Fully Finite Element Adaptive Algebraic Multigrid Method for Time-Space Caputo-Riesz Fractional Diffusion Equations

Xiaoqiang Yue\textsuperscript{a}, Weiping Bu\textsuperscript{a}, Shi Shu\textsuperscript{b,*}, Menghuan Liu\textsuperscript{a}, Shuai Wang\textsuperscript{a}

\textsuperscript{a}School of Mathematics and Computational Science, Xiangtan University, Hunan 411105, P.R. China
\textsuperscript{b}Hunan Key Laboratory for Computation and Simulation in Science and Engineering, Xiangtan University, Hunan 411105, P.R. China

Abstract

The paper aims to establish a fully discrete finite element (FE) scheme and provide cost-effective solutions for one-dimensional time-space Caputo-Riesz fractional diffusion equations on a bounded domain $\Omega$. Firstly, we construct a fully discrete scheme of the linear FE method in both temporal and spatial directions, derive many characterizations on the coefficient matrix and numerically verify that the fully FE approximation possesses the saturation error order under $L^2(\Omega)$ norm. Secondly, we theoretically prove the estimation $1 + \mathcal{O}(\tau^\alpha h^{-2\beta})$ on the condition number of the coefficient matrix, in which $\tau$ and $h$ respectively denote time and space step sizes. Finally, on the grounds of the estimation and fast Fourier transform, we develop and analyze an adaptive algebraic multigrid (AMG) method with low algorithmic complexity, reveal a reference formula to measure the strength-of-connection tolerance which severely affect the robustness of AMG methods in handling fractional diffusion equations, and illustrate the well robustness and high efficiency of the proposed algorithm compared with the classical AMG, conjugate gradient and Jacobi iterative methods.

Keywords: Caputo-Riesz fractional diffusion equation, fully time-space FE scheme, condition number estimation, algorithmic complexity, adaptive AMG method

2010 MSC: 35R11, 65F10, 65F15, 65N55

1. Introduction

In recent years, there has been an explosion of research interest in numerical solutions for fractional differential equations, mainly due to the following two aspects: (i) the huge majority can’t be solved analytically, (ii) the analytical solution (if luckily derived) always involve certain infinite
series which sharply drives up the costs of its evaluation. Various numerical methods have been proposed to approximate more accurately and faster, such as finite difference (FD) method \[1–10\], finite element (FE) method \[11–17\], finite volume \[18\] method and spectral (element) method \[19–25\]. An essential challenge against standard differential equations lies in the presence of the fractional differential operator, which gives rise to nonlocality (space fractional, nearly dense or full coefficient matrix) or memory-requirement (time fractional, the entire time history of evaluations) issue, resulting in a vast computational cost.

Preconditioned Krylov subspace methods are regarded as one of the potential solutions to the aforementioned challenge. Numerous preconditioners with various Krylov-subspace methods have been constructed respectively for one- and two-dimensional, linear and nonlinear space-fractional diffusion equations (SFDE) \[26–30\]. Multigrid method has been proven to be a superior solver and preconditioner for ill-conditioned Toeplitz systems as well as SFDE. Pang and Sun propose an efficient and robust geometric multigrid (GMG) with fast Fourier transform (FFT) for one-dimensional SFDE by an implicit FD scheme \[31\]. Bu et al. employ the GMG to one-dimensional multi-term time-fractional advection-diffusion equations via a fully discrete scheme by FD method in temporal and FE method in spatial directions \[32\]. Jiang and Xu construct optimal GMG for two-dimensional SFDE to get FE approximations \[33\]. Chen et al. make the first attempt to present an algebraic multigrid (AMG) method with line smoothers to the fractional Laplacian through localizing it into a nonuniform elliptic equation \[34\]. Zhao et al. invoke GMG for one-dimensional Riesz SFDE by an adaptive FE scheme using hierarchical matrices \[35\]. From the survey of references, in spite of quite a number of contributions to numerical methods and preconditioners, there are no calculations taking into account of fully discrete FE schemes and AMG methods for time-space Caputo-Riesz fractional diffusion equations.

In this paper, we are concerned with the following time-space Caputo-Riesz fractional diffusion equation (CR-FDE)

\[
\frac{\mathcal{D}_t^\alpha}{\mathcal{D}_t^\beta} u(x,t) = \frac{\partial^{2\beta}}{\partial |x|^{2\beta}} u(x,t) + f(x,t), \quad t \in (0, T], \quad x \in \Omega = (a, b)
\]

(1)

\[
u(x,t) = 0, \quad t \in I, \quad x \in \partial \Omega
\]

(2)

\[
u(x,0) = \psi_0(x), \quad x \in \Omega
\]

