{a}_{n-k}^{(n)} + r_{k-1}\mathsf{b}_{n-k+1}^{(n)} - \mathsf{b}_{n-k}^{(n)}, \text{ for } 1 \le k \le n,$$

where

$$\mathsf{a}_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{\min\{t_k, t_{n-\sigma}\}} \omega_{1-\alpha}(t_{n-\sigma} - s)\mathrm{d}s, \text{ for } 1 \le k \le n-1,$$

$$\mathsf{b}_{n-k}^{(n)} := \frac{2}{\tau_k(\tau_k + \tau_{k+1})} \int_{t_{k-1}}^{t_k} (s - t_{k-\frac{1}{2}})\omega_{1-\alpha}(t_{n-\sigma} - s)\mathrm{d}s, \text{ for } 1 \le k \le n-1,$$

with setting $\mathsf{b}_0^{(n)} = \mathsf{b}_n^{(n)} \equiv 0$.

Our analysis focuses on the temporal discrete scheme, thus the spacial discretization can couple with finite element methods (FEMs) [3, 26], finite difference methods (FDMs) [25], spectral methods (SMs) [22, 27] and so on. Since without any restriction for time scales, the time adaptive strategy can used.

There are some discussions on the regularity of the solutions to the subdiffusion equations in recent literature. For the linear subdiffusion problem, Stynes et al [23] proved that while $u_0 \in D(\mathcal{L}^{5/2})$ and $\|f(\cdot, t)\|_{\mathcal{L}^{5/2}} + \|f_t(\cdot, t)\|_{\mathcal{L}^{1/2}} + t^\rho \|f_{tt}(\cdot, t)\|_{\mathcal{L}^{1/2}} \le C$, then $|\partial_t^m u(\cdot, t)| \lesssim 1 + t^{\alpha - m}$, for $m = 0, 1, 2$. Jin et al [7] analyzed that when $u_0 \in H_0^1(\Omega) \cap H^2(\Omega)$ and $\partial_t^m f \in (0, T; L^2(\Omega))$ then $\|\partial_t^m (u(\cdot, t) - u_0)\|_{L^2(\Omega)} \lesssim t^{\alpha - m}, m \ge 0$. Al-Maskari et al [1] suggested that the initial data $u_0 \in \dot{H}^\nu(\Omega)$ can prove $\|\partial_t u(t)\|_{L^2(\Omega)} \lesssim t^{\nu\alpha/2 - 1}$. This weakly singular solution can also be a more generalized form, Zhang et al [29] gave error analysis based on a general assumption $u(t) - u(0) = \sum_{k=1}^m \hat{u}_k t^{\beta_k} + \tilde{u}(t)t^{\beta_{m+1}}$, where $0 < \beta_1 < \cdots < \beta_m < \beta_{m+1}$ and $\tilde{u}(t) \in L^2([0, T]; X)$. To this end, we give general assumptions on the regularity of the solution

$$\|\partial_t^m u\| \lesssim 1 + t^{\beta - m}, \text{ for } m = 1, 2, \text{ and } \beta \in (0, 1) \cup (1, 2],$$

where β is considered to be the regularity parameter in the general case [14]. When $\beta = \alpha$, it is the most common case.

The error order of a time discrete scheme is not constant, that is, the error order on various time levels has varying error orders, due to the non-local character of the Caputo derivative. Many academics have expressed concern about this aspect. When $f(x, t, u) = 0$, J. Gracia et al [5] analyzed them numerical scheme have error result $\tau t_n^{\alpha - 1}$. Jin et al [6] show that the time error of L_1 scheme should be τt_n^{-1} with $u_0 \in L^2(\Omega)$. These error estimates suggest that time factor t_n should be included in the error boundaries of time-fractional. Currently, there are two basic categories of analytical methods that may be applied to provide a result that is more cohesive. For one thing, N. Kopteva et al provided a general theoretical framework for pointwise-in-time error estimation of time-fractional diffusion problems using the discrete comparison principle which includes $L1, L2, L21 - \sigma$ schemes [10, 12, 13] and also semi-linear problem [11]. For another, Liao et al analyzed the

error of the Caputo derivative by constructing DCC kernels and developing the fractional version of the discrete Grönwall inequality [15, 18, 18], which is a popular tool for time evolution equations. However, the pointwise-in-time error estimation under the general regularity of the solution using the L1 discrete scheme with non-uniform time mesh is still very sparse in related work under the frame of DCC.

The numerical techniques may be stated as a DCS since the Fourier transform can be used to infer that the definition of the Caputo fractional order derivative originates from the continuous convolution form. Thus, the sharp estimation of DCS is a current topic of interest for related academics [24]. There are certain works by C. Lubich [19, 20, 21] that are linked to the investigation of this historical question. Numerous applications of this finding may be made to the research of numerical stability and convergence in time-fractional order problems.

The rest of the paper is organized as follows. In section 2, we will prove the asymptotical estimate for DCC and then obtain the asymptotically compatible fractional Grönwall inequality. In section 3, we present some applications, e.g. the numerical stability and pointwise-in-time error of L1 scheme of time-fractional parabolic equations. Section 4 will show several numerical examples to confirm the theoretical results. There are two appendices at the end of this article including appendix A give a sharp estimation of DCS and rigorous proof and appendix B prove the necessary condition $\frac{1}{2}D_\tau^\alpha \|u^n\|^2 \le (D_\tau^\alpha u^n, u^n)$ with discrete Caputo operator.

2 Asymptotically Compatible Fractional Grönwall Inequality

In the first subsection of this section, we will rigorously prove the asymptotic expression of DCC. The proof in uniform grids has been given by [8, Theorem 3.3], but for general non-uniform grids, no one has yet completed. Our proof method is relatively concise and more general. Continuing with this result, we will provide the asymptotically compatible Grönwall inequalities in the second subsection and their applications in stability and error analysis of semi-discrete schemes for time-fractional parabolic equations in the next section.

2.1 The asymptotic analysis of the DCC

In this subsection, we provide an asymptotic estimate of DCC for general non-uniform mesh and its continuous counterpart will be used to prove the asymptotically compatible fractional Grönwall inequality in the next subsection.

Firstly, we review the definition of DCC and express it in matrix form. The DCC defined by $(n-k)$-terms recurrence relations:

$$p_{n-k}^{(n)} := \frac{1}{a_0^{(n)}} \begin{cases} 1, & \text{if } k = n, \\ \sum\limits_{j=k+1}^{n} p_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right), & \text{if } 1 \le k < n. \end{cases} \tag{4}$$

The equivalent matrices form is

$$\mathbf{P}_n \mathbf{A}_n \mathbf{D}_n = \mathbf{I}_n, \tag{5}$$

where

$$
\mathbf{P}_n := \begin{bmatrix} p_0^{(1)} & 0 & \dots & 0 \\ p_1^{(2)} & p_0^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ p_{n-1}^{(n)} & p_{n-2}^{(n)} & \dots & p_0^{(n)} \end{bmatrix}, \mathbf{A}_n := \begin{bmatrix} a_0^{(1)} & 0 & \dots & 0 \\ a_1^{(2)} & a_0^{(2)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n-1}^{(n)} & a_{n-2}^{(n)} & \dots & a_0^{(n)} \end{bmatrix},
$$

with the difference matrices

$$
\mathbf{D}_n := \begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{bmatrix} \quad \text{and identity matrix } \mathbf{I}_n.
$$

For all $1 \le k \le n$, the relation $\sum_{j=k}^{n} p_{n-j}^{(n)} a_{j-k}^{(j)} = 1$ constantly holds is equivalent to

$$
\mathbf{P}_n \mathbf{A}_n = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 1 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \dots & 1 \end{bmatrix} = \mathbf{D}_n^{-1},
$$

which can obtain from (5) directly.

