



A second-order fast compact scheme with unequal time-steps for subdiffusion problems

Xin Li^{1,2} · Hong-lin Liao² · Luming Zhang²

Received: 17 August 2019 / Accepted: 11 March 2020 / Published online: 26 March 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

In consideration of the initial singularity of the solution, a temporally second-order fast compact difference scheme with unequal time-steps is presented and analyzed for simulating the subdiffusion problems in several spatial dimensions. On the basis of sum-of-exponentials technique, a fast Alikhanov formula is derived on general nonuniform meshes to approximate the Caputo's time derivative. Meanwhile, the spatial derivatives are approximated by the fourth-order compact difference operator, which can be implemented by a fast discrete sine transform via the FFT algorithm. So the proposed algorithm is computationally efficient with the computational cost about $O(MN \log M \log N)$ and the storage requirement $O(M \log N)$, where M and N are the total numbers of grids in space and time, respectively. With the aids of discrete fractional Grönwall inequality and global consistency analysis, the unconditional stability and sharp H^1 -norm error estimate reflecting the regularity of solution are established rigorously by the discrete energy approach. Three numerical experiments are included to confirm the sharpness of our analysis and the effectiveness of our fast algorithm.

Keywords Subdiffusion problems · Fast Alikhanov formula · Compact difference operator · Fractional Grönwall inequality · Sharp H^1 -norm error estimate

✉ Hong-lin Liao
liaohl@csrc.ac.cn; liaohl@nuaa.edu.cn

Xin Li
garryxin@163.com

Luming Zhang
zhanglm@nuaa.edu.cn

¹ Department of Mathematics, Anhui Science and Technology University, Bengbu, 233000, Anhui, China

² Department of Mathematics, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, Jiangsu, People's Republic of China

1 Introduction

In recent decades, fractional partial differential equations have attracted many attentions for its superiority in describing the phenomenon related to nonlocality and spatial heterogeneity. The time fractional differential equations are one of the important fractional models developed in various fields [1]. We consider the following reaction-subdiffusion problem

$$\partial_t^\alpha u - \Delta u = \kappa u + f(\mathbf{x}, t) \quad \text{for } \mathbf{x} \in \Omega \text{ and } 0 < t \leq T, \tag{1.1}$$

$$u(\mathbf{x}, t) = 0 \quad \text{for } \mathbf{x} \in \partial\Omega \text{ and } 0 < t \leq T, \tag{1.2}$$

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \bar{\Omega}, \tag{1.3}$$

where $\Omega \subset \mathbb{R}^d$ is a bounded interval ($d = 1$), rectangle ($d = 2$), or cube ($d = 3$), and $\partial\Omega$ denotes its boundary. Δ be the d -dimensional Laplacian operator and the reaction coefficient κ is a positive constant. In (1.1), $\partial_t^\alpha = {}^C_0\mathcal{D}_t^\alpha$ represents the Caputo fractional derivative of order α with respect to t , defined by

$$(\partial_t^\alpha v)(t) := \int_0^t \omega_{1-\alpha}(t-s)v'(s)ds, \quad 0 < \alpha < 1, \tag{1.4}$$

where the weakly singular kernel $\omega_{1-\alpha}(t-s)$ is defined by $\omega_\mu(t) := t^{\mu-1}/\Gamma(\mu)$ for $t > 0$.

Despite there are some theoretical methods to solve the fractional diffusion equations analytically, we may still have to explore numerical methods to obtain the solutions in general. In previous literatures, a large number of numerical methods are presented to approximate the Caputo derivative (1.4) availably. For instance, the well-known L1 formula [2, 3], which is constructed by using a piecewise linear interpolation on each subinterval and has the accuracy of order $2 - \alpha$. High-order approximations are constructed recently in [4–7] by using the piecewise quadratic polynomial interpolation. Most of discrete approximations for the Caputo derivative depend on the uniform time-steps and restrict the solutions to be sufficiently smooth. Nevertheless, studies in [8–12] show that the existence of the singular kernel makes the exact solution u always has an initial layer at $t = 0$, which is typically for fractional differential equations and has great influence on the convergence rate. In [10], the authors prove that the uniform L1 formula is accurate of orders $O(t_n^{\alpha-1}\tau)$ and $O(t_n^{-1}\tau)$ for smooth and nonsmooth initial data respectively, which are far from the desired goals. Recently, there appears several interesting works, including the specified mesh techniques [12–16] and the (generally) nonuniform mesh approaches [17–20], to fill this gap by resolving the initial regularity and restoring the optimal convergence rate.

However, due to the intrinsically nonlocal property and historical dependence of the fractional derivative, numerical applications of the aforementioned methods are always time-consuming, especially for the problems on two- or three-dimensional space domain. Therefore, it is imperative and necessary to develop efficient algorithms to reduce the huge storage and computational cost. Many efforts have been made in the literatures to develop fast time-stepping methods. Ke et al. [21] apply the fast Fourier transform for block triangular Toeplitz-like systems and propose a fast

direct method for solving the time-fractional PDEs. Another idea is suggested by Baffet and Hesthaven [22], where the compression is carried out in the Laplace domain by solving the equivalent ODE with some one-step A-stable scheme. Recently, Jiang et al. [23] use the sum-of-exponentials (SOEs) technique to approximate the kernel and reduce the computational complexity significantly. The overall computational cost $O(MN^2)$ and storage $O(MN)$ of the L1 formula are reduced to $O(MN \log N)$ and $O(M \log N)$, respectively, where M is the grid points in space and N is the total levels in time. Yan et al. [24] propose a second-order fast formula on uniform meshes to approximate the Caputo derivative. They also investigate the stability and convergence of the resulting second-order time-stepping scheme for the subdiffusion equations, corresponding to the case of $\kappa = 0$ of (1.1) by assuming that the solution is sufficiently smooth. Moreover, a fast L1 finite difference scheme on the graded mesh is reported in [25] for time-fractional diffusion equations by considering both the initial singularity and the long-time historical memory. Liao et al. [26] propose a two-level nonuniform L1 scheme for solving semilinear subdiffusion equations and prove that the fast linearized scheme is unconditionally stable on generally nonuniform meshes. To the best of our knowledge, no nonuniform fast algorithms for high-order numerical Caputo derivatives have been published yet.

The main contribution of this paper is to construct and analyze a computationally efficient difference algorithm on a general class of nonuniform time meshes for linear reaction-subdiffusion equations in multi-dimensions. More precisely, the nonuniform Alikhanov formula with the SOEs approximation is employed for the Caputo fractional derivative, and the fourth-order compact difference operator with a fast discrete sine transform (DST) is utilized for spatial discretization so that our algorithm is computationally efficient in both time direction and spatial dimensions. Consider the (possibly nonuniform) time levels $0 = t_0 < t_1 < \dots < t_{k-1} < t_k < \dots < t_N = T$ with the weighted time $t_{n-\theta} := \theta t_{n-1} + (1 - \theta)t_n$ for an offset parameter $\theta \in [0, 1)$. Denote the time-step sizes $\tau_k = t_k - t_{k-1}$ for $1 \leq k \leq N$ and the maximum time-step size $\tau := \max_{1 \leq k \leq N} \tau_k$. Let the step size ratios $\rho_k := \tau_k / \tau_{k+1}$ for $1 \leq k \leq N - 1$ and the maximum time-step ratio

$$\rho := \max_{1 \leq k \leq N-1} \rho_k.$$

Our accelerated discrete Caputo formula for the Caputo derivative (1.4) takes a form of

$$(\partial_{at}^\alpha v)^{n-\theta} := \sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau v^k, \tag{1.5}$$

where $\nabla_\tau v^k = v^k - v^{k-1}$ and the discrete convolution kernels $A_{n-k}^{(n)}$ are determined later. Let $\bar{\Omega}_h$ be the discrete spatial grid covered by $\bar{\Omega}$ with uniform spacings in each direction. For $0 \leq n \leq N$, denote the numerical solution $u_h^n \approx u(\mathbf{x}_h, t_n)$, and $u_h^{n-\theta} := \theta u_h^{n-1} + (1 - \theta)u_h^n \approx u(\mathbf{x}_h, t_{n-\theta})$ for $\mathbf{x}_h \in \bar{\Omega}_h$. We take $\theta = \alpha/2$ and have

the following θ -weighted compact difference scheme for the reaction-subdiffusion problem (1.1)–(1.3)

$$(\partial_{\alpha\tau}^\theta u_h)^{n-\theta} - \mathcal{D}_h u_h^{n-\theta} = \kappa u_h^{n-\theta} + f(\mathbf{x}_h, t_{n-\theta}) \quad \text{for } \mathbf{x}_h \in \Omega_h \text{ and } 1 \leq n \leq N, \quad (1.6)$$

$$u_h^n = 0 \quad \text{for } \mathbf{x}_h \in \partial\Omega_h \text{ and } 1 \leq n \leq N, \quad (1.7)$$

$$u_h^0 = u_0(\mathbf{x}_h) \quad \text{for } \mathbf{x}_h \in \bar{\Omega}_h. \quad (1.8)$$

Here, the compact difference operator \mathcal{D}_h will be defined in the next section, in which the numerical implementation of (1.6)–(1.8) is also described.

The stability and convergence analysis of (1.6)–(1.8) rely on the recently developed fractional Grönwall inequality in [18], which is applicable for any discrete fractional derivative having the discrete convolution form (1.5). For the completeness of presentation, we combine three previous results from [18, Lemma 2.2, Theorems 3.1, and 3.2] into the following lemma.

Lemma 1.1 *Assume that discrete convolution kernels $A_{n-k}^{(n)}$ satisfy the following two criteria:*

A1. *There is a constant $\pi_A > 0$ such that $A_{n-k}^{(n)} \geq \frac{1}{\pi_A} \int_{t_{k-1}}^{t_k} \frac{\omega_{1-\alpha}(t_n-s)}{\tau_k} ds$ for $1 \leq k \leq n \leq N$.*

A2. *The discrete kernels are monotone, i.e., $A_{n-k-1}^{(n)} - A_{n-k}^{(n)} \geq 0$ for $1 \leq k \leq n-1 \leq N-1$.*

Define also a sequence of complementary discrete convolution kernels $P_{n-k}^{(n)}$ by

$$P_0^{(n)} := \frac{1}{A_0^{(n)}}, \quad P_{n-j}^{(n)} := \frac{1}{A_0^{(j)}} \sum_{k=j+1}^n (A_{k-j-1}^{(k)} - A_{k-j}^{(k)}) P_{n-k}^{(n)} \quad \text{for } 1 \leq j \leq n-1. \quad (1.9)$$

Then the complementary kernels $P_{n-k}^{(n)} \geq 0$ are well-defined and fulfill

$$\sum_{j=k}^n P_{n-j}^{(n)} A_{j-k}^{(j)} = 1 \quad \text{for } 1 \leq k \leq n \leq N. \quad (1.10)$$

$$\sum_{j=1}^n P_{n-j}^{(n)} \omega_{1+(m-1)\alpha}(t_n) \leq \pi_A \omega_{1+m\alpha}(t_n) \quad \text{for } m = 0, 1 \text{ and } 1 \leq n \leq N. \quad (1.11)$$

Suppose that the offset parameter $0 \leq \nu < 1$, λ is a nonnegative constant independent of the time-steps and the maximum step size

$$\tau \leq 1/\sqrt[2]{2\Gamma(2-\alpha)\lambda\pi_A}.$$

If the nonnegative sequences $(v^k)_{k=0}^N$ and $(\eta^k)_{k=1}^N$ satisfy

$$\sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau v^k \leq \lambda v^{n-\nu} + \eta^n \quad \text{for } 1 \leq n \leq N, \quad (1.12)$$

or

$$\sum_{k=1}^n A_{n-k}^{(n)} \nabla_{\tau} (v^k)^2 \leq \lambda (v^{n-\nu})^2 + v^{n-\nu} \eta^n \quad \text{for } 1 \leq n \leq N, \tag{1.13}$$

then it holds that for $1 \leq n \leq N$,

$$\begin{aligned} v^n &\leq 2E_{\alpha} \left(2 \max\{1, \rho\} \lambda \pi_A t_n^{\alpha} \left(v^0 + \max_{1 \leq k \leq n} \sum_{j=1}^k P_{k-j}^{(k)} \eta^j \right) \right) \\ &\leq 2E_{\alpha} \left(2 \max\{1, \rho\} \lambda \pi_A t_n^{\alpha} \left(v^0 + \pi_A \Gamma(1 - \alpha) \max_{1 \leq k \leq n} \{t_k^{\alpha} \eta^k\} \right) \right), \end{aligned}$$

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

This fractional discrete Grönwall lemma suggests that these two criteria **A1-A2** should be verified to guarantee the stability of the fully discrete scheme (1.6)–(1.8). In Section 3, we will prove some properties of the discrete convolution kernels $A_{n-k}^{(n)}$ by assuming

M1. The maximum time-step ratio $\rho = 3/2$.

Then Lemma 1.1 leads to the unconditional stability, see Theorem 3.1. Note that, **M1** is a mild mesh condition, but it implies that a sudden, drastic reduction of time-steps should be avoided to ensure the stability.