(3)

with orders $\alpha \in (0, 1)$ and $\beta \in (1/2, 1)$, the Caputo and Riesz fractional derivatives are respectively
defined by
\[ C_0 D_t^\alpha u = \frac{1}{\Gamma(1 - \alpha)} \int_0^t (t - s)^{-\alpha} \frac{\partial u}{\partial s} \, ds, \quad \partial_x^{2\beta} u = -\frac{1}{2 \cos(\beta \pi)} (x D_L^{2\beta} u + x D_R^{2\beta} u), \]
where
\[ x D_L^{2\beta} u = \frac{1}{\Gamma(2 - 2\beta)} \frac{\partial^2}{\partial x^2} \int_a^x (s - x)^{1 - 2\beta} u \, ds, \quad x D_R^{2\beta} u = \frac{1}{\Gamma(2 - 2\beta)} \frac{\partial^2}{\partial x^2} \int_x^b (s - x)^{1 - 2\beta} u \, ds. \]

The remainder of this paper proceeds as follows. A fully discrete FE method of (1)-(3) is developed in Section 2. Section 3 comes up with the theoretical estimation and verification experiments on the condition number of the coefficient matrix. The classical AMG method is introduced in Section 4 followed by its uniform convergence analysis and the construction of an adaptive AMG method. Section 5 reports and analyzes numerical results to show the benefits. We close in Section 6 with some concluding remarks.

2. Fully discrete finite element scheme for the CR-FDE

For simplicity, following [36], we will use the symbols \( \lesssim \), \( \gtrsim \), and \( \simeq \) throughout the paper. \( u_1 \lesssim v_1 \) means \( u_1 \leq C_1 v_1 \), \( u_2 \gtrsim v_2 \) means \( u_2 \geq c_2 v_2 \) while \( u_3 \simeq v_3 \) means \( c_3 v_3 \leq u_3 \leq C_3 v_3 \), where \( C_1, c_2, c_3 \) and \( C_3 \) are generic positive constants independent of variables, time and space step sizes.

2.1. Reminder about fractional calculus

In this subsection, we briefly introduce some fractional derivative spaces and several auxiliary results. Here the \( L^2 \) inner product and norm are denoted by
\[ (u, v)_{L^2(\Omega)} = \int_\Omega uv \, dx, \quad \| u \|_{L^2(\Omega)} = (u, u)_{L^2(\Omega)}^{1/2}. \]

Definition 1. (Left and right fractional derivative spaces) For constant \( \mu > 0 \), define norms
\[ \| u \|_{J_\mu^L(\Omega)} := (\| u \|_{L^2(\Omega)}^2 + \| x D_L^{\mu} u \|_{L^2(\Omega)}^2)^{1/2}, \quad \| u \|_{J_\mu^R(\Omega)} := (\| u \|_{L^2(\Omega)}^2 + \| x D_R^{\mu} u \|_{L^2(\Omega)}^2)^{1/2}, \]
and let \( J_{\mu,0}^L(\Omega) \) and \( J_{\mu,0}^R(\Omega) \) be closures of \( C_0^\infty(\Omega) \) under \( \| \cdot \|_{J^L_\mu(\Omega)} \) and \( \| \cdot \|_{J^R_\mu(\Omega)} \), respectively.

Definition 2. (Fractional Sobolev space) For constant \( \mu > 0 \), define the norm
\[ \| u \|_{H^\mu(\Omega)} := (\| u \|_{L^2(\Omega)}^2 + \| \xi^\mu \hat{u} \|_{L^2(\xi^\mu)}^2)^{1/2}, \]
and let \( H_{0}^\mu(\Omega) \) be the closure of \( C_0^\infty(\Omega) \) under \( \| \cdot \|_{H^\mu(\Omega)} \), where \( \hat{u} \) is the Fourier transform of \( u \).
Remark 1. The equivalence between (4) and the general definition of the norm has been established in [14], which implies the reasonability of Definition 2.

Lemma 1. (see [12], Proposition 1) If constant \( \mu \in (0, 1) \), \( u, v \in J_{L,0}^{2\mu}(\Omega) \) (or \( J_{R,0}^{2\mu}(\Omega) \)), then
\[
(x D_L^{2\mu} u, v)_{L^2(\Omega)} = (x D_L^{2\mu} u, x D_R^{2\mu} v)_{L^2(\Omega)}, \quad (x D_R^{2\mu} u, v)_{L^2(\Omega)} = (x D_R^{2\mu} u, x D_L^{2\mu} v)_{L^2(\Omega)}.
\]

Lemma 2. (see [11], Lemma 2.4) For constant \( \mu > 0 \), we have
\[
(x D_L^{\mu} u, x D_R^{\mu} u)_{L^2(\Omega)} = \cos(\pi \mu) \| x D_L^{\mu} u \|_{L^2(\Omega)}^2.
\]