Before starting the discussion, it is necessary to declare some notations in this article, the adjoint time step ratio $\rho = \max_{1 \le n \le N-1} \{ \frac{\tau_{n+1}}{\tau_n} \}$, and maximum time step scale $\tau = \max_{1 \le n \le N} \{ \tau_n \}$. The asymptotic expression of DCC is

$$
\tilde{p}_{n-k}^{(n)} := \int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - s) \mathrm{d}s.
$$

Lemma 1 *For all $1 \le k+1 \le j \le n$,*

$$
\max_{k+1 \le j \le n} \left\{ \frac{\tilde{p}_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right)}{-\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \int_{t_{k-1}}^{t_k} \omega'_{1-\alpha}(t - s) \mathrm{d}s \mathrm{d}t} \right\} := C_{\rho,\tau} \le \frac{\rho+1}{2} + \mathcal{O}(\tau).
$$

Proof For convenience, we denote

$$
q_{n-k,j}^{(n)} := \frac{\tilde{p}_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right)}{-\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \int_{t_{k-1}}^{t_k} \omega'_{1-\alpha}(t - s) \mathrm{d}s \mathrm{d}t},
$$

which illustrate the similarity between $\tilde{p}_{n-j}^{(n)}(a_{j-k-1}^{(j)} - a_{j-k}^{(j)})$ and the denominator part. Equivalently,

$$
q_{n-k,j}^{(n)} = \frac{\frac{1}{\tau_j} \tilde{p}_{n-j}^{(n)} \cdot \frac{1}{\tau_k} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right)}{-\frac{1}{\tau_j} \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \cdot \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} \omega'_{1-\alpha}(t - s) \mathrm{d}s \mathrm{d}t}.
$$

Notice that

$$\frac{\tilde{p}_{n-j}^{(n)}}{\tau_j} = \int_0^1 \omega_\alpha(t_n - t_j + \theta\tau_j)\mathrm{d}\theta = \omega_\alpha(t_n - t_j) + \mathcal{O}(\tau_j),$$

$$\frac{a_{j-k-1}^{(j)} - a_{j-k}^{(j)}}{\tau_k} = \frac{1}{\tau_k}\int_0^1 \sum_{m=1}^{\infty} \frac{\omega_{1-\alpha}^{(m)}(t_j - t_{k+1})}{m!}[(\theta\tau_{k+1})^m - (\tau_{k+1} + \theta\tau_k)^m]\mathrm{d}\theta$$

$$= -\frac{1}{2}\omega_{1-\alpha}'(t_j - t_{k+1})(r_{k+1} + 1) + \mathcal{O}(\tau_k),$$

and

$$-\frac{1}{\tau_j}\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \cdot \frac{1}{\tau_k}\int_{t_{k-1}}^{t_k} \omega_{1-\alpha}'(t - s)\mathrm{d}s\mathrm{d}t$$

$$= -\int_0^1 \omega_\alpha(t_n - t_j + (1-\varphi)\tau_j)\int_0^1 \omega_{1-\alpha}'(t_j - (1-\varphi)\tau_j - t_{k+1} + \tau_{k+1} + \theta\tau_k)\mathrm{d}\theta\mathrm{d}\varphi.$$

Then

$$q_{n-k,j}^{(n)} = \frac{(\omega_\alpha(t_n - t_j) + \mathcal{O}(\tau_j)) \cdot \left(-\frac{1}{2}\omega_{1-\alpha}'(t_j - t_{k+1})(r_{k+1} + 1) + \mathcal{O}(\tau_k)\right)}{-(\omega_\alpha(t_n - t_j) + \mathcal{O}(\tau_j)) \cdot (\omega_{1-\alpha}'(t_j - t_k) + \mathcal{O}(\tau_j) + \mathcal{O}(\tau_{k+1}) + \mathcal{O}(\tau_k))}$$

$$= \frac{r_{k+1} + 1}{2} + \mathcal{O}(\tau_j) + \mathcal{O}(\tau_{k+1}) + \mathcal{O}(\tau_k).$$

Therefore, $q_{n-k,j}^{(n)} = (r_{k+1} + 1)/2 + \mathcal{O}(\tau_j) + \mathcal{O}(\tau_{k+1}) + \mathcal{O}(\tau_k)$ and it follows that $q_{n-k,j}^{(n)} \leq (\rho + 1)/2 + \mathcal{O}(\tau)$.

In practice, we need to bound the maximum mesh scales and maximum ratio of adjacent meshes as a prior assumption. This hypothesis is reasonable and intuitive, for which the sufficiently uniform and sufficiently fine mesh implies the discrete problem becomes closer to the continuous one, and at the same time, the DCC will close to its continuous counterpart.

Lemma 2 (The asymptotic analysis of DCC) *For all $1 \leq k \leq n$, the asymptotic expression $\tilde{p}_{n-k}^{(n)} = \int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - s)\mathrm{d}s$ of DCC yields that*

$$p_{n-k}^{(n)} \leq \tilde{p}_{n-k}^{(n)}, \tag{6}$$

Proof From the construction of DCC kernels which want to illustrate the discontinuous semi-group property and the semi-group property on each right part of the following relations, we know that

$$\sum_{j=k}^{n} p_{n-j}^{(n)} a_{j-k}^{(j)} = 1 = \sum_{j=k}^{n} \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\omega_{1-\alpha}(s - t_{k-1})\mathrm{d}s, \tag{7a}$$

$$\sum_{j=k+1}^{n} p_{n-j}^{(n)} a_{j-k-1}^{(j)} = 1 = \sum_{j=k+1}^{n} \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\omega_{1-\alpha}(s - t_k)\mathrm{d}s. \tag{7b}$$

Subtract equation (7a) over (7b), then for all $k < n$ has

$$0 = p_{n-k}^{(n)} a_0^{(k)} - \sum_{j=k+1}^{n} p_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right) \tag{8a}$$

$$= \int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t) \omega_{1-\alpha}(t - t_{k-1}) \mathrm{d}t$$

$$- \sum_{j=k+1}^{n} \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \left[\omega_{1-\alpha}(t - t_k) - \omega_{1-\alpha}(t - t_{k-1}) \right] \mathrm{d}t \tag{8b}$$

$$= \left(q_{n-k}^{(n)} \tilde{p}_{n-k}^{(n)} \right) a_0^{(k)} - \sum_{j=k+1}^{n} \left(q_{n-j}^{(n)} \tilde{p}_{n-j}^{(n)} \right) \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right), \tag{8c}$$

where

$$q_{n-k}^{(n)} := \begin{cases} 1/\tilde{p}_0^{(n)} a_0^{(n)}, & \text{if } k = n, \\ \displaystyle\sum_{j=k+1}^{n} \dfrac{\left(q_{n-j}^{(n)} \tilde{p}_{n-j}^{(n)} \right) \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)} \right)}{\tilde{p}_{n-k}^{(n)} a_0^{(k)}}, & \text{if } k < n. \end{cases}$$

Each of the above three relations has its individual effect:

- (8a) comes from the process of constructing DCC and actually has similar semi-group properties in discrete cases
- (8b) is the significant process of finding the continuous corresponding one of DCC. The analysis here requires a bit of intuition to observe the properties of continuous semi-groups and discrete analogs, which have strong similarities in their constituent forms In the subsequent proof, what we need to do is to quantify the similarity.
- (8c) can exist independently of the previous two relations, and its purpose to characterize the ratios between $p_{n-k}^{(n)}$ and $\tilde{p}_{n-k}^{(n)}$ via unknown parameters $q_{n-k}^{(n)}$ for all $1 \le k \le n$. From the matrix form (5), it can be seen that because the monotonicity of $a_{n-k}^{(n)}$ indicates that the $\mathbf{A}_n \mathbf{D}_n$ matrix is an M-matrix, then it can be determined that $\mathbf{P}_n$ is a positive matrix. Obviously, since $\tilde{p}_{n-k}^{(n)}$ is positive through its specific expressions, and then $q_{n-k}^{(n)}$ is also a positive number. The definition of this ratio may be similar to the definition of DCC, but there are differences in specific expressions. The process is to define the value of k through the term k to n. By repeating this, all the values of $q_{n-k}^{(n)}$ can be defined.

Then the matrix form of equations (8a) and (8b) is

$$\mathbf{I}_n = \mathbf{P}_n \mathbf{A}_n \mathbf{D}_n = \left(\mathbf{Q}_n \circ \widetilde{\mathbf{P}}_n \right) \mathbf{A}_n \mathbf{D}_n,$$

where $\mathbf{Q}_n \circ \widetilde{\mathbf{P}}_n$ is the Hadamard product of matrices and

$$\mathbf{Q}_n := \begin{bmatrix} q_0^{(1)} & 0 & \cdots & 0 \\ q_1^{(2)} & q_0^{(2)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ q_{n-1}^{(n)} & q_{n-2}^{(n)} & \cdots & q_0^{(n)} \end{bmatrix}.$$

Obviously, the inverse of matrices $\mathbf{A}_n \mathbf{D}_n$ exists and $(\mathbf{A}_n \mathbf{D}_n)^{-1} = \mathbf{P}_n$, thus $0 < \mathbf{P}_n = \mathbf{Q}_n \circ \widetilde{\mathbf{P}}_n$, where the M-matrices property of $\mathbf{A}_n \mathbf{D}_n$ guarantees that $\mathbf{P}_n$ is a positive matrix. The rest of this proof will discuss the uniform upper bound of $q_{n-k}^{(n)}$ from its definition,

$$
\begin{aligned}
q_{n-k}^{(n)} &= \sum_{j=k+1}^{n} \frac{\left(q_{n-j}^{(n)} \tilde{p}_{n-j}^{(n)}\right)\left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)}\right)}{\tilde{p}_{n-k}^{(n)} a_0^{(k)}} \\
&\leq \sum_{j=k+1}^{n} \frac{\left(q_{n-j}^{(n)} \tilde{p}_{n-j}^{(n)}\right)\left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)}\right)}{\int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t)\omega_{1-\alpha}(t - t_{k-1})\mathrm{d}t} \\
&\leq C_{r,\tau} \sum_{j=k+1}^{n} q_{n-j}^{(n)} \frac{\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t)\left[\omega_{1-\alpha}(t - t_k) - \omega_{1-\alpha}(t - t_{k-1})\right]\mathrm{d}t}{\int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t)\omega_{1-\alpha}(t - t_{k-1})\mathrm{d}t},
\end{aligned}
$$

where the first inequality from the Chebyshev sorting inequality

$$
\begin{aligned}
&\int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t)\omega_{1-\alpha}(t - t_{k-1})\mathrm{d}t \\
&\leq \int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t)\mathrm{d}t \cdot \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} \omega_{1-\alpha}(t - t_{k-1})\mathrm{d}t = \tilde{p}_{n-k}^{(n)} a_0^{(k)}.
\end{aligned}
$$

and the second one from the maximum of non-negative numbers

$$
C_{r,\tau} := \max_{k+1 \leq j \leq n} \left\{ \frac{\tilde{p}_{n-j}^{(n)}\left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)}\right)}{-\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t) \int_{t_{k-1}}^{t_k} \omega_{1-\alpha}'(t - s)\mathrm{d}s\mathrm{d}t} \right\}.
$$