The present analysis is also built on a new concept named global consistency analysis, developed recently by Liao et al. in [19], where a sharp L^2 -norm error estimate of the nonuniform Alikhanov scheme for the problem (1.1)–(1.3) is established systematically. To make the analysis extendable (such as for distributed-order subdiffusion problems), we assume that the continuous solution u satisfies the regularity assumptions

$$\|u\|_{H^4(\Omega)} \leq C_u, \quad \|u^{(\nu)}(t)\|_{H^2(\Omega)} \leq C_u(1 + t^{\sigma-\nu}), \tag{1.14}$$

where $\nu \in \{1, 2, 3\}$ and $0 < t \leq T$, $\sigma \in (0, 1) \cup (1, 2)$ is a regularity parameter and C_u is a positive constant. For instance [9, 11, 12], this assumption holds with $\sigma = \alpha$ for the original problem (1.1) if $f(\mathbf{x}, t) = 0$ and $u_0 \in H_0^1(\Omega) \cap H^2(\Omega)$. Invoking to the new global consistency error analysis of the accelerated Alikhanov formula presented in Lemma 4.2 and the fractional discrete Grönwall lemma described earlier, we establish sharp H^1 -norm error estimate by means of the time-space error splitting technique, see Theorem 4.1. This theorem implies that the convergence analysis is applicable to a general family of nonuniform time meshes satisfying **M1**. To resolve the initial singularity of solution and solve it efficiently, we choose the time-steps that satisfy the following condition:

AssG. There is a positive constant C_{γ} independent of k , such that $\tau_k \leq C_{\gamma} \tau \min\{1, t_k^{1-1/\gamma}\}$ for $1 \leq k \leq N$, with $t_k \leq C_{\gamma} t_{k-1}$ and $\tau_k/t_k \leq C_{\gamma} \tau_{k-1}/t_{k-1}$ for $2 \leq k \leq N$.

Here, the parameter $\gamma \geq 1$ controls the extent to which the time levels are concentrated near $t = 0$. If the mesh is quasi-uniform, then **AssG** holds with $\gamma = 1$.

As γ increases, the initial step sizes become smaller compared with the later ones. A simple example satisfying **AssG** is the graded mesh $t_k = T(k/N)^\gamma$. Under the assumption of **AssG**, Theorem 4.1 shows the error estimate of the time-stepping scheme (1.6)–(1.8) as

$$|u(t_n) - u_h^n|_1 \leq \frac{C_u}{\sigma(1 - \alpha)} \left(\tau^{\min\{\gamma\sigma, 2\}} + h^4 + \epsilon \right) \quad \text{for } 1 \leq n \leq N,$$

and the second-order accuracy in time is achieved if $\gamma \geq 2/\sigma$, see Corollary 4.1 and Remark 2. In the last section, three numerical examples are given to confirm the theoretical results and verify the accuracy and effectiveness of the fast scheme. Throughout the paper, any subscripted C , such as C_u , C_v , and C_Ω , denotes a general positive constant, which has different values in different circumstances, and always depends on the given data and the continuous solution but is independent of the time and spatial steps.

2 Accelerated Alikhanov formula and fully discrete scheme

In this section, based on the SOEs approximation, fast Alikhanov formula on generally nonuniform time meshes is considered at first. Then, we develop a fully discrete second-order compact scheme for the problem (1.1)–(1.3) combined with the compact difference operator. Furthermore, an efficient numerical implementation is given at the end of this section.

2.1 Fast Alikhanov formula on nonuniform grids

Denote $\Pi_{1,k}v$ and $\Pi_{2,k}v$ are the linear interpolation of a function v with respect to t_{k-1}, t_k and the quadratic interpolation of v with respect to t_{k-1}, t_k , and t_{k+1} , respectively. Recalling the definition $\nabla_\tau v^k$, it is easy to check that for $k \geq 1$,

$$\begin{aligned} (\Pi_{1,k}v)'(t) &= \frac{\nabla_\tau v^k}{\tau_k}, \\ (\Pi_{2,k}v)' &= \frac{\nabla_\tau v^k}{\tau_k} + \frac{2(t - t_{k-1/2})}{\tau_k(\tau_k + \tau_{k+1})}(\rho_k \nabla_\tau v^{k+1} - \nabla_\tau v^k). \end{aligned}$$

For simplicity of presentation, we always denote $\varpi_n(\xi) := -\omega_{2-\alpha}(t_{n-\theta} - \xi)$ for $0 \leq \xi \leq t_{n-\theta}$ and $\varpi_n'(\xi) = \omega_{1-\alpha}(t_{n-\theta} - \xi)$ for $0 \leq \xi < t_{n-\theta}$. The nonuniform Alikhanov approximation to the Caputo fractional derivative (1.4) is defined as [19]

$$\begin{aligned} (\partial_\tau^\alpha v)^{n-\theta} &:= \sum_{k=1}^{n-1} \int_{t_{k-1}}^{t_k} \varpi_n'(\xi)(\Pi_{2,k}v)'(\xi)d\xi + \int_{t_{n-1}}^{t_{n-\theta}} \varpi_n'(\xi)(\Pi_{1,n}v)'(\xi)d\xi \\ &= a_0^{(n)} \nabla_\tau v^n + \sum_{k=1}^{n-1} \left(a_{n-k}^{(n)} \nabla_\tau v^k + \rho_k b_{n-k}^{(n)} \nabla_\tau v^{k+1} - b_{n-k}^{(n)} \nabla_\tau v^k \right), \quad (2.1) \end{aligned}$$

where the discrete coefficients $a_{n-k}^{(n)}$ and $b_{n-k}^{(n)}$ are defined by

$$a_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{\min\{t_{n-\theta}, t_k\}} \varpi_n'(\xi) d\xi \quad \text{for } 1 \leq k \leq n, \tag{2.2}$$

$$b_{n-k}^{(n)} := \frac{2}{\tau_k(\tau_k + \tau_{k+1})} \int_{t_{k-1}}^{t_k} (\xi - t_{k-\frac{1}{2}}) \varpi_n'(\xi) d\xi \quad \text{for } 1 \leq k \leq n - 1. \tag{2.3}$$

Rearranging the terms in (2.1), we reformulate (2.1) in the following discrete convolution form

$$(\partial_\tau^\alpha v)^{n-\theta} = \sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau v^k, \tag{2.4}$$

where the discrete convolution kernels $A_{n-k}^{(n)}$ are defined as: $A_0^{(1)} := a_0^{(1)}$ if $n = 1$, and for $n \geq 2$,

$$A_{n-k}^{(n)} := \begin{cases} a_0^{(n)} + \rho_{n-1} b_1^{(n)} & \text{for } k = n, \\ a_{n-k}^{(n)} + \rho_{k-1} b_{n-k+1}^{(n)} - b_{n-k}^{(n)} & \text{for } 2 \leq k \leq n - 1, \\ a_{n-1}^{(n)} - b_{n-1}^{(n)} & \text{for } k = 1. \end{cases}$$

Due to the essentially nonlocal property of the fractional derivative, the formula mentioned above involves the solution information at all historical time-levels and it is time-consuming in practical calculations. Here, the SOEs technique is employed to deduce an accelerated Alikhanov formula to improve the efficiency. The key point of the SOEs approximation (see [23, Theorem 2.5] or [26, Lemma 2.2]) is as follows:

Lemma 2.1 *Given $\alpha \in (0, 1)$, an absolute tolerance error $\epsilon \ll 1$, a cut-off time $\Delta t > 0$, and a final time T , there exist a positive integer N_q , positive quadrature nodes s^ℓ , and positive weights ϑ^ℓ ($1 \leq \ell \leq N_q$) such that*

$$\left| \omega_{1-\alpha}(t) - \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell t} \right| \leq \epsilon \quad \forall t \in [\Delta t, T],$$

where the quadrature nodes number N_q satisfies

$$N_q = O\left(\log \frac{1}{\epsilon} \left(\log \log \frac{1}{\epsilon} + \log \frac{T}{\Delta t}\right) + \log \frac{1}{\Delta t} \left(\log \log \frac{1}{\epsilon} + \log \frac{1}{\Delta t}\right)\right).$$

In [24], the authors illustrate that $N_q = O(\log N)$ for $T \gg 1$ and $N_q = O(\log^2 N)$ for $T \approx 1$. Moreover, they also list a table to show the relationships between N_q and various parameters α , τ , and ϵ , see Table 1 in [24]. These results indicate the SOEs strategy efficiently reduces the computational complexity and storage requirement. By virtue of this lemma, we divide (2.1) into a sum of a local part (an integral over $[t_{n-1}, t_{n-\theta}]$) and a historical part (an integral over $[0, t_{n-1}]$). In the former part, linear interpolation is utilized to approximate $v'(\xi)$ ($\xi \in (t_{n-1}, t_{n-\theta})$) as before, and in the

Table 1 Temporal error of Example 1 for different α with $\sigma = 0.8, \gamma = 1$

N	$\alpha = 0.2$			$\alpha = 0.5$			$\alpha = 0.8$		
	Alikhanov	AccA	Order	Alikhanov	AccA	Order	Alikhanov	AccA	Order
80	2.73e-04	2.73e-04		8.32e-04	8.32e-04		1.70e-03	1.70e-03	
160	1.57e-04	1.57e-04	0.80	4.78e-04	4.78e-04	0.80	9.79e-04	9.79e-04	0.80
320	9.00e-05	9.00e-05	0.80	2.75e-04	2.75e-04	0.80	5.62e-04	5.62e-04	0.80
640	5.17e-05	5.17e-05	0.80	1.58e-04	1.58e-04	0.80	3.23e-04	3.23e-04	0.80
1280	2.97e-05	2.97e-05	0.80	9.06e-05	9.06e-05	0.80	1.85e-04	1.85e-04	0.80
min{ $\sigma, 3 - \alpha$ }			0.80				0.80		

latter one, SOEs technique is applied to the convolution kernel $\varpi'_n(\xi)$, i.e.,

$$\begin{aligned}
 (\partial_\tau^\alpha v)^{n-\theta} &\approx \int_{t_{n-1}}^{t_n-\theta} \varpi'_n(\xi) \frac{\nabla_\tau v^n}{\tau_n} d\xi + \int_0^{t_{n-1}} \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(t_{n-\theta}-\xi)} v'(\xi) d\xi \\
 &= a_0^{(n)} \nabla_\tau v^n + \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(1-\theta)\tau_n} \mathcal{H}^\ell(t_{n-1}) \quad \text{for } n \geq 1,
 \end{aligned}$$

where

$$\mathcal{H}^\ell(t_k) = \int_0^{t_k} e^{-s^\ell(t_k-\xi)} v'(\xi) d\xi \quad \text{with } \mathcal{H}^\ell(t_0) = 0 \quad \text{for } 1 \leq \ell \leq N_q.$$

To compute $\mathcal{H}^\ell(t_k)$ efficiently, we attempt to obtain the recurrence formulation. Applying the quadratic interpolation to approximate $v'(\xi)$ in each subinterval $[t_{k-1}, t_k]$ ($1 \leq k \leq n - 1$) yields

$$\begin{aligned}
 \mathcal{H}^\ell(t_k) &\approx e^{-s^\ell \tau_k} \mathcal{H}^\ell(t_{k-1}) + \int_{t_{k-1}}^{t_k} e^{-s^\ell(t_k-\xi)} (\Pi_{2,k} v)'(\xi) d\xi \\
 &= e^{-s^\ell \tau_k} \mathcal{H}^\ell(t_{k-1}) + a^{(k,\ell)} \nabla_\tau v^k + b^{(k,\ell)} (\rho_k \nabla_\tau v^{k+1} - \nabla_\tau v^k),
 \end{aligned}$$

where the coefficients are given by

$$\begin{aligned}
 a^{(k,\ell)} &= \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} e^{-s^\ell(t_k-\xi)} d\xi, \\
 b^{(k,\ell)} &= \frac{2}{\tau_k(\tau_k + \tau_{k+1})} \int_{t_{k-1}}^{t_k} (\xi - t_{k-\frac{1}{2}}) e^{-s^\ell(t_k-\xi)} d\xi.
 \end{aligned}$$

Note that these coefficients are calculated exactly based on the integral property of exponential functions. Overall, the accelerated nonuniform Alikhanov formula can

be formulated as

$$(\partial_{\tau}^{\alpha} v)^{n-\theta} := a_0^{(n)} \nabla_{\tau} v^n + \sum_{\ell=1}^{N_q} \vartheta^{\ell} e^{-s^{\ell}(1-\theta)\tau_n} \mathcal{H}^{\ell}(t_{n-1}) \quad \text{for } n \geq 1, \tag{2.5}$$

where $\mathcal{H}^{\ell}(t_k)$ satisfies $\mathcal{H}^{\ell}(t_0) = 0$ and the recurrence relation

$$\begin{aligned} \mathcal{H}^{\ell}(t_k) &= e^{-s^{\ell}\tau_k} \mathcal{H}^{\ell}(t_{k-1}) + a^{(k,\ell)} \nabla_{\tau} v^k \\ &+ b^{(k,\ell)} (\rho_k \nabla_{\tau} v^{k+1} - \nabla_{\tau} v^k) \quad \text{for } k \geq 1, 1 \leq \ell \leq N_q. \end{aligned}$$

2.2 Second-order compact scheme and implementation

We discretize the spatial domain $\Omega \in \mathbb{R}^d$ by equal spatial step sizes $h_k = (x_{k,R} - x_{k,L})/M_k$ for the positive integers M_k ($1 \leq k \leq d$) and denote $h := \max_{1 \leq k \leq d} h_k$ is the maximum spatial length. Let $x_{k,i_k} = x_{k,L} + i_k h_k$ for $0 \leq i_k \leq M_k$, $1 \leq k \leq d$, then the fully discrete grids in space can be defined as $\tilde{\Omega}_h := \{\mathbf{x}_h = (x_{1,i_1}, x_{2,i_2}, \dots, x_{d,i_d}) | 0 \leq i_k \leq M_k, 1 \leq k \leq d\}$. Set $\Omega_h = \tilde{\Omega}_h \cap \Omega$ and the boundary $\partial\Omega_h = \tilde{\Omega}_h \cap \partial\Omega$. Also, $M = \prod_{k=1}^d (M_k + 1)$ represents the total number of grid points in space. Denote an index vector $h = (i_1, i_2, \dots, i_d)$ and for any grid function v_h at the k th position ($1 \leq k \leq d$), we introduce the following difference operators:

$$\delta_k v_{i_k+\frac{1}{2}} := \frac{v_{i_k+1} - v_{i_k}}{h_k}, \quad \delta_k^2 v_{i_k} := \frac{v_{i_k+1} - 2v_{i_k} + v_{i_k-1}}{h_k^2}, \quad \mathcal{I}_k v_{i_k} := (1 + \frac{h_k^2}{12} \delta_k^2) v_{i_k},$$

and the compact difference operator is denoted as $\mathfrak{D}_k v_{i_k} := \frac{\delta_k^2}{\mathcal{I}_k} v_{i_k}$. As a result, the second-order and fourth-order spatial approximations of $\Delta v(\mathbf{x}_h)$ for $\mathbf{x}_h \in \Omega_h$ can be denoted separately as $\Delta_h v_h := \sum_{k=1}^d \delta_k^2 v_h$ and $\mathfrak{D}_h v_h := \sum_{k=1}^d \mathfrak{D}_k v_h$.