(8b) implies

$$
\sum_{j=k+1}^{n} \frac{\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - t)\left[\omega_{1-\alpha}(t - t_k) - \omega_{1-\alpha}(t - t_{k-1})\right]\mathrm{d}t}{\int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - t)\omega_{1-\alpha}(t - t_{k-1})\mathrm{d}t} \equiv 1,
$$

and then the discrete Grönwall inequality [4] with its $a_n = \varepsilon \to 0$ with loss of generality, shows that

$$
q_{n-k}^{(n)} \leq \varepsilon(1 + C_{r,\tau})^{n-k} \leq 1,
$$

where the estimation of $C_{r,\tau}$ has been analyzed in lemma 1.

Remark 1 When $r_{n+1} \equiv 1$ and $\tau \to 0$ for all $1 \leq n \leq N - 1$, then $q_{n-j,k}^{(n)} \to 1$ for all $1 \leq k + 1 \leq j \leq n \leq N$. Meanwhile,

$$
q_{n-k,k}^{(n)} \tilde{p}_{n-k}^{(n)} a_0^{(k)} - \sum_{j=k+1}^{n} q_{n-j,k}^{(n)} \tilde{p}_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)}\right)
$$

$$
\xrightarrow{\tau \to 0} \tilde{p}_{n-k}^{(n)} a_0^{(k)} - \sum_{j=k+1}^{n} \tilde{p}_{n-j}^{(n)} \left(a_{j-k-1}^{(j)} - a_{j-k}^{(j)}\right),
$$

namely $p_{n-k}^{(n)} \to \tilde{p}_{n-k}^{(n)}$ from entries-wise congruence of matrix. From the point view, when $q_{n-j,k}^{(n)} \to 1$, the $q_{n-k}^{(n)}$ bounded by the convex combination of term from $k + 1$ to n, and then the above inequality can be optimized.
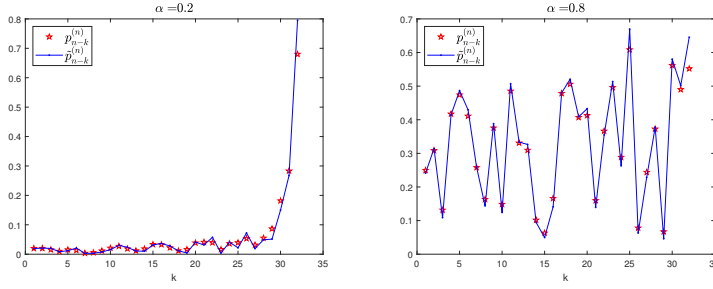
Fig. 1 The diagram of the $p_{n-k}^{(n)}$ and $\tilde{p}_{n-k}^{(n)}$ with fixed $n = 32$

Remark 2 Unlike usual habits, the construction of DCC defines k-th term from the larger part of indicator n to $k + 1$. The construction process of DCC tells us that the right-side part of each sub-figure in the figure 1 is the initial defined part, and the approximation of asymptotic estimation is not very significant in this part. The left-side parts, since the more terms are associated, require asymptotic estimation the most.

In fact, the mesh ratio r illustrates the degree of mesh smoothness. If time scales τ_k satisfies some proper smooth function with respect to k, e.g. $\tau_k = 0.4 \sin(\frac{3k\pi}{n}) + 0.41$, the graph of $p_{n-k}^{(n)}$ was basically identical with the $\tilde{p}_{n-k}^{(n)}$ in the left sub-figure of figure 2.
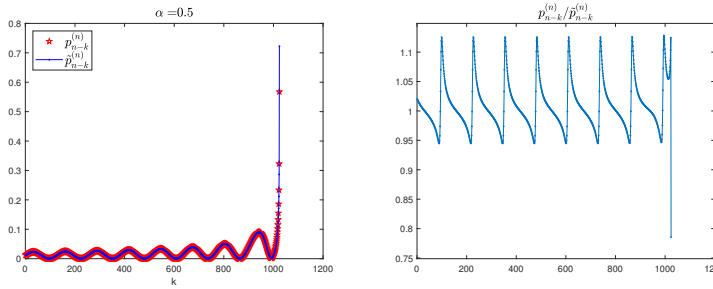

Fig. 2 The graphs of $p_{n-k}^{(n)}$ and $\tilde{p}_{n-k}^{(n)}$ with fixed $n = 1024$.

2.2 The Asymptotically compatible fractional Grönwall inequality

Theorem 1 *For arbitrary positive constant κ, the non-negative sequence $\{V_k\}_{k=0}^N$ and $\{F_k\}_{k=0}^N$,*

$$\mathrm{D}_\tau^\alpha V_n \leq \kappa V_n + F_n,$$

then

$$V_n \leq E_\alpha(\kappa t_n^\alpha) \left(V_0 + \max_{1 \leq n \leq N} \left\{ \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\mathrm{d}s \cdot F_j \right\} \right).$$

Proof Adjust the index n as j of original inequality and take $\Sigma_{j=1}^n p_{n-j}^{(n)}$ on both sides,

$$\sum_{j=1}^n p_{n-j}^{(n)} \sum_{k=1}^j a_{j-k}^{(j)} \nabla_\tau V_k \le \kappa \sum_{j=1}^n p_{n-j}^{(n)} V_j + \sum_{j=1}^n p_{n-j}^{(n)} F_j,$$

then

$$V_n \le V_0 + \kappa \sum_{j=1}^n p_{n-j}^{(j)} V_j + \sum_{j=1}^n p_{n-j}^{(n)} F_j \le V_0 + \kappa \sum_{j=1}^n \tilde{p}_{n-j}^{(j)} V_j + \sum_{j=1}^n \tilde{p}_{n-j}^{(n)} F_j.$$

The classical Grönwall inequality yields that

$$V_n \le E_\alpha(\kappa t_{n-1}^\alpha) \left(V_0 + \max_{1 \le \nu \le n} \left\{ \sum_{j=1}^\nu \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\mathrm{d}s \cdot F_j \right\} \right),$$

This completes the proof.

Corollary 1 *Comparing with classical Grönwall inequality that*

$$\int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\mathrm{d}s \xrightarrow{\alpha \to 1} \tau_j,$$

the theorem 1 gives an asymptotically compatible result that

$$\begin{cases} \text{If } \mathrm{D}_\tau^\alpha V_n \le \kappa V_n + F_n, \text{ then} \\ V_n \le E_\alpha(\kappa t_{n-1}^\alpha) \left(V_0 + \max_{1 \le \nu \le n} \left\{ \sum_{j=1}^\nu \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\mathrm{d}s \cdot F_j \right\} \right), \\ \\ \xrightarrow{\alpha \to 1} \begin{cases} \text{If } \delta_\tau V_n \le \kappa V_n + F_n, \text{ then} \\ V_n \le \exp(\kappa t_{n-1}) \left(V_0 + \sum_{j=1}^n \tau_j F_j \right). \end{cases} \end{cases}$$

3 Error analysis

In this section, we will provide applications of Grönwall inequalities in numerical analysis. Since FEMs, Galerkin SMs, and energy estimation for continuous problems, they are all analyzed under the L^2-norm. Therefore, in this section, we mainly consider estimation and analysis under the L^2-norm or discrete L^2-norm.

Lemma 3 (L^2**-stability**) *The temporal semi-discretization* (3) *have*

$$\|u^n\| \le E_\alpha(\kappa C t_{n-1}^\alpha) \left(\|u^0\| + C \max_{1 \le \nu \le n} \left\{ \sum_{j=1}^\nu \int_{t_{j-1}}^{t_j} \omega_\alpha(t_n - s)\mathrm{d}s \cdot \|f(\mathbf{x}, t_j)\| \right\} \right).$$

Proof Directly use theorem 1 and the monotonicity of discrete convolution kernel, this proof can be done.

Denote $u_h(t_n)$ the spacial semi-discretization solution at time t_n for all $n \in \{1, 2, \dots, N\}$, and the truncation error derived from temporal discrete $|D_\tau^\alpha u_h^n - \partial_t^\alpha u_h(t_n)| = |\sum_{k=1}^n T_{n-k}^{(n)}| =: |\mathtt{T}_\tau^n|$ with

$$T_{n-k}^{(n)} := \int_{t_{k-1}}^{t_k} \omega_{1-\alpha}(t_n - s) \left[\frac{\nabla_\tau u_m^k}{\tau_k} - \frac{\partial u}{\partial s}(x_m, s) \right] ds.$$

Lemma 4 (Lemma 3.1 [18]) *Under the assumption $\|\partial_t^m u\| \lesssim 1 + t^{\beta - m}$, for $m = 1, 2$, and $\beta \in (0, 1) \cup (1, 2]$, the truncated error on general non-uniform meshes gives*

$$|\mathtt{T}_\tau^n| \le a_0^{(n)} G^n + \sum_{k=1}^{n-1} \left(a_{n-k-1}^{(n)} - a_{n-k}^{(n)} \right) G^k, \quad n \ge 1. \tag{9}$$

where $G^k := 2 \int_{t_{k-1}}^{t_k} (t - t_{k-1}) |u''(t)| \, dt$. Since the positivity of $p_{n-j}^{(n)}$, then

$$\sum_{j=1}^n p_{n-j}^{(n)} \left| \mathtt{T}_\tau^j \right| \le 2 \sum_{k=1}^n p_{n-k}^{(n)} a_0^{(k)} \int_{t_{k-1}}^{t_k} (t - t_{k-1}) |u_{tt}| \, dt,$$

which is asymptotically compatible with the global truncation error when $\alpha \to 1^-$.

Remark 3 If we want to obtain the pointwise-in-time error, the right-hand side can not include max operator in theorem 1. The main reason is that the classical Grönwall inequality requires $\sum_{j=1}^n \tilde{p}_{n-j}^{(n)} F_j$ should be non-decreasing sequence with respect to n. The generic F_j may not satisfy this condition. However, graded mesh or some special mesh establishes its non-decreasing property with the explicit bound of truncated error.

Theorem 2 (pointwise-in-time error) *On graded or quasi-graded meshes, which denote the discrete-time nodes as $t_n \sim T(n/N)^r$ and its time steps fulfill $t_n \sim \tau n^{r-1}$ with the smallest scale $\tau = T/N^r$, if*

$$D_\tau^\alpha \mathcal{E}^n \le \kappa \mathcal{E}^n + |\mathtt{T}_h^n| + |\mathtt{T}_\tau^n|, n = 1, 2, \cdots, N,$$

where $\mathcal{E}^n$ is the L^2-norm error on the n-th time level and $\mathtt{T}_h^n, \mathtt{T}_\tau^n$ are corresponding spacial and temporal truncated errors respectively, then

$$\mathcal{E}^n \lesssim \begin{cases} \tau^{(2-\alpha)/r} t_n^{\beta - (2-\alpha)/r}, & r > \frac{2-\alpha}{1+\beta-\alpha}, \\ \tau^{1+\beta-\alpha} t_n^{\alpha-1}(1 + \ln(n)), & r = \frac{2-\alpha}{1+\beta-\alpha}, \\ \tau^{1+\beta-\alpha} t_n^{\alpha-1}, & r < \frac{2-\alpha}{1+\beta-\alpha}, \end{cases} \tag{10}$$

is valid under the hypothesis $\|\partial_t^m u\| \lesssim t^{\beta - m}$ for $m = 1, 2$ and $\beta \in (0, 1) \cup (1, 2]$.

Proof By the result of lemmas 5 and 4, we can directly prove that

$$2\sum_{k=1}^{n} p_{n-k}^{(n)} a_0^{(k)} \int_{t_{k-1}}^{t_k} (t - t_{k-1}) |u_{tt}| \, \mathrm{d}t$$

$$\lesssim 2\sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \omega_\alpha(t_n - s)\mathrm{d}s \cdot \omega_{2-\alpha}(\tau_k) \cdot \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} (t - t_{k-1}) t^{\beta-2} \mathrm{d}t$$

$$\leq \sum_{k=2}^{n} \omega_\alpha(t_n - t_k)\omega_{2-\alpha}(\tau_k)\tau_k^2 t_{k-1}^{\beta-2} + 2\omega_\alpha(t_n - t_1)\omega_{2-\alpha}(\tau_1)\frac{\tau_1^\beta}{\beta}$$

$$\lesssim \sum_{k=1}^{n} (n^r - k^r)^{\alpha-1} k^{(r-1)(3-\alpha)+r(\beta-2)}$$

$$= \begin{cases} \mathcal{O}(n^{r\beta-(2-\alpha)}), & (r-1)(3-\alpha) + r(\beta-2) > -1, \\ \mathcal{O}(n^{r\beta-(2-\alpha)}(1+\ln(n))), & (r-1)(3-\alpha) + r(\beta-2) = -1, \\ \mathcal{O}(n^{r(\alpha-1)}), & (r-1)(3-\alpha) + r(\beta-2) < -1, \end{cases}$$

$$\sim \begin{cases} \tau^{(2-\alpha)/r} t_n^{\beta-(2-\alpha)/r}, & r > \frac{2-\alpha}{1+\beta-\alpha}, \\ \tau^{1+\beta-\alpha} t_n^{\alpha-1}(1+\ln(n)), & r = \frac{2-\alpha}{1+\beta-\alpha}, \\ \tau^{1+\beta-\alpha} t_n^{\alpha-1}, & r < \frac{2-\alpha}{1+\beta-\alpha}. \end{cases}$$

When $n = 1$, it is clear that the second line equals $\mathcal{O}(\tau_1^\beta)$ by direct calculation which consists of the above conclusion.

Remark 4 The above poinwise-in-time error estimate for pure dissipative system [10] can be obtain more directly, meanwhile the Grönwal is not required.

$$\mathrm{D}_\tau^\alpha \mathcal{E}^n \leq |\mathrm{T}_h^n| + |\mathrm{T}_\tau^n|, n = 1, 2, \cdots, N,$$

$$\Rightarrow \quad \mathcal{E}^n \leq \sum_{j=1}^{n} p_{n-j}^{(j)} \left(|\mathrm{T}_h^n| + |\mathrm{T}_\tau^n|\right) \leq \sum_{j=1}^{n} \tilde{p}_{n-j}^{(j)} \left(|\mathrm{T}_h^n| + |\mathrm{T}_\tau^n|\right).$$

Consequently, the explicit expression of the right-hand side can be obtained similarly.

4 Numerical Verification

In this section, we will provide several typical numerical examples to verify our theoretical analysis results. Verify the convergence of the L1 scheme at different times under different graded grid parameters r with the given decay rate of the solution $\beta \in (0, 2]$.

4.1 The growing linear ODE

To avoid disturbance of spacial discrete error, we consider

$$\partial_t^\alpha u(t) = \kappa u + f(t), \text{ with } u(0) = 0, \tag{11}$$

N_τ	$r = 1$		$r = \frac{2-\alpha}{1+\beta-\alpha}$		$r = \frac{2-\alpha}{1+\beta-\alpha}+1$	
	$\|\cdot\|_\infty$	c.o.	$\|\cdot\|_\infty$	c.o.	$\|\cdot\|_\infty$	c.o.
64	1.335e-01	-	3.574e-02	-	1.021e-02	-
128	1.056e-01	0.338	2.350e-02	0.605	5.467e-03	0.901
256	8.429e-02	0.325	1.548e-02	0.602	2.929e-03	0.900
512	6.771e-02	0.316	1.021e-02	0.601	1.570e-03	0.900
Theo. c.o.	-	0.300	-	0.600	-	0.900

Table 1 The error and convergence order of beginning time level with $\alpha = 0.3, \beta = 0.6$ for (11)

N_τ	$r = 1$		$r = \frac{2-\alpha}{1+\beta-\alpha}$		$r = \frac{2-\alpha}{1+\beta-\alpha}+1$	
	$\|\cdot\|_\infty$	c.o.	$\|\cdot\|_\infty$	c.o.	$\|\cdot\|_\infty$	c.o.
64	1.653e-01	-	2.831e-02	-	1.546e-02	-
128	1.039e-01	0.670	1.203e-02	1.235	5.882e-03	1.394
256	6.484e-02	0.680	5.043e-03	1.254	2.233e-03	1.398
512	4.027e-02	0.687	2.092e-03	1.269	8.466e-04	1.399
Theo. c.o.	-	0.700	-	1.400^-	-	1.400

Table 2 The error and convergence order of ending time level with $\alpha = 0.3, \beta = 0.6$ for (11)

with exact solution $u(t) = \omega_{1+\beta}(t)$ and $f(t) = \omega_{1+\beta-\alpha}(t) - \kappa\omega_{1+\beta}(t)$ for testing the pointwise-in-time error. From the accurate-order tables 1 2, we can know that the factor $(1 + \ln(n))$ is little effect on the error order at the initial time when $r = \frac{2-\alpha}{1+\beta-\alpha}$, whereas the large n will lead to loss of convergence order.