Let $\mathbb{V}_h = \{v = (v_h)_{\mathbf{x}_h \in \Omega_h} | v_h = 0 \text{ when } \mathbf{x}_h \in \partial\Omega_h\}$ be the space of grid functions. For any $v, w \in \mathbb{V}_h$, we define the discrete inner product $\langle v, w \rangle_h := (\prod_{k=1}^d h_k) \sum_{\mathbf{x}_h \in \Omega_h} v_h w_h$ and the corresponding discrete L^2 -norm $\|v\| := \sqrt{\langle v, v \rangle_h}$.

Also, the discrete H^1 semi-norm is defined as $|v|_1 := \sqrt{\sum_{k=1}^d \|\delta_k v_h\|^2}$. Based on the embedding theorem, there exist a positive constant C_{Ω} dependent on the domain Ω such that

$$\|v\| \leq C_{\Omega} |v|_1 \quad \text{for } v \in \mathbb{V}_h. \tag{2.6}$$

Combining with the fast Alikhanov formula derived earlier, one has a fully discrete second-order compact scheme (1.6)–(1.8) for the problem (1.1)–(1.3). To speed up the numerical computations, we introduce the fast discrete sine transform (DST) [27, 28] in spatial direction and bypass the direct calculations. Based on the DST and compact difference operator, for any grid function v_h at the k th position, fourth-order

spatial approximation can be drawn:

$$\begin{aligned} & \sum_{j_k=1}^{M_k-1} \widehat{v}_{j_k}'' \left[\sin\left(\frac{(i_k+1)j_k\pi}{M_k}\right) + 10 \sin\left(\frac{i_k j_k \pi}{M_k}\right) + \sin\left(\frac{(i_k-1)j_k\pi}{M_k}\right) \right] \\ \approx & \frac{12}{h_k^2} \sum_{j_k=1}^{M_k-1} \widehat{v}_{j_k} \left[\sin\left(\frac{(i_k+1)j_k\pi}{M_k}\right) - 2 \sin\left(\frac{i_k j_k \pi}{M_k}\right) + \sin\left(\frac{(i_k-1)j_k\pi}{M_k}\right) \right]. \end{aligned}$$

After simple calculations, for $1 \leq j_k \leq M_k - 1$ and $1 \leq k \leq d$, we have

$$\widehat{v}_{j_k}'' \approx \widehat{v}_{j_k} \frac{12}{h_k^2} \left(\cos\left(\frac{j_k\pi}{M_k}\right) - 1 \right) \left(\cos\left(\frac{j_k\pi}{M_k}\right) + 5 \right)^{-1} = \widehat{v}_{j_k} \widehat{h}^{(j_k, M_k)}.$$

Denote an index set $v = \{(j_1, j_2, \dots, j_d) | 1 \leq j_k \leq M_k - 1, 1 \leq k \leq d\}$. Resorting to the fast Alikhanov formula in (2.5) and the fast DST mentioned above, numerical scheme (1.6) can be reformulated in the following way for implementation:

$$\begin{aligned} \left(\mathbf{a}_0^{(n)} - (1-\theta)\kappa - (1-\theta) \sum_{k=1}^d \widehat{h}^{(j_k, M_k)} \right) \nabla_\tau \widehat{u}_v^n &= \left(\sum_{k=1}^d \widehat{h}^{(j_k, M_k)} + \kappa \right) \widehat{u}_v^{n-1} + \widehat{f}^{n-\theta} - \widehat{\mathcal{M}}^{n-1}, \\ \widehat{\mathcal{M}}^{n-1} &= \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(1-\theta)\tau_n} \widehat{\mathcal{H}}^\ell(t_{n-1}) \\ &= \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(1-\theta)\tau_n} \left[e^{-s^\ell\tau_n} \widehat{\mathcal{H}}^\ell(t_{n-2}) + (a^{(n-1, \ell)} - b^{(n-1, \ell)}) \nabla_\tau \widehat{u}_v^{n-1} + b^{(n-1, \ell)} \rho_{n-1} \nabla_\tau \widehat{u}_v^n \right], \end{aligned}$$

where $\widehat{\mathcal{H}}^\ell(t_0) = 0$. \widehat{u}_v^{n-1} and $\widehat{f}^{n-\theta}$ are obtained from u_h^{n-1} and $f^{n-\theta}$ via the DST, respectively. The current solution u_h^n is computed from \widehat{u}_v^n via the inverse DST.

The fast DST is essentially implemented via the fast Fourier transform (FFT). In [29], the authors illustrate that FFT algorithm can reduce the numerical cost from $O(M^2)$ to $O(M \log M)$. Therefore, compared with traditional algorithms, the presented one reduces the overall computational cost from $O(M^2 N^2)$ to $O(MN \log M \log N)$ and the overall storage from $O(MN)$ to $O(M \log N)$.

3 Properties of discrete convolution kernels and stability

To prepare for the subsequent analysis, in this section, we show that the discrete kernels $\mathbf{A}_{n-k}^{(n)}$ of the fast approximation $(\partial_{\alpha\tau}^\alpha v)^{n-\theta}$ satisfy **A1-A2** in Lemma 1.1. Then, the stability in the H^1 -norm can be verified rigorously via the fractional discrete Grönwall lemma.

3.1 Properties of discrete kernels

We eliminate the historical term $\mathcal{H}^\ell(t_{n-1})$ in (2.5) and reformulate the recursive relationship into the discrete convolution form. The recurrence relation in (2.5) leads

to

$$\mathcal{H}^\ell(t_k) = \sum_{j=1}^k e^{-s^\ell(t_k-t_j)} \left(a^{(j,\ell)} \nabla_\tau v^j + \rho_j b^{(j,\ell)} \nabla_\tau v^{j+1} - b^{(j,\ell)} \nabla_\tau v^j \right). \tag{3.1}$$

Replacing k by $n - 1$ and plugging (3.1) into (2.5), we use the definitions (2.2)–(2.3) to obtain an alternative formula

$$(\partial_{a\tau}^\alpha v)^{n-\theta} = a_0^{(n)} \nabla_\tau v^n + \sum_{k=1}^{n-1} \left(a_{n-k}^{(n)} \nabla_\tau v^k + \rho_k b_{n-k}^{(n)} \nabla_\tau v^{k+1} - b_{n-k}^{(n)} \nabla_\tau v^k \right), \tag{3.2}$$

where the positive coefficients $a_{n-k}^{(n)}$ and $b_{n-k}^{(n)}$ are defined by

$$a_0^{(n)} := a_0^{(n)} \quad \text{and} \quad a_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(t_n-\theta-\xi)} d\xi \quad \text{for } 1 \leq k \leq n - 1, \tag{3.3}$$

$$b_{n-k}^{(n)} := \frac{2}{\tau_k(\tau_k + \tau_{k+1})} \int_{t_{k-1}}^{t_k} \sum_{\ell=1}^{N_q} \vartheta^\ell e^{-s^\ell(t_n-\theta-\xi)} (\xi - t_{k-\frac{1}{2}}) d\xi \quad \text{for } 1 \leq k \leq n - 1. \tag{3.4}$$

Correspondingly, we rearrange (3.2) and obtain the following discrete convolution form

$$(\partial_{a\tau}^\alpha v)^{n-\theta} = \sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau v^k,$$

where the discrete convolution kernels $A_{n-k}^{(n)}$ are defined as: $A_0^{(1)} := a_0^{(1)}$ if $n = 1$, and for $n \geq 2$,

$$A_{n-k}^{(n)} = \begin{cases} a_0^{(n)} + \rho_{n-1} b_1^{(n)} & \text{for } k = n, \\ a_{n-k}^{(n)} + \rho_{k-1} b_{n-k+1}^{(n)} - b_{n-k}^{(n)} & \text{for } 2 \leq k \leq n - 1, \\ a_{n-1}^{(n)} - b_{n-1}^{(n)} & \text{for } k = 1. \end{cases}$$

Comparing the coefficients between (2.2)–(2.3) and (3.3)–(3.4), we see that the latter ones are obtained by replacing the SOEs approximations with the integral kernels in the formers. Then Lemma 2.1 gives

$$\left| a_{n-k}^{(n)} - a_{n-k}^{(n)} \right| \leq \epsilon \quad \text{for } 1 \leq k \leq n - 1, \tag{3.5}$$

and when **M1** ($\rho_k \leq \rho = \frac{3}{2}$) holds, for $1 \leq k \leq n - 1$,

$$\left| b_{n-k}^{(n)} - b_{n-k}^{(n)} \right| \leq \frac{2\epsilon}{\tau_k(\tau_k + \tau_{k+1})} \int_{t_{k-1}}^{t_k} \left| \xi - t_{k-\frac{1}{2}} \right| d\xi = \frac{\tau_k^2 \epsilon}{2\tau_k(\tau_k + \tau_{k+1})} = \frac{\rho_k \epsilon}{2(1 + \rho_k)} \leq \frac{3\epsilon}{10}. \tag{3.6}$$

Next, we show that discrete convolution kernels $A_{n-k}^{(n)}$ satisfy the assumptions **A1**–**A2**. The following bounds which can be proved by the integral mean-value theorem will be utilized in several places,

$$a_{n-k-1}^{(n)} > \omega_{1-\alpha}(t_{n-\theta} - t_k) > a_{n-k}^{(n)} \quad \text{for } 1 \leq k \leq n - 1. \tag{3.7}$$

Lemma 3.1 *Let M1 holds. Denote $\epsilon^* = \frac{\alpha}{2(1-\alpha)}\omega_{1-\alpha}(T)$ and $\epsilon^{**} = \frac{1}{26}\omega_{1-\alpha}(T)$. If the tolerance error $\epsilon \leq \min\{\epsilon^*, \epsilon^{**}\}$, then the discrete convolutional kernels $A_{n-k}^{(n)}$ satisfy **A1** with $\pi_A = 2$ for $1 \leq k \leq n$. Furthermore, $A_{n-k}^{(n)}$ has an upper bound with $A_{n-k}^{(n)} \leq \frac{2}{\tau_n}\omega_{2-\alpha}(\tau_n)$.*

Proof The definition of $a_{n-k}^{(n)}$ in (3.3) and the integral mean-value theorem yield

$$a_{n-k-1}^{(n)} > a_{n-k}^{(n)} \quad \text{for } 1 \leq k \leq n - 2,$$

where s^ℓ, ϑ^ℓ are positive in Lemma 2.1 and the monotonicity of exponential function was also used. For $k = n - 1$, the definition (2.2) and (3.7) yield

$$a_0^{(n)} - a_1^{(n)} > \frac{1}{\tau_n}\omega_{2-\alpha}(t_{n-\theta} - t_{n-1}) - \omega_{1-\alpha}(t_{n-\theta} - t_{n-1}) = \frac{\alpha - \theta}{1 - \alpha}(1 - \theta)^{-\alpha}\omega_{1-\alpha}(\tau_n).$$

Since $e^x > 1 + x$ for $x \in \mathbb{R}$ and $\ln(1 - \frac{x}{2}) < 0$ for $0 < x < 1$, one has

$$a_0^{(n)} - a_1^{(n)} > \frac{\alpha}{2(1 - \alpha)}e^{-\alpha \ln(1 - \frac{\alpha}{2})}\omega_{1-\alpha}(\tau_n) > \frac{\alpha}{2(1 - \alpha)}\omega_{1-\alpha}(\tau_n) \geq \epsilon^* \geq \epsilon.$$

Therefore, we obtain $a_0^{(n)} = a_0^{(n)} > a_1^{(n)} + \epsilon \geq a_1^{(n)}$ by using the definitions in (3.3) and the inequality (3.5). It means that $a_{n-k-1}^{(n)} > a_{n-k}^{(n)}$ for $1 \leq k \leq n - 1$.

In consideration of [19, Lemma 4.3] and $0 < \theta < \frac{1}{2}$, it is easy to check

$$b_{n-k}^{(n)} < \frac{\theta \tau_k}{2(t_{n-\theta} - t_k)} \frac{\rho_k}{1 + \rho_k} a_{n-k}^{(n)} \leq \frac{\theta \tau_k}{2(1 - \theta)\tau_{k+1}} \frac{\rho_k}{1 + \rho_k} a_{n-k}^{(n)} < \frac{\rho_k^2}{2(1 + \rho_k)} a_{n-k}^{(n)}. \tag{3.8}$$

For $1 \leq k \leq n - 1$, the definition of $A_{n-k}^{(n)}$ and the inequalities (3.5)–(3.6) yield

$$A_{n-k}^{(n)} \geq a_{n-k}^{(n)} - b_{n-k}^{(n)} > a_{n-k}^{(n)} - \epsilon - \frac{\rho_k^2}{2(1 + \rho_k)} a_{n-k}^{(n)} - \frac{3\epsilon}{10} = \frac{2 + 2\rho_k - \rho_k^2}{2(1 + \rho_k)} a_{n-k}^{(n)} - \frac{13\epsilon}{10}.$$

Note that $\frac{2+2x-x^2}{2(1+x)}$ is increasing for $x \in (0, \frac{3}{2}]$, and $\epsilon \leq \epsilon^{**} \leq \frac{1}{26}\omega_{1-\alpha}(t_{n-\theta} - t_{k-1}) < \frac{1}{26}a_{n-k}^{(n)}$, then we find

$$A_{n-k}^{(n)} \geq \frac{11}{20}a_{n-k}^{(n)} - \frac{1}{20}a_{n-k}^{(n)} = \frac{1}{2}a_{n-k}^{(n)} > \frac{1}{2} \int_{t_{k-1}}^{t_k} \frac{\omega_{1-\alpha}(t_n - \xi)}{\tau_k} d\xi \quad \text{for } 1 \leq k \leq n - 1.$$

For the case of $k = n$, it is clear that $A_0^{(n)} \geq a_0^{(n)} = a_0^{(n)} > \frac{1}{2} \int_{t_{k-1}}^{t_k} \frac{\omega_{1-\alpha}(t_n - \xi)}{\tau_k} d\xi$. In summary, by choosing $\pi_A = 2$, the discrete convolution kernels $A_{n-k}^{(n)}$ satisfy **A1** for $1 \leq k \leq n$.

Now we derive an upper bound of $A_{n-k}^{(n)}$. From (3.6) and (3.8), one gets

$$A_0^{(n)} = a_0^{(n)} + \rho_{n-1}b_1^{(n)} < a_0^{(n)} + \frac{\rho_{n-1}^3}{2(1 + \rho_{n-1})} a_1^n + \frac{3\epsilon}{10}\rho_{n-1}.$$

Note that $\frac{x^3}{2(1+x)}$ is increasing for $x \in (0, \frac{3}{2}]$, we apply (3.7) and **M1** to get

$$A_0^{(n)} < a_0^{(n)} + \frac{27}{40}a_1^{(n)} + \frac{9\epsilon}{20} < a_0^{(n)} + \frac{27}{40}a_0^{(n)} + \frac{9}{20} \times \frac{1}{26}a_0^{(n)} = \frac{22}{13}a_0^{(n)} < 2a_0^{(n)}.$$

Since $a_0^{(n)} = \frac{1}{\tau_n}(1-\theta)^{1-\alpha} \int_{t_{n-1}}^{t_n} \omega_{1-\alpha}(t_n - \xi) d\xi \leq \frac{1}{\tau_n} \int_{t_{n-1}}^{t_n} \omega_{1-\alpha}(t_n - \xi) d\xi$, we obtain an upper bound, i.e., $A_{n-k}^{(n)} \leq \frac{2}{\tau_n} \int_{t_{n-1}}^{t_n} \omega_{1-\alpha}(t_n - \xi) d\xi = \frac{2}{\tau_n} \omega_{2-\alpha}(\tau_n)$. This completes the proof. \square

Next, we establish the monotonicity of discrete kernels. For simplicity of presentation, two positive coefficients are introduced [19]

$$I_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} (t_k - \xi) \varpi_n''(\xi) d\xi, \quad J_{n-k}^{(n)} := \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} (\xi - t_{k-1}) \varpi_n''(\xi) d\xi \quad \text{for } 1 \leq k \leq n - 1. \tag{3.9}$$

Some properties of the coefficients are also given as follows:

Lemma 3.2 *The positive coefficients in (3.9) satisfy*

$$\begin{aligned} (i) \quad & I_{n-k}^{(n)} \geq \frac{1 + \rho_k}{\rho_k} b_{n-k}^{(n)}, \quad (ii) \quad J_{n-k}^{(n)} \geq \frac{2(1 + \rho_k)}{\rho_k} b_{n-k}^{(n)} \quad \text{for } 1 \leq k \leq n - 1. \\ (iii) \quad & J_{n-k-1}^{(n)} \geq \frac{1}{\rho_k} J_{n-k}^{(n)} \quad \text{for } 1 \leq k \leq n - 2. \end{aligned}$$

Proof We refer to [19, Lemma 4.4] for (i)–(ii) and [19, Lemma 4.5] for (iii). \square

Next, we are ready to prove the following monotonicity lemma.

Lemma 3.3 *Let M1 holds. Denote $\epsilon_k = \frac{\alpha(1+5\rho_k)\tau_k}{105T} \omega_{1-\alpha}(T)$ for $1 \leq k \leq n - 1$. If the tolerance error $\epsilon \leq \min_{1 \leq k \leq n-1} \epsilon_k$, then the discrete convolutional kernels $A_{n-k}^{(n)}$ satisfy A2 for $1 \leq k \leq n$.*

Proof Comparing the discrete convolution kernels between two formulas and invoking to the inequalities (3.5)–(3.6), one has $A_0^{(1)} = A_0^{(1)}$ if $n = 1$, and for $n \geq 2$,

$$\left| A_{n-k}^{(n)} - A_{n-k}^{(n)} \right| \leq \begin{cases} \epsilon + \frac{3\epsilon}{10} \rho_{n-1} \leq \frac{29}{20} \epsilon & \text{for } k = n, \\ \epsilon + \frac{3\epsilon}{10} \rho_{k-1} + \frac{3\epsilon}{10} \leq \frac{7}{4} \epsilon & \text{for } 2 \leq k \leq n - 1, \\ \epsilon + \frac{3\epsilon}{10} \leq \frac{13}{10} \epsilon & \text{for } k = 1, \end{cases} \tag{3.10}$$

where **M1** was also used. Also, we can derive a similar estimate of $\left| A_{n-k-1}^{(n)} - A_{n-k-1}^{(n)} \right|$. Combining these two inequalities, we have $A_0^{(1)} = A_0^{(1)}$ if $n = 1$. For $n \geq 2$, $A_0^{(n)} \leq A_0^{(n)} + \frac{29\epsilon}{20}$ and

$$A_{n-k-1}^{(n)} - A_{n-k}^{(n)} + A_{n-k}^{(n)} - A_{n-k-1}^{(n)} \leq \begin{cases} 16\epsilon/5 & \text{for } k = n - 1, \\ 7\epsilon/2 & \text{for } 2 \leq k \leq n - 2, \\ 61\epsilon/20 & \text{for } k = 1. \end{cases} \tag{3.11}$$

According to [19, Theorem 2.2 (II)] and Lemma 3.2 (i), it holds that

$$A_{n-k-1}^{(n)} - A_{n-k}^{(n)} \geq (1 + \rho_k) b_{n-k}^{(n)} + \frac{1}{5} I_{n-k}^{(n)} \geq (1 + \rho_k) \left(\frac{1}{5\rho_k} + 1 \right) b_{n-k}^{(n)} \quad \text{for } 1 \leq k \leq n - 1. \tag{3.12}$$

Recalling the definition of $\varpi_n(\xi)$, we get the second derivative $\varpi_n''(\xi) = -\omega_{-\alpha}(t_{n-\theta} - \xi) > 0$ and the third derivative $\varpi_n'''(\xi) = \omega_{-\alpha-1}(t_{n-\theta} - \xi) > 0$ for $0 \leq \xi < t_{n-\theta}$. By means of [19, Lemma 2.1], for $1 \leq k \leq n - 1$, it is not difficult to check

$$\begin{aligned} b_{n-k}^{(n)} &= \int_{t_{k-1}}^{t_k} \frac{(t_k - \xi)(\xi - t_{k-1})}{\tau_k(\tau_k + \tau_{k+1})} \varpi_n''(\xi) d\xi \geq \frac{\tau_k^2}{6(\tau_k + \tau_{k+1})} \varpi_n''(t_{k-1}) \\ &= \frac{\rho_k \tau_k}{6(1 + \rho_k)} \varpi_n''(t_{k-1}). \end{aligned} \tag{3.13}$$

Substituting (3.13) into (3.12) and applying (3.11), one gets

$$A_{n-k-1}^{(n)} - A_{n-k}^{(n)} \geq A_{n-k-1}^{(n)} - A_{n-k}^{(n)} - \frac{7\epsilon}{2} \geq \frac{(1 + 5\rho_k)\tau_k}{30} \varpi_n''(t_{k-1}) - \frac{7\epsilon}{2}.$$

When the hypothesis holds, i.e., $\epsilon \leq \min_{1 \leq k \leq n-1} \frac{\alpha(1+5\rho_k)\tau_k}{105T} \omega_{1-\alpha}(T)$, we obtain

$$A_{n-k-1}^{(n)} - A_{n-k}^{(n)} > -\frac{(1 + 5\rho_k)\tau_k}{30} \omega_{-\alpha}(T) - \frac{7\epsilon}{2} = \frac{\alpha(1 + 5\rho_k)\tau_k}{30T} \omega_{1-\alpha}(T) - \frac{7\epsilon}{2} \geq 0.$$

This completes the proof of Lemma 3.3. □

Remark 1 In numerical simulations, the SOEs tolerance error ϵ is always set to be small enough so that the convergence order should not be degraded. Hence, the condition of ϵ in Lemma 3.3 is allowed to guarantee the monotonicity of discrete convolutional kernels $A_{n-k}^{(n)}$.

In addition to the boundedness and monotonicity described above, the following auxiliary lemma is also indispensable in the subsequent analysis.

Lemma 3.4 *Under the conditions of Lemma 3.3, it holds that*

$$A_0^{(n)} - A_1^{(n)} > \theta(2A_0^{(n)} - A_1^{(n)}) \quad \text{for } n \geq 2.$$

Proof The conclusion is equivalent to $\frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} > 0$. For $n \geq 2$, (3.10) yields

$$\frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} \geq \frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} - \frac{1-2\theta}{1-\theta} \cdot \frac{29\epsilon}{20} - \frac{7\epsilon}{4} > \frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} - \frac{16\epsilon}{5}.$$

Resorting to the proof of [19, Theorem 2.2 (III)], we have

$$\frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} > \begin{cases} J_1^{(n)} + b_1^{(n)} & \text{for } n = 2, \\ J_1^{(n)} - \rho_{n-2} b_2^{(n)} + b_1^{(n)} & \text{for } n \geq 3. \end{cases}$$

By means of Lemma 3.2 (ii)–(iii) with $k = n - 2$, it is easy to check

$$J_1^{(n)} = \frac{\rho_{n-2}^3}{2(1 + \rho_{n-2})} J_1^{(n)} + \frac{2 + 2\rho_{n-2} - \rho_{n-2}^3}{2(1 + \rho_{n-2})} J_1^{(n)} \geq \rho_{n-2} b_2^{(n)} + \frac{13}{40} J_1^{(n)},$$

where **M1** and $\frac{2+2x-x^3}{2(1+x)} \geq \frac{13}{40}$ for $x \in (0, \frac{3}{2}]$ were used. Then, it follows that

$$\begin{aligned} J_1^{(n)} - \rho_{n-2}b_2^{(n)} + b_1^{(n)} - \frac{16\epsilon}{5} &\geq \frac{13 + 33\rho_{n-1}}{20\rho_{n-1}}b_1^{(n)} - \frac{16\epsilon}{5} \\ &\geq \frac{(13 + 33\rho_{n-1})\tau_{n-1}}{120(1 + \rho_{n-1})} \varpi_n''(t_{n-2}) - \frac{16\epsilon}{5}, \end{aligned}$$

where Lemma 3.2 (ii) with $k = n - 1$ and (3.13) were used. Under the conditions of Lemma 3.3, one has $\epsilon \leq \epsilon_{n-1} = \frac{\alpha(1+3\rho_{n-1})\tau_{n-1}}{105T} \omega_{1-\alpha}(T)$. This together with **M1** leads to

$$J_1^{(n)} - \rho_{n-2}b_2^{(n)} + b_1^{(n)} - \frac{16\epsilon}{5} > \frac{\alpha(13 + 33\rho_{n-1})\tau_{n-1}}{120(1 + \rho_{n-1})T} \omega_{1-\alpha}(T) - \frac{16\epsilon}{5} > 0.$$

Overall, for $n \geq 2$, we have $\frac{1-2\theta}{1-\theta} A_0^{(n)} - A_1^{(n)} > 0$. This completes the proof of Lemma 3.4. □

If the tolerance error $\epsilon \leq \min\{\epsilon^*, \epsilon^{**}, \min_{1 \leq k \leq n-1} \epsilon_k\}$ and **M1** holds, Lemmas 3.1 and 3.3 state that the discrete convolution kernels $A_{n-k}^{(n)}$ satisfy the assumptions **A1-A2** and have the property in Lemma 3.4. According to Lemmas 3.3 and 3.4, one has the following result.