4.2 The reaction-diffusion equation

The spatial discretization adopts center difference with discrete scale $h = 2^{-9}\pi$. Consider $\partial_t^\alpha u - \Delta u = \kappa u + f(\mathbf{x}, t)$ with $\kappa > 0$ and verify the convergence order by setting exact solution $u = \sin(x)\omega_{1+\beta}(t)$ with the corresponding source term $f = \sin(x)(\omega_{1+\beta-\alpha}(t) + (1-\kappa)\omega_{1+\beta}(t))$. For simplicity, let $\kappa = 1$, $\mathbf{x} \in [-\pi, \pi]$, the final time $T = 1$ and $\mathcal{E}^n(N_\tau) := \sqrt{h\Sigma_h(u_h(t_n) - U_h^n)^2}$.

N_τ	$\gamma = 1$		$\gamma = \frac{2-\alpha}{1+\beta-\alpha}$		$\gamma = \frac{2-\alpha}{1+\beta-\alpha}+1$	
	$\mathcal{E}^1(N_\tau)$	c.o.	$\mathcal{E}^1(N_\tau)$	c.o.	$\mathcal{E}^1(N_\tau)$	c.o.
64	2.192e-01	-	6.296e-02	-	3.374e-02	-
128	1.781e-01	0.300	4.154e-02	0.600	2.006e-02	0.750
256	1.446e-01	0.300	2.740e-02	0.600	1.193e-02	0.750
512	1.175e-01	0.300	1.808e-02	0.600	7.093e-03	0.750
Theo. c.o.	-	0.300	-	0.600	-	0.750

Table 3 The error and convergence order of beginning time level with $\alpha = 0.3, \beta = 0.6$.

N_τ	$r = 1$		$r = \frac{2-\alpha}{1+\beta-\alpha}$		$r = \frac{2-\alpha}{1+\beta-\alpha} + 1$	
	$\mathcal{E}^{N_\tau}(N_\tau)$	c.o.	$\mathcal{E}^{N_\tau}(N_\tau)$	c.o.	$\mathcal{E}^{N_\tau}(N_\tau)$	c.o.
64	4.528e-02	-	8.350e-03	-	6.046e-03	-
128	2.783e-02	0.702	3.480e-03	1.263	2.337e-03	1.371
256	1.711e-02	0.702	1.433e-03	1.280	8.934e-04	1.387
512	1.052e-02	0.701	5.816e-04	1.301	3.354e-04	1.413
Theo. c.o.	-	0.700	-	1.400^-	-	1.400

Table 4 The error and convergence order of ending time level with $\alpha = 0.3, \beta = 0.6$.

5 Conclusions

In this article, we answered an open question about discrete complementary convolution kernels in [28] and demonstrated the asymptotic estimation of DCC under general non-uniform grids using a relatively concise proof. By utilizing the continuous counterpart of DCC, we proved the asymptotically compatible Grönwall inequality and applied it to the numerical analysis of time fractional parabolic equations. The required estimation formula for DCS in a graded mesh provides rigorous theoretical proof and illustrates the pointwise error estimation in graded mesh.The strict proof of the necessity for the Caputo discrete coefficient to hold true for $\mathrm{D}_\tau^\alpha |u^n|^2 \leq (\mathrm{D}_\tau^\alpha u^n, u^n)$ is also given in this paper incidentally. For the stability analysis of higher-order Caputo BDF_k, a unified analysis framework will be provided in the subsequent work.

A The estimate of DCS

In this part, we need to propose a beneficial estimate for DCS. Since the definition of Caputo fractional derivatives is based on the convolution formula, the estimation of discrete convolution integrals has been mentioned in many kinds of literature [19,23,5]. However, the estimation methods they use are relatively complicated, and the results are not very concise [23,24]. Here, we want to obtain fairly precise and easily analyzed results. Compared with Jin et al [9, Lemma 3.11], our result can be regarded as its graded mesh version.

Lemma 5 (the estimate of DCS) *For arbitrary positive real number r, the discrete convolution summation $S_{r,p,q}^{(n)} := \sum\limits_{k=1}^{n-1} (n^r - k^r)^p k^q$ have an estimate*

$$S_{r,p,q}^{(n)} \sim \begin{cases} \mathcal{O}\left(n^{rp+q+1}\right), & \min\{p,q\} > -1, \\ \mathcal{O}\left(n^{rp+q+1}(1 + \ln(n))\right), & \min\{p,q\} = -1, \\ \mathcal{O}\left(n^{\max\{rp,(r-1)p+q\}}\right), & \min\{p,q\} < -1. \end{cases}$$

Proof To begin with the proof, we want to explain the main idea of analysis. If $S_{1,p,q}^{(n)}$ have equal and invariant convergence rate with $n^{C_{1,p,q}}$ when $n \to +\infty$, where the $C_{1,p,q}$ is a fixed constant, then

$$\lim_{n \to \infty} \frac{S_{1,p,q}^{(n)}}{n^{C_{1,p,q}}} \sim 1.$$

Thus, we should find the highest-order terms of $S_{1,p,q}^{(n)}$. Even though each term of $S_{1,p,q}^{(n)}$ has a different order, we can ensure the range of its. Moreover, the highest order and the lowest order should come from the first term and the last term, or inverse ones which is the feature of finite terms polynomials. Without loss of generality, we just need to consider the following four cases.

1. When $p > -1$ and $q > -1$

$$\lim_{n \to \infty} \frac{S_{1,p,q}^{(n)}}{n^{p+q+1}} = \lim_{n \to \infty} \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)^p \cdot \left(\frac{k}{n}\right)^q \cdot \frac{1}{n} = B(p+1, q+1) \sim 1.$$

Hence, $S_n^{(1,p,q)} \sim n^{p+q+1}$ while $p > -1$ and $q > -1$.

2. When $p > -1 = q$

$$\frac{S_{1,p,-1}^{(n)}}{n^p(1 + \ln n)} = \frac{1}{(1 + \ln n)} \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)^p \frac{1}{k}$$

$$= \frac{1}{(1 + \ln n)} \sum_{k=1}^{n-1} \left(1 - \frac{k}{n}\right)^{p+1} \left(\frac{1}{n-k} + \frac{1}{k}\right).$$

The $\frac{1}{n^{p+1}} < \left(1 - \frac{k}{n}\right)^{p+1} < 1$ with $p + 1 > 0$, and the equality $\sum_{k=1}^{n-1} \frac{n}{(n-k)k} = 2 \sum_{k=1}^{n-1} \frac{1}{k}$, we know that

$$\frac{1}{(1 + \ln n)} \frac{2\gamma}{n^{p+1}} \left(\frac{1}{\gamma}(\sum_{k=1}^{n} \frac{1}{k} - \ln n) - \frac{1}{\gamma n} + \ln n\right)$$

$$< \frac{1}{(1 + \ln n)} \frac{2}{n^{p+1}} \left((\sum_{k=1}^{n} \frac{1}{k} - \ln n) - \frac{1}{n} + \ln n\right)$$

$$= \frac{1}{(1 + \ln n)} \frac{2}{n^{p+1}} \sum_{k=1}^{n-1} \frac{1}{k} < \frac{S_n^{(p,-1)}}{n^p(1 + \ln n)} < \frac{2}{(1 + \ln n)} \sum_{k=1}^{n-1} \frac{1}{k}$$

$$= \frac{2}{(1 + \ln n)} \left((\sum_{k=1}^{n-1} \frac{1}{k} - \ln n) + \ln n\right)$$

$$< \frac{2}{(1 + \ln n)} \left(\frac{1}{\gamma}(\sum_{k=1}^{n} \frac{1}{k} - \ln n) + \ln n\right),$$

where $\sum_{k=1}^{n} \frac{1}{k} - \ln n > 0$ and $\lim_{n \to \infty}(\sum_{k=1}^{n} \frac{1}{k} - \ln n)$ converge to the Euler constant $\gamma \approx 0.57721 < 1$. Take the limit $n \to \infty$, then

$$0 < \lim_{n \to \infty} \frac{S_{1,p,-1}^{(n)}}{n^p(1 + \ln n)} < 2.$$

If $S_{1,p,-1}^{(n)} \nsim n^p(1 + \ln n)$, then the limitation converges to either 0 or ∞. For this reason, $S_{1,p,q}^{(n)} \sim n^p(1 + \ln n)$ while $p = -1$ or $q = -1$.