Corollary 3.1 *Under the conditions of Lemma 3.3, the fast Alikhanov formula (2.5) satisfies*

$$v^{n-\theta} (\partial_{\alpha\tau}^\alpha v)^{n-\theta} \geq \frac{1}{2} \sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau \left(|v^k|^2 \right) \quad \text{for } 1 \leq n \leq N.$$

Proof Lemma 3.3 implies that the discrete convolution kernels $A_{n-k}^{(n)}$ are monotone, and Lemma 3.4 shows that $\theta^{(n)} \geq \theta$ for $1 \leq n \leq N$, where

$$\theta^{(1)} := \frac{1}{2} \quad \text{and} \quad \theta^{(n)} := \frac{A_0^{(n)} - A_1^{(n)}}{2A_0^{(n)} - A_1^{(n)}} \quad \text{for } n \geq 2.$$

Then, the proof of [18, Lemma 4.1] leads to the claimed result. □

3.2 H^1 -norm stability

In the following analysis, we also need some results on the spatial approximation \mathcal{D}_h . Recalling the compact difference operator, it is well-known that the matrix corresponding to \mathcal{I}_k is a tridiagonal, real, symmetric, and positive definite matrix, and the eigenvalues are in the form of

$$\lambda_{k,j_k} = \frac{5}{6} + \frac{1}{6} \cos\left(\frac{j_k \pi}{M_k}\right) \quad \text{for } 1 \leq j_k \leq M_k - 1, 1 \leq k \leq d.$$

Denote H_k is the matrix associated with the operator \mathcal{I}_k^{-1} and λ_{H_k, j_k} represents the eigenvalues of H_k . For $1 \leq k \leq d$, it is easy to check $1 < \lambda_{H_k, j_k} < \frac{3}{2}$ and

$$\frac{3}{2} \langle \Delta_h v_h, v_h \rangle_h < \langle \mathfrak{D}_h v_h, v_h \rangle_h = \left\langle \sum_{k=1}^d H_k \delta_k^2 v_h, v_h \right\rangle_h < \langle \Delta_h v_h, v_h \rangle_h. \tag{3.14}$$

Now we establish the stability of the time-stepping scheme (1.6)–(1.8). Taking the discrete inner product of (1.6) with $2(\partial_{\alpha\tau}^\alpha u_h)^{n-\theta}$ and applying Cauchy-Schwarz inequality yield

$$\begin{aligned} 2\|(\partial_{\alpha\tau}^\alpha u_h)^{n-\theta}\|^2 &= 2\langle \mathfrak{D}_h u_h^{n-\theta}, (\partial_{\alpha\tau}^\alpha u_h)^{n-\theta} \rangle_h + 2\kappa \langle u_h^{n-\theta}, (\partial_{\alpha\tau}^\alpha u_h)^{n-\theta} \rangle_h \\ &\quad + 2\langle f^{n-\theta}, (\partial_{\alpha\tau}^\alpha u_h)^{n-\theta} \rangle_h \\ &\leq 2\langle \mathfrak{D}_h u_h^{n-\theta}, (\partial_{\alpha\tau}^\alpha u_h)^{n-\theta} \rangle_h + \kappa^2 \|u_h^{n-\theta}\|^2 \\ &\quad + \|f^{n-\theta}\|^2 + 2\|(\partial_{\alpha\tau}^\alpha u_h)^{n-\theta}\|^2. \end{aligned}$$

With the aids of (3.14) and embedding inequality (2.6), we apply Corollary 3.1 with $v := \Delta_h^{1/2} u_h$ and obtain

$$\sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau (|u_h^k|_1^2) \leq \kappa^2 C_\Omega |u_h^{n-\theta}|_1^2 + \|f^{n-\theta}\|^2 \quad \text{for } 1 \leq n \leq N,$$

which has the form of (1.12) with $\lambda := \kappa^2 C_\Omega$, $v^k := |u_h^k|_1^2$, and $\eta^n := \|f^{n-\theta}\|^2$. Under the conditions of Lemmas 3.1 and 3.3, we see that Lemma 1.1 holds with $\pi_A = 2$. Consequently, Lemma 1.1 indicates the H^1 -norm stability of the presented scheme in the following sense.

Theorem 3.1 *Let the conditions of Lemmas 3.1 and 3.3 hold. If the restriction of maximum time-step $\tau \leq 1/\sqrt{4\kappa^2\Gamma(2-\alpha)C_\Omega}$, then the numerical solution u_h^n of the scheme (1.6)–(1.8) is stable in the sense that*

$$\begin{aligned} |u_h^k|_1^2 &\leq 2E_\alpha \left(6\kappa^2 t_n^\alpha C_\Omega \right) \left(|u_h^0|_1^2 + \max_{1 \leq k \leq n} \sum_{j=1}^k P_{k-j}^{(k)} \|f^{j-\theta}\|^2 \right) \\ &\leq 2E_\alpha \left(6\kappa^2 t_n^\alpha C_\Omega \right) \left(|u_h^0|_1^2 + 2\Gamma(1-\alpha) \max_{1 \leq k \leq n} \{t_k^\alpha \|f^{k-\theta}\|^2\} \right) \quad \text{for } 1 \leq n \leq N. \end{aligned}$$

4 Global consistency analysis and convergence

4.1 Global consistency error analysis

We now proceed with the global consistency error analysis of the fast Alikhanov formula (2.5). First of all, the local consistency error of the standard Alikhanov formula on nonuniform grids in (2.1) is given (slightly simplified) from [19, Theorem 3.4].

Lemma 4.1 *Let M1 holds. Assume that $v \in C^3(0, T]$, and there exists a positive constant C_v such that $|v'''(t)| \leq C_v(1 + t^{\sigma-3})$ for $0 < t \leq T$, where $\sigma \in (0, 1) \cup (1, 2)$ is a regularity parameter. For the standard Alikhanov formula on nonuniform grids, there exist*

$$|(\partial_t^\alpha v)(t_{n-\theta}) - (\partial_\tau^\alpha v)^{n-\theta}| \leq A_0^{(n)} G_{\text{loc}}^n + \sum_{k=1}^{n-1} (A_{n-k-1}^{(n)} - A_{n-k}^{(n)}) G_{\text{his}}^k \quad \text{for } 1 \leq n \leq N,$$

where two quantities G_{loc}^n and G_{his}^k are defined as

$$G_{\text{loc}}^n := \frac{3}{2} \int_{t_{n-1}}^{t_{n-1/2}} (\xi - t_{n-1})^2 |v'''(\xi)| d\xi + \frac{3\tau_n}{2} \int_{t_{n-1/2}}^{t_n} (t_n - \xi) |v'''(\xi)| d\xi,$$

$$G_{\text{his}}^k := \frac{5}{2} \int_{t_{k-1}}^{t_k} (\xi - t_{k-1})^2 |v'''(\xi)| d\xi + \frac{5}{2} \int_{t_k}^{t_{k+1}} (t_{k+1} - \xi)^2 |v'''(\xi)| d\xi.$$

The main difference between the fast Alikhanov formula and the standard one is that the convolution kernels are approximated by SOEs with the tolerance error ϵ . Denote $\hat{t}_j = \max\{1, t_j\}$ for $1 \leq j \leq N$. When $j \geq 2$, the regularity assumptions (1.14) and (3.10) lead to

$$\begin{aligned} |(\partial_{a\tau}^\alpha v)^{j-\theta} - (\partial_\tau^\alpha v)^{j-\theta}| &\leq \sum_{k=1}^{j-1} |\nabla_\tau v^k| |A_{j-k}^{(j)} - A_{j-k}^{(j)}| \\ &\leq \frac{7\epsilon}{4} \sum_{k=1}^{j-1} \int_{t_{k-1}}^{t_k} |v'(\xi)| d\xi \leq \frac{C_v}{\sigma} t_{j-1}^\sigma \epsilon \leq \frac{C_v}{\sigma} \hat{t}_{j-1}^2 \epsilon. \end{aligned}$$

From Lemma 4.1 and (3.11), we have $|(\partial_t^\alpha v)(t_{1-\theta}) - (\partial_\tau^\alpha v)^{1-\theta}| \leq A_0^{(1)} G_{\text{loc}}^1$ if $j = 1$, and for $j \geq 2$,

$$\begin{aligned} |(\partial_t^\alpha v)(t_{j-\theta}) - (\partial_\tau^\alpha v)^{j-\theta}| &\leq A_0^{(j)} G_{\text{loc}}^j + \sum_{k=1}^{j-1} (A_{j-k-1}^{(j)} - A_{j-k}^{(j)}) G_{\text{his}}^k \\ &\quad + \frac{29\epsilon}{20} G_{\text{loc}}^j + \frac{7\epsilon}{2} \sum_{k=1}^{j-1} G_{\text{his}}^k. \end{aligned}$$

In view of the regularity assumptions (1.14) and the definition of G_{loc}^n , it is not difficult to find $G_{\text{loc}}^1 \leq C_v \sigma^{-1} \tau_1^\sigma$ and $G_{\text{loc}}^k \leq C_v t_{k-1}^{\sigma-3} \tau_k^3$ for $2 \leq k \leq N$. Accordingly, the regularity assumptions (1.14) and the definition of G_{his}^k mentioned above imply that $G_{\text{his}}^1 \leq C_v (\sigma^{-1} \tau_1^\sigma + t_1^{\sigma-3} \tau_2^3)$ and $G_{\text{his}}^k \leq C_v (t_{k-1}^{\sigma-3} \tau_k^3 + t_k^{\sigma-3} \tau_{k+1}^3)$ for $2 \leq k \leq N - 1$. Hence, we have

$$\begin{aligned} \frac{29\epsilon}{20} G_{\text{loc}}^j + \frac{7\epsilon}{2} \sum_{k=1}^{j-1} G_{\text{his}}^k &\leq C_h \left(t_{j-1}^\sigma \epsilon + \frac{t_1^\sigma}{\sigma} \epsilon + \sum_{k=2}^{j-1} (t_{k-1}^\sigma + t_k^\sigma) \epsilon \right) \\ &\leq \frac{C_v}{\sigma} \hat{t}_{j-1}^2 \epsilon \quad \text{for } j \geq 2. \end{aligned}$$

Denote the local consistency error

$$\Upsilon^{n-\theta}[v] := (\partial_t^\alpha v)(t_{n-\theta}) - (\partial_{\tau t}^\alpha v)^{n-\theta} \tag{4.1}$$

a triangle inequality leads to $|\Upsilon^{1-\theta}[v]| \leq A_0^{(1)} G_{\text{loc}}^1$ if $j = 1$, and for $j \geq 2$,

$$|\Upsilon^{j-\theta}[v]| \leq A_0^{(j)} G_{\text{loc}}^j + \sum_{k=1}^{j-1} \left(A_{j-k-1}^{(j)} - A_{j-k}^{(j)} \right) G_{\text{his}}^k + \frac{C_v}{\sigma} \hat{t}_{j-1}^2 \epsilon. \tag{4.2}$$

Multiplying the inequality (4.2) by $P_{n-j}^{(n)}$ and summing the index j from 1 to n , we exchange the order of summation and apply the definition (1.9) of $P_{n-j}^{(n)}$ to get

$$\begin{aligned} \sum_{j=1}^n P_{n-j}^{(n)} |\Upsilon^{j-\theta}[v]| &\leq \sum_{j=1}^n P_{n-j}^{(n)} A_0^{(j)} G_{\text{loc}}^j + \sum_{j=2}^n P_{n-j}^{(n)} \sum_{k=1}^{j-1} \left(A_{j-k-1}^{(j)} - A_{j-k}^{(j)} \right) G_{\text{his}}^k + \frac{C_v}{\sigma} \sum_{j=2}^n P_{n-j}^{(n)} \hat{t}_{j-1}^2 \epsilon \\ &\leq \sum_{j=1}^n P_{n-j}^{(n)} A_0^{(j)} G_{\text{loc}}^j + \sum_{k=1}^{n-1} G_{\text{his}}^k \sum_{j=k+1}^n P_{n-j}^{(n)} \left(A_{j-k-1}^{(j)} - A_{j-k}^{(j)} \right) + \frac{C_v \hat{t}_{n-1}^2 \epsilon}{\sigma} \sum_{j=2}^n P_{n-j}^{(n)} \\ &\leq \sum_{k=1}^n P_{n-k}^{(n)} A_0^{(k)} G_{\text{loc}}^k + \sum_{k=1}^{n-1} P_{n-k}^{(n)} A_0^{(k)} G_{\text{his}}^k + \frac{C_v}{\sigma} \hat{t}_n^2 \epsilon. \end{aligned} \tag{4.3}$$

where (1.11) with $m = 1$ and $\pi_A = 2$ was used in the last step. Under the conditions of Lemma 3.1, the boundedness implies that $A_0^{(k)} \leq \frac{2}{\tau_k} \omega_{2-\alpha}(\tau_k)$, $A_{k-2}^{(k)} \geq \frac{1}{2} \omega_{1-\alpha}(t_k - t_1)$ and

$$\frac{A_0^{(k)}}{A_{k-2}^{(k)}} \leq \frac{4\omega_{2-\alpha}(\tau_k)}{\tau_k \omega_{1-\alpha}(t_k - t_1)} \leq \frac{4}{1-\alpha} \frac{(t_k - t_1)^\alpha}{\tau_k^\alpha} \leq \frac{C_v}{1-\alpha} t_k^\alpha \tau_k^{-\alpha} \quad \text{for } 2 \leq k \leq n \leq N.$$

Moreover, the identity (1.10) for the complementary discrete kernels $P_{n-j}^{(n)}$ gives

$$P_{n-1}^{(n)} A_0^{(1)} \leq 1 \quad \text{and} \quad \sum_{k=2}^{n-1} P_{n-k}^{(n)} A_{k-2}^{(k)} \leq \sum_{k=2}^n P_{n-k}^{(n)} A_{k-2}^{(k)} = 1.$$

Then, it follows from (4.3) that

$$\sum_{j=1}^n P_{n-j}^{(n)} |\Upsilon^{j-\theta}[v]| \leq P_{n-1}^{(n)} A_0^{(1)} \left(G_{\text{loc}}^1 + G_{\text{his}}^1 \right) + \sum_{k=2}^n P_{n-k}^{(n)} A_0^{(k)} G_{\text{loc}}^k + \sum_{k=2}^{n-1} P_{n-k}^{(n)} A_0^{(k)} G_{\text{his}}^k + \frac{C_v}{\sigma} \hat{t}_n^2 \epsilon.$$

The first term on the right-hand side is bounded by $C_v \left(\sigma^{-1} \tau_1^\sigma + t_1^{\sigma-3} \tau_2^3 \right) \leq C_v \frac{\tau_1^\sigma}{\sigma}$, and two terms in the middle can be bounded by

$$\begin{aligned} &\frac{C_v}{1-\alpha} \left(\sum_{k=2}^n P_{n-k}^{(n)} A_{k-2}^{(k)} t_k^\alpha \tau_k^{-\alpha} G_{\text{loc}}^k + \sum_{k=2}^{n-1} P_{n-k}^{(n)} A_{k-2}^{(k)} t_k^\alpha \tau_k^{-\alpha} G_{\text{his}}^k \right) \\ &\leq \frac{C_v}{1-\alpha} \left(\max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_{k-1}^{3-\alpha} + \max_{2 \leq k \leq n-1} t_k^\alpha (t_{k-1}^{\sigma-3} \tau_k^{3-\alpha} + t_k^{\sigma-3} \tau_{k+1}^3 \tau_k^{-\alpha}) \right) \\ &\leq \frac{C_v}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_{k-1}^{3-\alpha} (\rho_{k-1}^\alpha + 1) \leq \frac{C_v}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_{k-1}^{3-\alpha} \quad \text{for } 2 \leq n \leq N. \end{aligned}$$

Combining these inequalities together, one has

$$\sum_{j=1}^n P_{n-j}^{(n)} |\Upsilon^{j-\theta}[v]| \leq C_v \left(\frac{\tau_1^\sigma}{\sigma} + \frac{1}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} + \frac{\epsilon}{\sigma} t_n^\alpha \hat{\tau}_{n-1}^2 \right).$$

In conclusion, by taking the initial singularity into account, we have the following global consistency error analysis of the fast Alikhanov approximation on nonuniform meshes.