3. When $-1 > p \geq q$

$$\frac{S_{1,p,q}^{(n)}}{n^p} = \sum_{k=1}^{n-1} \frac{1}{n^p} k^p (n-k)^q = \sum_{k=1}^{n-1} \frac{1}{n^p} k^p (n-k)^p (n-k)^{q-p}$$

$$< \sum_{k=1}^{n-1} \left(\frac{n}{k(n-k)}\right)^{-p} = \sum_{k=1}^{n-1} \left(\frac{1}{k} + \frac{1}{n-k}\right)^{-p}$$

$$= \sum_{k=1}^{[\frac{n}{2}]-1} \left(\frac{1}{k} + \frac{1}{n-k}\right)^{-p} + \sum_{k=[\frac{n}{2}]}^{n-1} \left(\frac{1}{k} + \frac{1}{n-k}\right)^{-p}$$

$$< 2^{-p} \sum_{k=1}^{[\frac{n}{2}]-1} \left(\frac{1}{k}\right)^{-p} + 2^{-p} \sum_{k=[\frac{n}{2}]}^{n-1} \left(\frac{1}{n-k}\right)^{-p} \leq 2^{1-p} \sum_{k=1}^{[\frac{n}{2}]} \left(\frac{1}{k}\right)^{-p},$$

then

$$0 < \lim_{n \to \infty} \frac{S_{1,p,q}^{(n)}}{n^p} < 2^{1-p} \sum_{k=1}^{\infty} \left(\frac{1}{k}\right)^{-p}.$$

4. When $p > -1 > q$

$$\frac{S_{1,p,q}^{(n)}}{n^p} \leq \frac{\sum_{k=1}^{n-1} k^p (n-k)^q}{n^p} \leq \frac{\left(\sum_{k=1}^{n-1} k^{\mu p}\right)^{\frac{1}{\mu}} \cdot \left(\sum_{k=1}^{n-1} (n-k)^{\nu q}\right)^{\frac{1}{\nu}}}{n^p}$$

$$\lesssim \left(\sum_{k=1}^{n-1} \left(\frac{k}{n}\right)^{\mu p}\right)^{\frac{1}{\mu}} \cdot \left(\sum_{k=1}^{n-1} (n-k)^{\nu q}\right)^{\frac{1}{\nu}},$$

where $\frac{1}{\mu} + \frac{1}{\nu} = 1$ and the Hölder inequality have been used in second line. Let $\mu \to \infty$, namely corresponding $\nu \to 1$, since the classical limit formula and the rest part is convergent series with $q < -1$, thence

$$\frac{S_{1,p,q}^{(n)}}{n^p} \lesssim 1.$$

The first two cases can imitate the above proof. Now, we consider the last one. $(n^r - 1)^p \sim n^{rp}$ and $(n^r - (n-1)^r)^p (n-1)^q \sim n^{(r-1)p+q}$ are terms while $k = 1$ and $k = n-1$ respectively. To achieve the whole convergence rate, we just need to find the highest order of this finite sum about n. The range of this order about n is from n^{rp} to $n^{(r-1)p+q}$, so the convolution summation is the same order with $n^{\max\{rp,(r-1)p+q\}}$. So far, the whole proof is accomplished.

Here, the Matlab code can be tested that our estimate is optimal,

```
N = 16;
q = zeros(N,1);
for j = 1:N
    q(j) = computing(2^j, -1, 2,.3);
    % close to a constant
end
function q = computing(n,p,q,r)
a = sum((n^r - (1:n-1).^r).^p.*(1:n-1).^q);
if min(p,q) > -1
    b = n^(r*p+q+1);
elseif min(p,q) == -1
    b = n^(r*p+q+1)*(1+log(n));
elseif min(p,q) < -1
    b = n^(max(r*p,(r-1)*p+q));
end
q = a/b;
end
```

B A proof of necessity

In classical parabolic problems, $\frac{1}{2}\partial_t \|u\|^2 = (\partial_t u, u)$ play a significant role for L^2-stability analysis. In time-fractional Caputo cases, A. A. Alikhanov [2, Lemma 1] has proved an analog $\frac{1}{2}\partial_t^\alpha \|u^n\|^2 \leq (\partial_t^\alpha u^n, u^n)$. Considering discrete version with proper discretization, as we well

know, the DC kernels $a_{n-k}^{(n)}$ monotonically increasing with respect to k can lead to the following inequality in L^2-stability analysis

$$\frac{1}{2}\mathrm{D}_\tau^\alpha \|u^n\|^2 \le (\mathrm{D}_\tau^\alpha u^n, u^n) \tag{12}$$

which includes L1-scheme, L21$_\sigma$-scheme with additional condition [16] and so on.

In this paper, we will propose rigorous proof the monotonicity of DC kernels is not only a sufficient condition of relation (12) but also a necessary condition. However, some high-order schemes, e.g. L1$^+$ or Caputo BDF-k do not have a relation (12), which is the main difficulty of stability analysis. The discrete L^2-norm estimation is essentially a series of inequalities established by quadratic form. The proof in this section also starts from this perspective, establishing the relationship between the positive inertia index of the representation matrix of the quadratic form and the uniform monotonicity.

Lemma 6 *For arbitrary positive real sequence $\{d_i\}_{i=1}^n$, define the real symmetrical matrix value function $\mathbf{M}(\mathbf{d}_n) := \mathbf{L}_n\mathbf{d}_n + \mathbf{d}_n\mathbf{L}_n^{\mathrm{T}} - \mathbf{d}_n$, whose $\mathbf{d}_n = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ and $\mathbf{L}_n$ is the all-one unite lower triangular matrix, then the matrix representation of quadratic form $\mathbf{M}(\mathbf{d}_n)$ is positive definite if only if $\{d_i\}_{i=1}^n$ must be uniformly strictly monotonically increasing.*

Proof Denote characteristic polynomial $p(\lambda; \mathbf{d}_n) := |\mathbf{M}(\mathbf{d}_n) - \lambda\mathbf{I}_n|$ and check its expression

$$p(\lambda; \mathbf{d}_n) = \begin{vmatrix} d_1 - \lambda & d_1 & \ldots & d_1 \\ d_1 & d_2 - \lambda & \ldots & d_2 \\ \vdots & \vdots & \ddots & \vdots \\ d_1 & d_2 & \ldots & d_n - \lambda \end{vmatrix} \xrightarrow{\textbf{step 1}} \begin{vmatrix} 1 & 0 & \ldots & 0 \\ -d_1 & d_1 - \lambda & \ldots & d_1 \\ \vdots & \vdots & \ddots & \vdots \\ -d_n & d_1 & \ldots & d_n - \lambda \end{vmatrix}$$

$$\xrightarrow{\textbf{step 2}} \begin{vmatrix} 1 & 1 & \ldots & 1 \\ -\delta_0^1 & -\lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\delta_0^n & -\delta_1^n & \ldots & -\lambda \end{vmatrix} \xrightarrow{\textbf{step 3}} \begin{vmatrix} x_0^{(n)} & 0 & \ldots & 0 \\ -\delta_0^1 & -\lambda & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\delta_0^n & -\delta_1^n & \ldots & -\lambda \end{vmatrix} = (-\lambda)^n x_0^{(n)},$$

where $\delta_j^k := d_k - d_j$ for all $0 \le j < k \le n$ with letting $d_0 = 0$.

step 1 Add one row and one column to the determinant and ensure that the values of the determinant do not change, and mark the added rows and columns as the zeroth row and zeroth column for easy description in the next;

step 2 Add the zeroth column of the determinant to each subsequent column and denote $d_k = \delta_0^k$ for all $1 \le k \le n$ as a unified form;

step 3 Retain the first element of the zeroth row in the determinant and reduce the remaining 1 to 0 through the following process. Firstly, by multiplying the elements in the nth row by $1/\lambda$ and adding them to the zeroth row, the elements in the nth column of the zeroth row are reduced to zero. At the same time, the kth element of the zeroth row is reduced to $1 - \frac{1}{\lambda}\delta_k^n, \forall\, 0 \le k \le n$. This step is completed through a similar process for subsequent rows.

In order to provide a more detailed description of **step 3**, which is the process of changing the elements in the zeroth row of the determinant, we record $x_k^{(n)}$ as the kth element in the zeroth row, and the superscript represents the nth change, such as $x_k^{(0)} \equiv 1, \forall\, 0 \le k \le n$ indicates that each element in the zeroth row in the initial state is one. So, the entire process is described as follows,

$$\begin{aligned} &\text{For all } 0 \le k \le n - 0,\ x_k^{(0)} = 1; \\ &\text{For all } 0 \le k \le n - 1,\ x_k^{(1)} = x_k^{(0)} - \tfrac{1}{\lambda}\delta_k^n x_{n-0}^{(0)}; \\ &\qquad\qquad \vdots \qquad\qquad\qquad \vdots \\ &\text{For all } 0 \le k \le 0,\qquad x_k^{(n)} = x_k^{(n-1)} - \tfrac{1}{\lambda}\delta_k^1 x_1^{(n-1)}; \end{aligned}$$

In fact, each step in the above process contains this linear transformation, and the coincidence of multiple linear transformations can be represented by corresponding matrix multiplication. So this process is represented as a matrix form,

$$x_0^{(n)} = \left(\prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k] \right) \mathbf{e}_{n+1}, \tag{13}$$

where $[\mathbf{I}_k, \mathbf{v}_k]$ represent $k \times k$ identity matrix stitching a k-dimensional column vector $\mathbf{v}_k := -\frac{1}{\lambda} \left[\delta_0^k, \delta_1^k \ldots, \delta_{k-1}^k \right]^{\mathrm{T}}$, e.g.

$$[\mathbf{I}_k, \mathbf{v}_k] = \begin{bmatrix} 1 & 0 & \ldots & 0 & -\frac{1}{\lambda}\delta_0^k \\ 0 & 1 & \ldots & 0 & -\frac{1}{\lambda}\delta_1^k \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -\frac{1}{\lambda}\delta_{k-1}^k \end{bmatrix}_{k \times (k+1)} .$$

and $\mathbf{e}_{n+1}$ is $(n+1)$-dimensional all-one column vector.