Lemma 4.2 *Assume that $v \in C^3(0, T]$ and there exist $C_v > 0$ such that $|v'(t)| \leq C_v(1 + t^{\sigma-1})$ and $|v'''(t)| \leq C_v(1 + t^{\sigma-3})$ for $0 < t \leq T$, where $\sigma \in (0, 1) \cup (1, 2)$ is a regularity parameter. Denote $\hat{t}_n = \max\{1, t_n\}$. Under the conditions of Lemma 3.1, the global consistency error of the fast Alikhanov approximation (2.5) satisfies*

$$\sum_{j=1}^n P_{n-j}^{(n)} |\Upsilon^{j-\theta}[v]| \leq C_v \left(\frac{\tau_1^\sigma}{\sigma} + \frac{1}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} + \frac{\epsilon}{\sigma} t_n^\alpha \hat{\tau}_{n-1}^2 \right) \quad \text{for } 1 \leq n \leq N.$$

4.2 Sharp H^1 -norm error analysis

As described in section 3.1 of [20], under the assumption of AssG, the traditional H^1 -norm analysis together with the discrete Grönwall inequality in Lemma 1.1 shows the error estimate of the time-stepping scheme (1.6)–(1.8) as

$$|u(t_n) - u_h^n|_1 \leq \frac{C_u}{\sigma(1-\alpha)} \left(\tau^{\min\{2-\alpha, \gamma(\sigma-\alpha/2)\}} + h^4 + \epsilon \right) \quad \text{for } 1 \leq n \leq N.$$

It is obvious that a loss of theoretical accuracy $O(\tau^{\gamma \frac{\alpha}{2}})$ exists in time under the regularity assumption (1.14), which leads to a suboptimal error estimate due to the initial singularity and the discrete convolution form. The authors utilize the time-space error splitting technique to overcome this problem (see [20] for more details). In the following analysis, an alternative two-stage error analysis is applied to establish sharp H^1 -norm error estimate for the fully discrete scheme (1.6)–(1.8) under the realistic assumption of solution.

Define the continuous inner product $(v, w) := \int_\Omega v(\mathbf{x})w(\mathbf{x})d\mathbf{x}$ with the associated L^2 -norm $\|v\|_{L^2} := \sqrt{(v, v)}$. Also, the H^1 semi-norm is defined as $|v|_{H^1} := \sqrt{(-\Delta v, v)}$. It is easy to check $\|v\|_{L^2} \leq C_\Omega |v|_{H^1}$ via the embedding inequality. Additionally, we have $|v|_1 \leq C_\Omega |v|_{H^1}$ by using the Cauchy-Schwarz inequality.

Stage 1: Temporal error analysis for time-discrete system We apply the fast Alikhanov formula (3.2) to approximate the problem (1.1)–(1.3) and obtain the time-discrete system

$$(D_\tau^\alpha u)^{n-\theta} = \Delta u^{n-\theta} + \kappa u^{n-\theta} + f(\mathbf{x}, t_{n-\theta}) \quad \text{for } \mathbf{x} \in \Omega \text{ and } 1 \leq n \leq N, \tag{4.4}$$

$$u^n(\mathbf{x}) = 0 \quad \text{for } \mathbf{x} \in \partial\Omega \text{ and } 1 \leq n \leq N, \tag{4.5}$$

$$u^0(\mathbf{x}) = u_0(\mathbf{x}) \quad \text{for } \mathbf{x} \in \bar{\Omega}. \tag{4.6}$$

The existence and uniqueness of the solution u^n can be proved straightforward since it is a linear elliptic problem. Denote the solution error $e^n(\mathbf{x}) := u(\mathbf{x}, t_n) - u^n$

for $\mathbf{x} \in \Omega$, then the governing equation reads

$$(D_\tau^\alpha e)^{n-\theta} = \Delta e^{n-\theta} + \kappa e^{n-\theta} + \Upsilon^{n-\theta}[u] + R_\tau^{n-\theta}, \tag{4.7}$$

where $\Upsilon^{n-\theta}[u]$ is the local consistency error described in (4.1), and

$$R_\tau^{n-\theta} := (\kappa + \Delta)[u^{n-\theta}(\mathbf{x}) - u(\mathbf{x}, t_{n-\theta})] \quad \text{for } \mathbf{x} \in \Omega. \tag{4.8}$$

Taking the continuous inner product of (4.7) with $-\Delta e^{n-\theta}$, we have

$$((D_\tau^\alpha e)^{n-\theta}, -\Delta e^{n-\theta}) + \|\Delta e^{n-\theta}\|_{L^2}^2 = (\kappa e^{n-\theta}, -\Delta e^{n-\theta}) - (\Upsilon^{n-\theta}[u], \Delta e^{n-\theta}) - (R_\tau^{n-\theta}, \Delta e^{n-\theta}).$$

In terms of the first Green formula and Cauchy-Schwarz inequality, we apply Corollary 3.1 with $v := \Delta^{1/2}e$ and obtain

$$\sum_{k=1}^n \mathbf{A}_{n-k}^{(n)} \nabla_\tau \left(|e^k|_{H^1}^2 \right) \leq 2\kappa |e^{n-\theta}|_{H^1}^2 + 2|e^{n-\theta}|_{H^1} \left(|\Upsilon^{n-\theta}[u]|_{H^1} + |R_\tau^{n-\theta}|_{H^1} \right),$$

which takes the form of (1.13) with $\lambda := 2\kappa$, $v^k := |e^k|_{H^1}$, and $\eta^n := 2 \left(|\Upsilon^{n-\theta}[u]|_{H^1} + |R_\tau^{n-\theta}|_{H^1} \right)$. By virtue of Lemmas 3.1 and 3.3, we see that Lemma 1.1 is satisfied with $\pi_A = 2$. Thus, under the restriction of maximum time-step size $\tau \leq 1/\sqrt[4]{8\kappa\Gamma(2-\alpha)}$, Lemma 1.1 gives

$$|e^n|_{H^1} \leq 2E_\alpha \left(12\kappa t_n^\alpha \right) \max_{1 \leq k \leq n} \sum_{j=1}^k \mathbf{P}_{k-j}^{(k)} \left(|\Upsilon^{j-\theta}[u]|_{H^1} + |R_\tau^{j-\theta}|_{H^1} \right).$$

Similar to the proof of [19, Lemma 3.8], it is easy to find

$$\sum_{j=1}^n \mathbf{P}_{n-j}^{(n)} |R_\tau^{j-\theta}| \leq C_u \left(\frac{\tau_1^{\sigma+\alpha}}{\sigma} + t_n^\alpha \max_{2 \leq k \leq n} t_{k-1}^{\sigma-2} \tau_k^2 \right) \quad \text{for } 1 \leq n \leq N. \tag{4.9}$$

Therefore, Lemma 4.2 and (4.9) show that

$$|e^n|_{H^1} \leq C_u \left(\frac{\tau_1^\sigma}{\sigma} + \frac{1}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} + t_n^\alpha \max_{2 \leq k \leq n} t_{k-1}^{\sigma-2} \tau_k^2 + \frac{\epsilon}{\sigma} t_n^\alpha \hat{t}_{n-1}^2 \right). \tag{4.10}$$

Stage 2: Spatial error analysis for fully discrete system Now return to the fully discrete system, which can be viewed as the spatial approximation of time-discrete system (4.4)–(4.6). Under the priori assumptions in (1.14), this system has a unique solution $u^n \in H^4(\Omega)$ for $1 \leq n \leq N$. Denote the solution error $e_h^n := u^n - u_h^n$ for $\mathbf{x}_h \in \Omega_h$, the governing equation reads

$$(\partial_{\tau}^\alpha e_h)^{n-\theta} = \mathfrak{D}_h e_h^{n-\theta} + \kappa e_h^{n-\theta} + R_s^{n-\theta}, \tag{4.11}$$

where $R_s^{n-\theta} = (\mathfrak{D}_h - \Delta)u^{n-\theta}$. By means of Taylor’s expansion with integral reminder, it is easy to check that $\|R_s^{n-\theta}\| \leq C_u h^4$.

Taking the discrete inner product of (4.11) with $2(\partial_{\tau}^\alpha e_h)^{n-\theta}$, one has

$$2\|(\partial_{\tau}^\alpha e_h)^{n-\theta}\|^2 - 2\langle \mathfrak{D}_h e_h^{n-\theta}, (\partial_{\tau}^\alpha e_h)^{n-\theta} \rangle_h = 2\langle \kappa e_h^{n-\theta}, (\partial_{\tau}^\alpha e_h)^{n-\theta} \rangle_h + 2\langle R_s^{n-\theta}, (\partial_{\tau}^\alpha e_h)^{n-\theta} \rangle_h.$$

By means of (3.14) and the embedding inequality (2.6), we apply Corollary 3.1 with $v := \Delta_h^{1/2} e_h$ and obtain

$$\sum_{k=1}^n A_{n-k}^{(n)} \nabla_\tau \left(|e_h^k|_1^2 \right) \leq \kappa^2 \|e_h^{n-\theta}\|^2 + \|R_s^{n-\theta}\|^2 \leq \kappa^2 C_\Omega |e_h^{n-\theta}|_1^2 + \|R_s^{n-\theta}\|^2,$$

which has the form of (1.12) with $\lambda := \kappa^2 C_\Omega$, $v^k := |e_h^k|_1^2$, and $\eta^n = \|R_s^{n-\theta}\|^2$. Under the conditions of Lemmas 3.1 and 3.3, we see that Lemma 1.1 holds with $\tau_A = 2$. If the maximum time-step size $\tau \leq 1/\sqrt[4]{4\kappa^2 \Gamma(2-\alpha) C_\Omega}$, Lemma 1.1 yields

$$|e_h^n|_1 \leq \sqrt{2E_\alpha(6\kappa^2 C_\Omega t_n^\alpha) \Gamma(1-\alpha)} \max_{1 \leq k \leq n} \{t_k^{\alpha/2} \|R_s^{k-\theta}\|\} \leq C_u t_n^{\alpha/2} h^4. \tag{4.12}$$

Combining (4.10) with (4.12), we utilize the triangle inequality

$$\begin{aligned} |u(t_n) - u_h^n|_1 &\leq |e^n|_1 + |e_h^n|_1 \leq C_\Omega |e^n|_{H^1} + |e_h^n|_1 \\ &\leq C_u \left(\frac{\tau_1^\sigma}{\sigma} + \frac{1}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} + t_n^\alpha \max_{2 \leq k \leq n} t_{k-1}^{\sigma-2} \tau_k^2 + t_n^{\alpha/2} h^4 + \frac{\epsilon}{\sigma} t_n^\alpha \tau_{n-1}^2 \right). \end{aligned}$$

Overall, we establish the convergence of the numerical solution.

Theorem 4.1 *Assume that the exact solution u satisfies the regularity property in (1.14) with the parameter $\sigma \in (0, 1) \cup (1, 2)$. Let **M1** holds and $\hat{t}_n = \max\{1, t_n\}$ for $1 \leq n \leq N$. Under the restriction of maximum time-step $\tau \leq 1/\sqrt[4]{4\kappa \max\{2, \kappa C_\Omega\} \Gamma(2-\alpha)}$ and the SOEs tolerance error $\epsilon \leq \min\{\epsilon^*, \epsilon^{**}, \min_{1 \leq k \leq n-1} \epsilon_k\}$, the numerical solution u_h^n is convergent with respect to the discrete H^1 -norm,*

$$|u(t_n) - u_h^n|_1 \leq C_u \left(\frac{\tau_1^\sigma}{\sigma} + \frac{1}{1-\alpha} \max_{2 \leq k \leq n} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} + t_n^\alpha \max_{2 \leq k \leq n} t_{k-1}^{\sigma-2} \tau_k^2 + t_n^{\alpha/2} h^4 + \frac{\epsilon}{\sigma} t_n^\alpha \tau_{n-1}^2 \right).$$

As mentioned in Section 1, the aforesaid analysis is applicable to a wider class of unequal time-steps satisfying **M1**. On the other hand, if **AssG** holds, one has

$$\begin{aligned} t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} &\leq C_\gamma t_k^{\alpha+\sigma-3} \tau_k^{3-\alpha} \leq C_\gamma t_k^{\alpha+\sigma-3} \tau_k^{3-\alpha-\beta} (\tau \min\{1, t_k^{1-1/\gamma}\})^\beta \\ &\leq C_\gamma t_k^{\sigma-\beta/\gamma} (\tau_k/t_k)^{3-\alpha-\beta} \tau^\beta \leq C_\gamma t_k^{\max\{0, \sigma-(3-\alpha)/\gamma\}} \tau^\beta \quad \text{for } 2 \leq k \leq n, \end{aligned} \tag{4.13}$$

where $\beta := \min\{2, \gamma\sigma\}$. Obviously, $\tau_1^\sigma \leq C_\gamma \tau^{\gamma\sigma} \leq C_\gamma \tau^\beta$. Moreover, it is easy to check

$$\begin{aligned} t_{k-1}^{\sigma-2} \tau_k^2 &\leq C_\gamma t_k^{\sigma-2} \tau_k^{2-\beta} (\tau \min\{1, t_k^{1-1/\gamma}\})^\beta \\ &\leq C_\gamma t_k^{\sigma-\beta/\gamma} (\tau_k/t_k)^{2-\beta} \tau^\beta \leq C_\gamma t_k^{\max\{0, \sigma-2/\gamma\}} \tau^\beta \quad \text{for } 2 \leq k \leq n. \end{aligned} \tag{4.14}$$

Then, Theorem 4.1 gives the following corollary.