Next, we will calculate the specific expression of $(-\lambda)^n \left(\prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k] \right) \mathbf{e}_{n+1}$. Denote $\gamma_{k-i}^k := -\frac{1}{\lambda}\delta_{k-i}^k$ for all $1 \leq i \leq k$ and observe its non-compact form

$$[\mathbf{I}_1, \mathbf{v}_1][\mathbf{I}_2, \mathbf{v}_2] \ldots [\mathbf{I}_n, \mathbf{v}_n]\mathbf{e}_{n+1}.$$

From right to left is a matrix representation of the change process of the elements in the zeroth row of the determinant in **step 3**, whose dimensions gradually decrease and eventually change to a one-dimensional number. From left to right will be our next calculation process, where the dimension of the row vector is gradually increased to $n+1$-dimension, but ultimately multiplying it with the column vector is equivalent to summing the $n+1$ components of the row vector, which also yields a one-dimensional number.

At this point, our problem transforms into determining each element in the row vector $\prod_{k=1}^{n}[\mathbf{I}_k, \mathbf{v}_k]$. Here, we will approach this computational problem from the perspective of combinatorial mathematics. To this end, we first explain the meaning of the following symbols in combinatorial mathematics

$$\gamma_{i_{j-1}}^{i_j} : \text{a single step from } i_{j-1} \text{ to } i_j,$$

$$\prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}} : \text{a multi-step from } i_0 \text{ to } i_k \text{ via } k \text{ steps}.$$

Particularly, $\gamma_0^0 = 1$ means the position has no movement.

Next, we use mathematical induction to prove that $\forall \, 1 \leq i \leq n+1$ has

$$\left\langle \prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k] \right\rangle_i = \sum_{k=1}^{n+1-i} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}}, \tag{14}$$

where $0 = i_0 < i_1 < i_2 < \ldots < i_k = n+1-i$ and $\langle \cdot \rangle_i$ represents the i-th component in the vector. In the above equation, the specific meaning of its abstract expression is the total sum of all possibilities which from $i_0 = 0$ departs, can only move in the direction of point $n+1-i$ each time, but does not limit the forward step size, and ultimately reaches point $n+1-i$.

When $n = 1$, $[1, \gamma_0^1] = [\mathbf{I}_1, \mathbf{v}_1]$, obviously holds. Assume (14) holds for n, and now, we will prove $n+1$ also holds.

$$\left(\prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k] \right) \cdot [\mathbf{I}_{n+1}, \mathbf{v}_{n+1}] = \left[\prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k], \left(\prod_{k=1}^{n} [\mathbf{I}_k, \mathbf{v}_k] \right) \mathbf{v}_{n+1} \right]$$

$$= \left[\prod_{k=1}^{n} [\mathbf{v}_k, \mathbf{I}_k], \sum_{i=1}^{n+1} \left(\sum_{k=1}^{n+1-i} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}} \right) \gamma_{n+1-i}^{n+1} \right] = \left[\prod_{k=1}^{n} [\mathbf{v}_k, \mathbf{I}_k], \sum_{k=1}^{n+1} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}} \right],$$

where the calculation of the last step can also start from the meaning of combination, for all $1 \leq i \leq n+1$,

$$\left(\sum_{k=1}^{n+1-i} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}} \right) \gamma_{n+1-i}^{n+1},$$

represent all possibilities from $i_0 = 0$ depart, reach point $n+1-i$, and continue moving forward to reach point $n+1$, thus consequently $\sum_{i=1}^{n+1}$ operation means that taking the sum of all the possibilities of starting from $i_0 = 0$ and reaching point $n+1$, which can be used to directly write the final result based on this meaning.

By the above analysis, we know that

$$x_0^{(n)} = \sum_{i=1}^{n+1} \sum_{k=1}^{n+1-i} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}}, \tag{15}$$

those value is represented as from $i_0 = 0$ depart and take $0 \leq k \leq n$ steps to reach all possible sums up to point n. Based on $\gamma_{k-i}^k := -\frac{1}{\lambda} \delta_{k-i}^k$ is the -1 degree of $-\lambda$, then $\prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}}$ is represented as the $-k$ degree of $-\lambda$. All $-k$ degrees in the expression (15) are represented as from $i_0 = 0$ depart, after k steps, and reach all possibilities up to point n, i.e

$$\sum_{1 \leq i_1 < ... < i_k \leq n} \prod_{j=0}^{k-1} \gamma_{i_j}^{i_{j+1}}.$$

Then

$$p(\lambda; \mathbf{d}_n) = (-\lambda)^n \sum_{k=1}^{n} \left(\sum_{1 \leq i_1 < ... < i_k \leq n} \prod_{j=1}^{k} \gamma_{i_{j-1}}^{i_j} \right) = \sum_{i=0}^{n} C_i (-\lambda)^i,$$

where C_i represents both the coefficient of the degree i of $-\lambda$ and from $i_0 = 0$ depart at, take $n-i$ steps, and reach all possibilities up to point n, namely C_i have specific expression

$$\begin{bmatrix} C_0 \\ \vdots \\ C_i \\ \vdots \\ C_{n-2} \\ C_{n-1} \\ C_n \end{bmatrix} = \begin{bmatrix} \prod_{k=1}^{n} \delta_{k-1}^k \\ \vdots \\ \sum_{1 \leq k_1 < ... < k_i \leq n} \prod_{j=1}^{i} \delta_{k_{j-1}}^{k_j} \\ \vdots \\ \sum_{1 \leq k_1 < k_2 \leq n} \delta_{k_1}^{k_2} \delta_0^{k_1} \\ \sum_{1 \leq k \leq n} \delta_0^k \\ 1 \end{bmatrix} \overset{\star}{=\!=} \begin{bmatrix} \prod_{k=1}^{n} \lambda_k \\ \vdots \\ \sum_{1 \leq k_1 < ... < k_i \leq n} \prod_{j=1}^{i} \lambda_{k_j} \\ \vdots \\ \sum_{1 \leq k_1 < k_2 \leq n} \lambda_{k_1} \lambda_{k_2} \\ \sum_{1 \leq k \leq n} \lambda_k \\ 1 \end{bmatrix}$$

According to all eigenvalues of a real symmetric matrix are real numbers and assume the kth real eigenvalue is λ_k, then its characteristic polynomial can be also represented as $C \prod_{k=1}^{n} (\lambda_k - \lambda) = p(\lambda; \mathbf{d}_n)$ where $C \neq 0$. In this problem, by comparing the coefficient of highest-order term $(-\lambda)^n$, it can be determined that $C = 1$. For more coefficient comparisons, we can refer to the identity $(\star)$.

Therefore, by comparing the coefficient of zeroth polynomial term, we can conclude that

$$\prod_{k=1}^{n} \lambda_k \equiv \prod_{k=1}^{n} (d_k - d_{k-1}), \text{ for all } 1 \leq k \leq n,$$

namely, for all $1 \leq k \leq n$ have $\lambda_k > 0 \Leftrightarrow d_k > d_{k-1}$. This completes the proof.

Theorem 3 *For all $1 \leq n \leq N$,*

$$\frac{1}{2}\sum_{k=1}^{n} a_{n-k}^{(n)} \nabla_{\tau} \left\| u^k \right\|^2 \leq \left(\sum_{k=1}^{n} a_{n-k}^{(n)} \nabla_{\tau} u^k, u^n \right),$$

if and only if the DC kernels $a_{n-k}^{(n)}$ is monotonically increasing with respect to index k.

Proof

$$\left(\sum_{i=1}^{n} a_{n-i}^{(n)} \nabla_{\tau} u^i, u^n \right) - \frac{1}{2} \sum_{i=1}^{n} a_{n-i}^{(n)} \nabla_{\tau} \left\| u^i \right\|^2 = \sum_{i=1}^{n} a_{n-i}^{(n)} \left(\nabla_{\tau} u^i, u^n - u^{i-\frac{1}{2}} \right)$$

$$= \sum_{i=1}^{n} a_{n-i}^{(n)} \left(\nabla_{\tau} u^i, u^n - u^{i-1} \right) + \sum_{i=1}^{n} a_{n-i}^{(n)} \left(\nabla_{\tau} u^i, u^{i-1} - u^{i-\frac{1}{2}} \right)$$

$$= \sum_{i=1}^{n} a_{n-i}^{(n)} \left(\nabla_{\tau} u^i, \sum_{j=i}^{n} \nabla_{\tau} u^j \right) - \frac{1}{2} \sum_{i=1}^{n} a_{n-i}^{(n)} \left\| \nabla_{\tau} u^i \right\|^2$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{j} a_{n-i}^{(n)} \nabla_{\tau} u^i, \nabla_{\tau} u^j \right) - \frac{1}{2} \sum_{i=1}^{n} a_{n-i}^{(n)} \left\| \nabla_{\tau} u^i \right\|^2$$

$$= \mathbf{v}_n^{\mathrm{T}} \left(\mathbf{L}_n \mathbf{A}_{\mathrm{diag}} \right) \mathbf{v}_n - \frac{1}{2} \mathbf{v}_n^{\mathrm{T}} \mathbf{A}_{\mathrm{diag}} \mathbf{v}_n = \frac{1}{2} \mathbf{v}_n^{\mathrm{T}} \left(\mathbf{L}_n \mathbf{A}_{\mathrm{diag}} + \mathbf{A}_{\mathrm{diag}} \mathbf{L}_n^{\mathrm{T}} - \mathbf{A}_{\mathrm{diag}} \right) \mathbf{v}_n,$$

where diagonal matrix $\mathbf{A}_{\mathrm{diag}} := \mathrm{diag}([a_{n-j}^{(n)}]_{j=1}^{n},)$ and column vector $\mathbf{v}_n := [\nabla_{\tau} u^j]_{j=1}^{n}$. According to the lemma 6, $\mathbf{M}(\mathbf{A}_{\mathrm{diag}})$ is a matrix representation of the positive definite quadratic form based on the monotonicity of $a_{n-k}^{(n)}$. Therefore, for arbitrary real vector $\mathbf{v}_n$, the sufficient and essential condition to ensure $\mathbf{v}_n^{\mathrm{T}} \mathbf{M}(\mathbf{A}_{\mathrm{diag}}) \mathbf{v}_n \geq 0$ is that $a_{n-k}^{(n)}$ should be monotonically increasing for all $1 \leq k \leq n$. The assertion have been finished.

Conflict of interest

The authors declared that they have no conflicts of interest to this work.

References

1. Al-Maskari, M., Karaa, S.: Numerical approximation of semilinear subdiffusion equations with nonsmooth initial data. SIAM Journal on Numerical Analysis **57**(3), 1524–1544 (2019). DOI 10.1137/18m1189750
2. Alikhanov, A.A.: A priori estimates for solutions of boundary value problems for fractional-order equations. Differential Equations **46**(5), 660–666 (2010). DOI 10.1134/s0012266110050058
3. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. Society for Industrial and Applied Mathematics (2002). DOI 10.1137/1.9780898719208. URL `https://epubs.siam.org/doi/abs/10.1137/1.9780898719208`
4. Clark, D.S.: Short proof of a discrete gronwall inequality. Discrete Applied Mathematics **16**(3), 279–281 (1987). DOI 10.1016/0166-218x(87)90064-3
5. Gracia, J.L., O'Riordan, E., Stynes, M.: Convergence in positive time for a finite difference method applied to a fractional convection-diffusion problem. Computational Methods in Applied Mathematics **18**(1), 33–42 (2017). DOI 10.1515/cmam-2017-0019
6. Jin, B., Lazarov, R., Zhou, Z.: Numerical methods for time-fractional evolution equations with nonsmooth data: A concise overview. Computer Methods in Applied Mechanics and Engineering **346**, 332–358 (2019). DOI 10.1016/j.cma.2018.12.011
7. Jin, B., Li, B., Zhou, Z.: Numerical analysis of nonlinear subdiffusion equations. SIAM Journal on Numerical Analysis **56**(1), 1–23 (2018). DOI 10.1137/16m1089320
8. Jin, B., Zhou, Z.: An analysis of galerkin proper orthogonal decomposition for subdiffusion **51**(1), 89–113 (2017). DOI 10.1051/m2an/2016017
9. Jin, B., Zhou, Z.: Numerical Treatment and Analysis of Time-Fractional Evolution Equations. Springer International Publishing (2023). DOI 10.1007/978-3-031-21050-1
10. Kopteva, N.: Error Analysis of the L1 Method on Graded and Uniform Meshes for a Fractional-Derivative Problem in Two and Three Dimensions. Mathematics of Computation **88**(319), 2135–2155 (2019). DOI 10.1090/mcom/3410
11. Kopteva, N.: Error analysis for time-fractional semilinear parabolic equations using upper and lower solutions. SIAM Journal on Numerical Analysis **58**(4), 2212–2234 (2020). DOI 10.1137/20m1313015
12. Kopteva, N.: Error analysis of an l2-type method on graded meshes for a fractional-order parabolic problem. Mathematics of Computation **90**(327), 19–40 (2020). DOI 10.1090/mcom/3552
13. Kopteva, N., Meng, X.: Error Analysis for a Fractional-Derivative Parabolic Problem on Quasi-Graded Meshes using Barrier Functions. SIAM Journal on Numerical Analysis **58**(2), 1217–1238 (2020). DOI 10.1137/19M1300686
14. Li, D., null, H.Q., Zhang, J.: Sharp pointwise-in-time error estimate of l1 scheme for nonlinear subdiffusion equations. Journal of Computational Mathematics **42**(3), 662–678 (2024). DOI 10.4208/jcm.2205-m2021-0316
15. Liao, H.L., Li, D.F., Zhang, J.W.: Sharp Error Estimate of the Nonuniform L1 Formula for Linear Reaction-Subdiffusion Equations. SIAM Journal on Numerical Analysis **56**(2), 1112–1133 (2018). DOI 10.1137/17m1131829
16. Liao, H.L., McLean, W., Zhang, J.: A second-order scheme with nonuniform time steps for a linear reaction-subdiffusion problem. Communications in Computational Physics **30**(2), 567–601 (2021). DOI 10.4208/cicp.oa-2020-0124
17. Liao, H.L., McLean, W., Zhang, J.W.: A Discrete Gronwall Inequality with Applications to Numerical Schemes for Subdiffusion Problems. SIAM Journal on Numerical Analysis **57**(1), 218–237 (2019). DOI 10.1137/16m1175742

18. Liao, H.l., Yan, Y.G., Zhang, J.W.: Unconditional Convergence of a Fast Two-Level Linearized Algorithm for Semilinear Subdiffusion Equations. Journal of Scientific Computing **80**(1), 1–25 (2019). DOI 10.1007/s10915-019-00927-0
19. Lubich, C.: Discretized Fractional Calculus. SIAM Journal on Mathematical Analysis **17**(3), 704–719 (1986). DOI 10.1137/0517050
20. Lubich, C.: Convolution quadrature and discretized operational calculus. I. Numerische Mathematik **52**(2), 129–145 (1988). DOI 10.1007/BF01398686
21. Lubich, C.: Convolution quadrature and discretized operational calculus. II. Numerische Mathematik **52**(4), 413–425 (1988). DOI 10.1007/BF01462237
22. Shen, J., Tang, T., Wang, L.L.: Spectral Methods. Springer Berlin Heidelberg (2011). DOI 10.1007/978-3-540-71041-7
23. Stynes, M., O'Riordan, E., Gracia, J.L.: Error analysis of a finite difference method on graded meshes for a time-fractional diffusion equation. SIAM Journal on Numerical Analysis **55**(2), 1057–1079 (2017). DOI 10.1137/16m1082329
24. Stynes, M., Wang, D.L.: A sharp discrete convolution sum estimate. Communications in Nonlinear Science and Numerical Simulation **117**, 106923 (2023). DOI 10.1016/j.cnsns.2022.106923
25. Thomas, J.W.: Numerical Partial Differential Equations: Finite Difference Methods. Springer New York (1995). DOI 10.1007/978-1-4899-7278-1
26. Thomée, V.: Galerkin Finite Element Methods for Parabolic Problems. Springer Berlin Heidelberg (2006). DOI 10.1007/3-540-33122-0
27. Trefethen, L.N.: Spectral Methods in MATLAB. Society for Industrial and Applied Mathematics (2000). DOI 10.1137/1.9780898719598. URL `https://epubs.siam.org/doi/abs/10.1137/1.9780898719598`
28. Yang, Y., Wang, J., Chen, Y., Liao, H.l.: Compatible l2 norm convergence of variable-step l1 scheme for the time-fractional mbe model with slope selection. Journal of Computational Physics **467**, 111467 (2022). DOI 10.1016/j.jcp.2022.111467
29. Zhang, H., Zeng, F., Jiang, X., Karniadakis, G.E.: Convergence analysis of the time-stepping numerical methods for time-fractional nonlinear subdiffusion equations. Fractional Calculus and Applied Analysis **25**(2), 453–487 (2022). DOI 10.1007/s13540-022-00022-6