Corollary 4.1 *Let the conditions of Theorem 4.1 hold. If the time mesh satisfies AssG, the discrete solution u_h^n is convergent in the sense that*

$$|u(t_n) - u_h^n|_1 \leq \frac{C_u}{\sigma(1-\alpha)} \left(\tau^{\min\{\gamma\sigma, 2\}} + h^4 + \epsilon \right) \quad \text{for } 1 \leq n \leq N.$$

It says that the optimal grading parameter $\gamma_{\text{opt}} = 2/\sigma$.

Remark 2 Actually, an alternative estimate of (4.13) can be obtained similarly

$$t_k^\alpha t_{k-1}^{\sigma-3} \tau_k^3 \tau_{k-1}^{-\alpha} \leq C_\gamma t_k^{\alpha+\sigma-3} \tau_k^{3-\alpha} \leq C_\gamma t_k^{\alpha+\sigma-3} (\tau \min\{1, t_k^{1-1/\gamma}\})^{3-\alpha} \leq C_\gamma t_k^{\sigma-(3-\alpha)/\gamma} \tau^{3-\alpha}.$$

This means that the fast nonuniform Alikhanov formula $(\partial_{\alpha\tau}^\alpha v)^{n-\theta}$ approximates $(\partial_t^\alpha v)(t_{n-\theta})$ to order $O(\tau^{\min\{\gamma\sigma, 3-\alpha\}})$, and the optimal grading parameter $\gamma_{\text{opt}} = (3-\alpha)/\sigma$. However, restricted by the term (4.14) arising from $R_\tau^{n-\theta}$ in (4.8), the temporal convergence rate of the new scheme is to order $O(\tau^2)$ if $\gamma \geq 2/\sigma$.

5 Numerical examples

In this section, three examples are presented to verify the effectiveness and accuracy of the new scheme. For comparisons, we denote the numerical scheme using the standard Alikhanov (Alikhanov) approximation (2.4) as Scheme 1 in the tables, while the fully discrete scheme (1.6)–(1.8) using the accelerated Alikhanov (AccA) formula is called Scheme 2 for simplicity. To illustrate superiority of the fast algorithm, comparisons of computational cost between Scheme 1 and Scheme 2 are also indicated.

Example 1 To test the validity and accuracy of the fast Alikhanov approximation, we begin by considering a fractional ordinary differential equation $\mathcal{D}_t^\alpha u = \omega_{1+\sigma-\alpha}(t)$. The time interval is $(0, 1)$ and the exact solution is $u = \omega_{1+\sigma}(t)$ which satisfies the regularity assumptions (1.14). To trade off the accuracy and efficiency, here and in what follows, we choose $\epsilon = 10^{-12}$ in the simulations for the fast algorithm.

Tables 1 and 2 display the temporal error and convergence order in uniform mesh ($\gamma = 1$). Due to the lack of smoothness near the initial time, we find that all convergence orders are hard to achieve $O(\tau^{3-\alpha})$. Orders in two tables can be described as $O(\tau^\sigma)$ which supports the sharp error estimate in Remark 2. By comparison test, we observe that the convergence rate is improved with the increase of the regularity parameter σ . It is also in agreement with the theoretical analysis. Additionally, computational errors in two tables indicate that the accelerated formula has the same accuracy as well as the standard one.

The graded grids are employed to improve the convergence order. Tables 3, 4, and 5 list the temporal error and convergence order for $\sigma = 0.8$ in nonuniform partitions with different parameters α and γ . The effectiveness of the fast Alikhanov formula is

Table 2 Temporal error of Example 1 for different α with $\sigma = 1.6, \gamma = 1$

N	$\alpha = 0.2$			$\alpha = 0.5$			$\alpha = 0.8$			
	Alikhanov AccA		Order	Alikhanov AccA		Order	Alikhanov AccA		Order	
	Error	Error		Error	Error		Error	Error		
80	3.96e−06	3.96e−06		1.18e−05	1.18e−05		2.59e−05	2.59e−05		
160	1.30e−06	1.30e−06	1.60	3.90e−06	3.90e−06	1.60	8.54e−06	8.54e−06	1.60	
320	4.30e−07	4.30e−07	1.60	1.29e−06	1.29e−06	1.60	2.82e−06	2.82e−06	1.60	
640	1.42e−07	1.42e−07	1.60	4.25e−07	4.25e−07	1.60	9.29e−07	9.29e−07	1.60	
1280	4.68e−08	4.68e−08	1.60	1.40e−07	1.40e−07	1.60	3.06e−07	3.06e−07	1.60	
min{ $\sigma, 3 - \alpha$ }			1.60				1.60			

still held in this situation due to the same temporal error between two formulas. Comparing with Table 1, it is not difficult to find that the convergence rate is increased with different γ . Numerical results in Tables 3, 4, and 5 witness the predicted time accuracy in Remark 2. Moreover, invoking to the data in three tables, one gets the optimal mesh parameter $\gamma_{opt} = (3 - \alpha)/\sigma$ which is also coincide with our analysis. All these results substantiate the validity of nonuniform meshes in resolving the initial singularity and the sharpness of error estimate in Remark 2. Also, we see that the fast Alikhanov approximation is effective no matter in uniform grid or nonuniform ones.

Example 2 We consider the problem (1.1)–(1.3) in $\Omega = (0, \pi)^2$ to witness the effectiveness and accuracy of Scheme 2. Let $T = 1, \kappa = 2$, and $f = \omega_{1+\sigma-\alpha}(t) \sin x \sin y$. The continuous problem has an exact solution $u = \omega_{1+\sigma}(t) \sin x \sin y$ which satisfies the regularity assumptions (1.14).

Table 3 Temporal error of Example 1 for different γ with $\alpha = 0.2, \sigma = 0.8$

N	$\gamma = 3$			$\gamma = 3.5 = \gamma_{opt}$			$\gamma = 4$			
	Alikhanov AccA		Order	Alikhanov AccA		Order	Alikhanov AccA		Order	
	Error	Error		Error	Error		Error	Error		
80	8.24e−07	8.24e−07		8.50e−07	8.50e−07		1.13e−06	1.13e−06		
160	1.56e−07	1.56e−07	2.40	1.36e−07	1.36e−07	2.64	1.83e−07	1.83e−07	2.63	
320	2.96e−08	2.96e−08	2.40	2.13e−08	2.13e−08	2.67	2.88e−08	2.88e−08	2.67	
640	5.61e−09	5.61e−09	2.40	3.26e−09	3.26e−09	2.71	4.45e−09	4.45e−09	2.69	
1280	1.06e−09	1.06e−09	2.40	4.95e−10	4.94e−10	2.72	6.79e−10	6.76e−10	2.72	
min{ $\gamma\sigma, 3 - \alpha$ }			2.40				2.80			

Table 4 Temporal error of Example 1 for different γ with $\alpha = 0.5, \sigma = 0.8$

N	$\gamma = 3$			$\gamma = 3.125 = \gamma_{\text{opt}}$		$\gamma = 3.5$	
	Alikhanov AccA		Alikhanov	AccA	Alikhanov AccA		
	Error	Error	Order Error	Error	Order Error	Error	Order
80	3.91e-06	3.91e-06	4.09e-06	4.09e-06	4.80e-06	4.80e-06	
160	7.41e-07	7.41e-07	2.40	7.55e-07	2.44	8.90e-07	8.90e-07
320	1.40e-07	1.40e-07	2.40	1.37e-07	2.46	1.62e-07	1.62e-07
640	2.66e-08	2.66e-08	2.40	2.48e-08	2.47	2.93e-08	2.93e-08
1280	5.04e-09	5.03e-09	2.40	4.43e-09	2.48	5.26e-09	5.25e-09
$\min\{\gamma\sigma, 3 - \alpha\}$			2.40		2.50		2.50

Define the $L_\infty(H^1)$ semi-norm solution error $e(N, h) = \max_{1 \leq k \leq N} |u^k - u_h^k|_1$. In numerical tests, temporal order and spatial order are investigated by supposing that $e(N, h) \approx C(\tau^p + h^q)$ and evaluating the convergence rates $p \approx \log[e(N, h)/e(2N, h)]$ and $q \approx \log[e(N, h)/e(N, h/2)]$. To check the spatial accuracy, we choose the parameters $\alpha = 0.4, \sigma = 1.6$ and set $N = 1000$ to avoid contamination of the temporal error. Spatial errors and convergence rates for both uniform partition and nonuniform ones are recorded in Table 6. We see that the spatial errors of Schemes 1 and 2 are equal and both two schemes preserve the fourth-order accuracy well. This is consistent with our predicted space accuracy in Corollary 4.1. Table 7 displays the errors and orders in time for different parameters γ with $\alpha = 0.4, \sigma = 1.6$ and a fine mesh size $h = \pi/200$. Temporal errors of Schemes 1 and 2 remain the same which confirms the effectiveness and accuracy of the fully discrete scheme. Unfortunately, numerical results in the third part of Table 7 are different from the first two cases. The convergence rates in this part are higher than 2, maybe can reach $\gamma\sigma$. This amusing phenomenon shows that there are certain differences

Table 5 Temporal error of Example 1 for different γ with $\alpha = 0.8, \sigma = 0.8$

N	$\gamma = 2$			$\gamma = 2.75 = \gamma_{\text{opt}}$		$\gamma = 3$	
	Alikhanov AccA		Alikhanov	AccA	Alikhanov AccA		
	Error	Error	Order Error	Error	Order Error	Error	Order
80	6.64e-05	6.64e-05	2.12e-05	2.12e-05	2.21e-05	2.21e-05	
160	2.19e-05	2.19e-05	1.60	4.82e-06	2.14	4.90e-06	4.90e-06
320	7.23e-06	7.23e-06	1.60	1.08e-06	2.16	1.08e-06	1.08e-06
640	2.38e-06	2.38e-06	1.60	2.39e-07	2.18	2.36e-07	2.36e-07
1280	7.86e-07	7.86e-07	1.60	5.27e-08	2.18	5.16e-08	5.16e-08
$\min\{\gamma\sigma, 3 - \alpha\}$			1.60		2.20		2.20

Table 6 Spatial error of Example 2 for different γ with $\alpha = 0.4, \sigma = 1.6, N = 1000$

h	$\gamma = 1$			$\gamma = 1.25 = \gamma_{opt}$			$\gamma = 1.5$		
	Scheme1		Scheme2	Scheme1		Scheme2	Scheme1		Scheme2
	Error	Error	Order	Error	Error	Order	Error	Error	Order
$\pi/5$	1.04e−03	1.04e−03		1.04e−03	1.04e−03		1.04e−03	1.04e−03	
$\pi/10$	6.40e−05	6.40e−05	4.02	6.40e−05	6.40e−05	4.02	6.40e−05	6.40e−05	4.02
$\pi/20$	4.00e−06	4.00e−06	4.00	3.99e−06	3.99e−06	4.00	3.99e−06	3.99e−06	4.00
$\pi/40$	1.40e−07	2.58e−07	3.95	2.51e−07	2.51e−07	3.99	2.50e−07	2.50e−07	4.00
			4.00			4.00			4.00

between theory and experiment. It seems that the truncation error of $u(t_{n-\theta}) - u^{n-\theta}$ in (4.9) can be further enhanced at least for the coarse meshes.

The left panel of Fig. 1 shows the errors with the parameters $\alpha = 0.5, \sigma = 0.5, h = \pi/10$ at $T = 1$ for different γ . We can conclude that the performance of nonuniform grids is better than the uniform one. Error curves on different nonuniform partitions are indicated in the right panel. These curves illustrate that the fully discrete scheme is unconditionally stable.

Since the computational cost in space had been proved in [29], we just examine the computational cost in time. Taking the parameters $\alpha = 0.6, \sigma = 1.6, \gamma = 1.5, h = \pi/10$, and $T = 1$, we consider the relationship between CPU time and the time-step size N . The left panel of Fig. 2 reports the CPU time in seconds for Schemes 1 and 2 versus N . We find that the CPU time increase via the number of N is linear with the slopes of one and two in the logarithm for Scheme 2 and Scheme 1, respectively. It verifies that the new scheme is much faster than the normal one.

Example 3 Consider the three-dimensional problems (1.1)–(1.3) in $\Omega = (0, \pi)^3$ and the time interval is $(0, 1]$. Let $\kappa = 3$ and $f = \omega_{1+\sigma-\alpha}(t) \sin x \sin y \sin z$, so that

Table 7 Temporal error of Example 2 for different γ with $\alpha = 0.4, \sigma = 1.6, h = \pi/200$

N	$\gamma = 1$			$\gamma = 1.25 = \gamma_{opt}$			$\gamma = 1.5$		
	Scheme1		Scheme2	Scheme1		Scheme2	Scheme1		Scheme2
	Error	Error	Order	Error	Error	Order	Error	Error	Order
10	3.89e−04	3.89e−04		1.78e−04	1.78e−04		1.39e−04	1.39e−04	
20	1.28e−04	1.28e−04	1.60	4.45e−05	4.45e−05	2.00	2.66e−05	2.66e−05	2.39
40	4.23e−05	4.23e−05	1.60	1.11e−05	1.11e−05	2.00	5.05e−06	5.05e−06	2.40
80	1.40e−05	1.40e−05	1.60	2.78e−06	2.78e−06	2.00	9.56e−07	9.56e−07	2.40
$\min\{\gamma\sigma, 2\}$			1.60			2.00			2.00

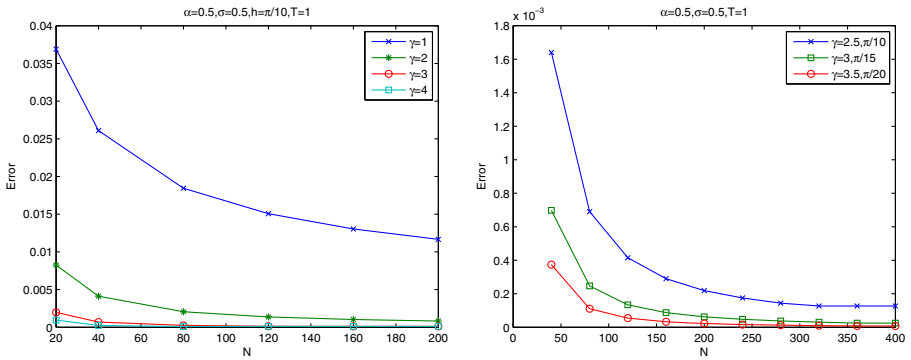


Fig. 1 Error curves of the fully discrete scheme (1.6)–(1.8) for different γ in Example 2 (left); Error curves of the fully discrete scheme (1.6)–(1.8) on nonuniform meshes for different γ and h in Example 2 (right)

the reaction-subdiffusion problem has an exact solution $u = \omega_{1+\sigma}(t) \sin x \sin y \sin z$. This solution still satisfies the regularity assumptions (1.14).

We choose the parameters $\alpha = 0.8, \sigma = 0.8$ and still test the temporal error and spatial error separately. For spatial accuracy, $N = 1000$ is taken as Example 2 to ignore the temporal error. Errors and convergence orders are reported in Table 8 by varying mesh sizes from $h = \pi/5$ to $\pi/40$. For the case of uniform partition, we observe that the errors are almost invariable when the mesh size is refined to $h = \pi/10$. This problem might be caused by too small regularity parameter σ , so that the temporal error contaminates the final data. On the contrary, numerical results of nonuniform partitions are close to the fourth-order accuracy which agrees well with our theoretical prediction. In general, the nonuniform grids outperform the uniform one for the problem with singularity solutions. Always, the same parameters are selected to check the temporal convergence rates in Table 9. We choose $h = \pi/200$ enable to avoid the errors in space and double increase the number of N . Numerical

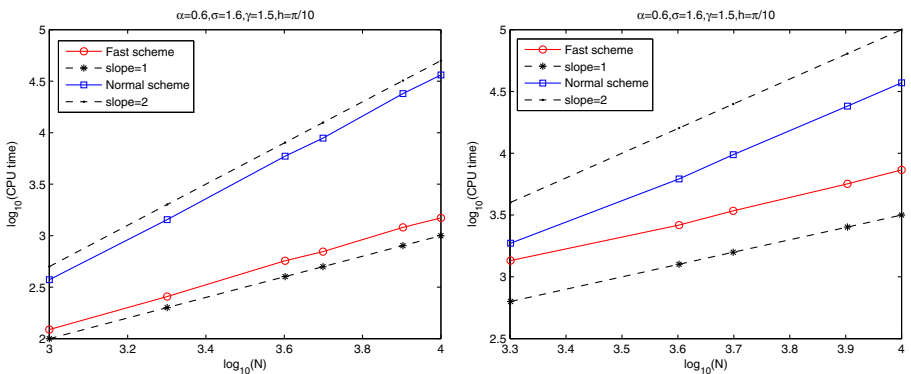


Fig. 2 Comparisons of computational cost between two schemes for 2D problem in Example 2 (left) and 3D problem in Example 3 (right)

Table 8 Spatial error of Example 3 for different γ with $\alpha = 0.8, \sigma = 0.8, N = 1000$

h	$\gamma = 1$			$\gamma = 2.5 = \gamma_{opt}$			$\gamma = 3$		
	Scheme1		Scheme2	Scheme1		Scheme2	Scheme1		Scheme2
	Error	Error	Order	Error	Error	Order	Error	Error	Order
$\pi/5$	2.60e-03	2.60e-03		2.72e-03	2.72e-03		2.72e-03	2.72e-03	
$\pi/10$	4.45e-04	4.45e-04	2.55	1.68e-04	1.68e-04	4.02	1.68e-04	1.68e-04	4.02
$\pi/20$	4.45e-04	4.45e-04	0	1.03e-05	1.03e-05	4.00	1.03e-05	1.03e-05	4.00
$\pi/40$	4.45e-04	4.45e-04	0	4.77e-07	4.77e-07	4.43	4.87e-07	4.87e-07	4.40
			0			4.00			4.00

results in the first two cases maintain the theoretical analysis in Corollary 4.1 for both uniform partition and nonuniform ones. However, the third part in this table arises an interesting phenomenon similar to Example 2, namely, the convergence rate is higher than 2, approaches $3 - \alpha$.

Analogously, comparisons of error curves between two types of meshes and unconditional stability of the proposed scheme are listed in Fig. 3. In the left panel, error curve of the uniform grid is still higher than the other nonuniform ones. On the other side, the performances of error curves on different nonuniform grids illustrate the unconditionally stability of the new scheme.

Comparisons of computational complexity are also reported for 3D case. CPU times with the parameters $\alpha = 0.6, \sigma = 1.6, h = \pi/10$ at $T = 1$ and different numbers of time steps N are shown in the right panel of Fig. 2. The fast scheme has almost linear complexity in N and is much more efficient than the normal one.

In conclusion, compared with Scheme 1, the fully discrete scheme (Scheme 2) not only has the same accuracy but also reduces the computational cost and storage requirement significantly for long time calculation and spatial multi-dimensional cases.

Table 9 Temporal error of Example 3 for different γ with $\alpha = 0.8, \sigma = 0.8, h = \pi/200$

N	$\gamma = 1$			$\gamma = 2.5 = \gamma_{opt}$			$\gamma = 3$		
	Scheme1		Scheme2	Scheme1		Scheme2	Scheme1		Scheme2
	Error	Error	Order	Error	Error	Order	Error	Error	Order
10	1.77e-02	1.77e-02		2.96e-03	2.96e-03		3.08e-03	3.08e-03	
20	1.02e-02	1.02e-02	0.80	7.70e-04	7.70e-04	1.94	7.62e-04	7.62e-04	2.02
40	5.84e-03	5.84e-03	0.80	1.93e-04	1.93e-04	2.00	1.79e-04	1.79e-04	2.09
80	3.35e-03	3.35e-03	0.80	4.82e-05	4.82e-05	2.00	4.09e-05	4.09e-05	2.13
$\min\{\gamma\sigma, 2\}$			0.80			2.00			2.00

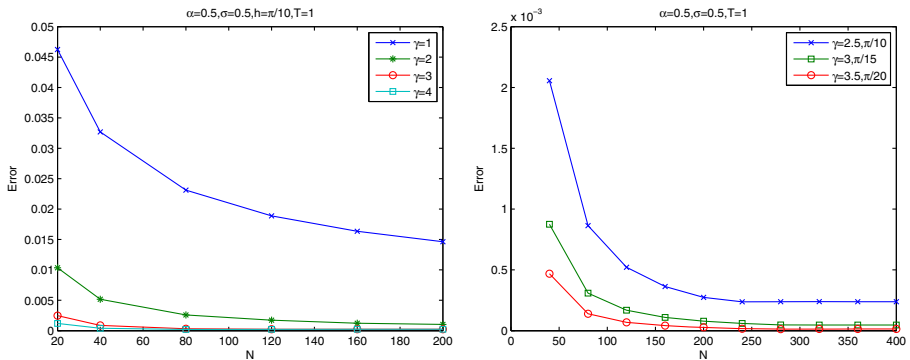


Fig. 3 Error curves of the fully discrete scheme (1.6)–(1.8) for different γ in Example 3 (left); Error curves of the fully discrete scheme (1.6)–(1.8) on nonuniform meshes for different γ and h in Example 3 (right)

6 Concluding remarks

By taking the intrinsically initial singularity of solution into account, an accelerated nonuniform version of Alikhanov formula is presented for approximating the Caputo fractional derivative. In the finite difference framework, we utilize the discrete sine transform and develop a second-order fast compact scheme for solving the subdiffusion problems. Unconditional stability and sharp H^1 -norm error estimate are established by employing the discrete fractional Grönwall inequality and global consistency analysis. Some numerical examples are presented to support the effectiveness of this scheme and the sharpness of our analysis. The theoretical results in time approximation together with their proofs here would be also valid for some other spatial discretizations such as finite element method and spectral method.

Funding information Xin Li is financially supported by a grant KJ2018A0523 from the University Natural Science Research Key Project of Anhui Province. Hong-lin Liao is financially supported by a grant 1008-56SYAH18037 from NUAAscientific Research Starting Fund of Introduced Talent and a grant DRA2015518 from 333 High-level Personal Training Project of Jiangsu Province.

References

1. Hilfer, R.: Applications of Fractional Calculus in Physics. World Scientific, Singapore (2000)
2. Sun, Z.Z., Wu, X.N.: A fully discrete difference scheme for a diffusion-wave system. Appl. Numer. Math. **56**, 193–209 (2006)
3. Lin, Y.M., Xu, C.J.: Finite difference/spectral approximations for the time-fractional diffusion equation. J. Comput. Phys. **225**, 1533–1552 (2007)
4. Li, C.P., Chen, A., Ye, J.J.: Numerical approaches to fractional calculus and fractional ordinary differential equation. J. Comput. Phys. **230**, 3352–3368 (2011)
5. Gao, G.H., Sun, Z.Z., Zhang, H.W.: A new fractional numerical differentiation formula to approximate the Caputo fractional derivative and its applications. J. Comput. Phys. **259**, 33–50 (2014)
6. Alikhanov, A.A.: A new difference scheme for the time fractional diffusion equation. J. Comput. Phys. **280**, 424–438 (2015)
7. Lv, C.W., Xu, C.J.: Error analysis of a high order method for time-fractional diffusion equations. SIAM J. Sci. Comput. **38**, A2699–A2724 (2016)

8. Brunner, H., Ling, L., Yamamoto, M.: Numerical simulations of 2D fractional subdiffusion problems. *J. Comput. Phys.* **229**, 6613–6622 (2010)
9. McLean, W.: Regularity of solutions to a time-fractional diffusion equation. *ANZIAM J.* **52**, 123–138 (2010)
10. Jin, B.T., Lazarov, R., Zhou, Z.: An analysis of the I1 scheme for the subdiffusion equation with nonsmooth data. *IMA J. Numer. Anal.* **36**(1), 197–221 (2016)
11. Sakamoto, K., Yamamoto, M.: Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *J. Math. Anal. Appl.* **382**, 426–447 (2011)
12. 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 J. Numer. Anal.* **55**, 1057–1079 (2017)
13. Yuste, S.B., Quintana-Murillo, J.: A finite difference method with non-uniform timesteps for fractional diffusion equation. *Comput. Phys. Commun.* **183**, 2594–2600 (2012)
14. Mustapha, K., Aïmoutawa, J.: A finite difference method for an anomalous sub-diffusion equation, theory and applications. *Numer. Algo.* **61**, 525–543 (2012)
15. Zhang, Y.N., Sun, Z.Z., Liao, H.-L.: Finite difference methods for the time fractional diffusion equation on nonuniform meshes. *J. Comput. Phys.* **265**, 195–210 (2014)
16. Li, C.P., Yi, Q., Chen, A.: Finite difference methods with non-uniform meshes for nonlinear fractional differential equations. *J. Comput. Appl. Math.* **316**, 614–631 (2016)
17. Liao, H.-L., Li, D.F., Zhang, J.W.: Sharp error estimate of the nonuniform I1 formula for linear reaction-subdiffusion equation. *SIAM J. Numer. Anal.* **56**, 1112–1133 (2018)
18. Liao, H.-L., McLean, W., Zhang, J.W.: A discrete grönwall inequality with application to numerical schemes for subdiffusion problems. *SIAM J. Numer. Anal.* **57**, 218–237 (2019)
19. Liao, H.-L., McLean, W., Zhang, J.W.: A second-order scheme with nonuniform time steps for a linear reaction-subdiffusion problem. *arXiv:1803.09873v2* (2018)
20. Ren, J.C., Liao, H.-L., Zhang, J.W., Zhang, Z.M.: Sharp h^1 -norm error estimates of two time-stepping schemes for reaction-subdiffusion problems. *arXiv:1811.08059v1* (2018)
21. Ke, R.H., Ng, M.K., Sun, H.W.: A fast direct method for block triangular Toeplitz-like with tri-diagonal block systems from time-fractional partial differential equation. *J. Comput. Phys.* **303**, 203–211 (2015)
22. Baffet, D., Hesthaven, J.S.: A kernel compression scheme for fractional differential equations. *SIAM J. Numer. Anal.* **55**, 496–520 (2017)
23. Jiang, S.D., Zhang, J.W., Zhang, Q., Zhang, Z.M.: Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun. Comput. Phys.* **21**, 650–678 (2017)
24. Yan, Y.G., Sun, Z.Z., Zhang, J.W.: Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations: a second-order scheme. *Commun. Comput. Phys.* **22**, 1028–1048 (2017)
25. Shen, J.Y., Sun, Z.Z., Du, R.: Fast finite difference schemes for time-fractional diffusion equations with a weak singularity at initial time. *East Asia J. Appl. Math.* **8**, 834–858 (2018)
26. Liao, H.-L., Yan, Y.G., Zhang, J.W.: Unconditional convergence of a fast two-level linearized algorithm for semilinear subdiffusion equations. *J. Sci. Comput.* **80**, 1–25 (2019)
27. Wang, H.Q., Zhang, Y., Ma, X., Qiu, J., Liang, Y.: An efficient implementation of fourth-order compact finite difference scheme for Poisson equation with Dirichlet boundary conditions. *Commun. Math. Appl.* **71**, 1843–1860 (2016)
28. Wang, H.Q., Ma, X., Lu, J.L., Cao, W.: An efficient time-splitting compact finite difference method for Gross-Pitaevskii equation. *Appl. Math. Comput.* **297**, 131–144 (2017)
29. Shen, J., Tang, T., Wang, L.L.: *Spectral Methods: Algorithms, Analysis and Applications*. Springer, Berlin (2011)

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.