Lemma 3. (Fractional Poincaré-Friedrichs inequality, see [11], Theorem 2.10) For \( u \in J_{L,0}^{\mu}(\Omega) \), we have
\[
\| u \|_{L^2(\Omega)} \lesssim \| x D_L^{\mu} u \|_{L^2(\Omega)}.
\]

2.2. Derivation of the fully discrete scheme

By Lemma 1, we get the variational (weak) formulation of (1)-(3): given \( f \in L^2(\Omega, I) \), \( \phi_0 \in L^2(\Omega) \), and \( Q_t := \Omega \times (0, t) \), to find \( u \in H \) subject to \( u(x, 0) = \psi_0(x) \) and
\[
\left( C_0 D_0^{\alpha} u, v \right)_{Q_t} + B_{Q_t}(u, v) = (f, v)_{Q_t}, \quad \forall v \in H^*,
\]
where \( H := H_0^\beta(\Omega) \times H^1(I) \), \( H^* := H_0^\beta(\Omega) \times L^2(I) \), and
\[
\begin{align*}
\left( C_0 D_0^{\alpha} u, v \right)_{Q_t} &= \int_0^t \left( C_0 D_0^{\alpha} u, v \right)_{L^2(\Omega)} \, d\sigma, \\
B_{Q_t}(u, v) &= \int_0^t \frac{1}{2 \cos(\beta \pi)} \left[ (x D_L^{\beta} u, x D_R^{\beta} v)_{L^2(\Omega)} + (x D_R^{\beta} u, x D_L^{\beta} v)_{L^2(\Omega)} \right] \, d\sigma.
\end{align*}
\]

In order to acquire numerical solutions of \( u \), we firstly make a (possibly nonuniform) temporal discretization by points \( 0 = t_0 < t_1 < \cdots < t_N = T \), and a uniform spatial discretization by points \( x_i = a + ih \) \((i = 0, 1, \cdots, M)\), where \( h = (b - a)/M \) represents the space step size. Let
\[
I_j = (t_{j-1}, t_j), \quad \tilde{I}_j = (0, t_j), \quad j = 1, 2, \cdots, N; \quad \Omega_h = \{ \Omega_l : \Omega_l = (x_{l-1}, x_l), \ l = 1, 2, \cdots, M \}.
\]

We observe that it is convenient to form the FE spaces in tensor products
\[
\mathcal{V}_n = \mathcal{V}_h^\beta(\Omega_h) \times \mathcal{V}_r(\tilde{I}_n), \quad \mathcal{V}_n^* = \mathcal{V}_h^\beta(\Omega_h) \times \mathcal{V}_r^*(I_n),
\]
where
\[
\mathcal{V}_h^b(\Omega_h) = \{ w_h \in H_0^b(\Omega) \cap C(\Omega) : w_h(x)|_{\Omega_h} \in \mathcal{P}_l(\Omega), \ l = 1, \cdots, M \},
\]
\[
\mathcal{V}_n(I_n) = \{ v_n \in C(I_n) : v_n(0) = 1, \ v_n(t)|_{I_j} \in \mathcal{P}_l(I_j), \ j = 1, \cdots, n \},
\]
\[
\mathcal{V}_n^b(I_n) = \{ v_n \in L^2(I_n) : v_n(t)|_{I_n} \in \mathcal{P}_0(I_n) \},
\]
and \( \mathcal{P}_k \) denotes the set of all polynomials of degree \( \leq k \).

**Remark 2.** Apparently, for a given \( u_{h,\tau}(x,t) \in \mathcal{V}_n \), we have \( \partial u_{h,\tau}/\partial t \in \mathcal{V}_n^* \), where \( \partial u_{h,\tau}/\partial t \) is obtained by differentiating \( u_{h,\tau} \) with respect to \( t \) on each subinterval \( I_j \) \( (j = 1, 2, \cdots, N) \).

We obtain a fully discrete FE scheme in temporal and spatial directions of problem (7): given \( Q_n := \Omega_h \times I_n \), to find \( u_{h,\tau} \in \mathcal{V}_n \) such that \( u_{h,\tau}(x,0) = \psi_{0,l}(x) \) and
\[
\left( \frac{\partial}{\partial t} u_{h,\tau}, v_{h,\tau} \right)_{Q_n} + B_{\Omega}^n(u_{h,\tau}, v_{h,\tau}) = (f, v_{h,\tau})_{Q_n}, \ \forall v_{h,\tau} \in \mathcal{V}_n^*,
\]
where \( \psi_{0,l}(x) \in \mathcal{V}_n \) satisfying \( \psi_{0,l}(x_i) = \psi_0(x_i) \ (i = 0, 1, \cdots, M) \), and
\[
B_{\Omega}^n(u_{h,\tau}, v_{h,\tau}) = \int_{t_{n-1}}^{t_n} \frac{1}{2 \cos(\beta \pi)} \left[ (x D_L^0 u_{h,\tau}, x D_L^0 v_{h,\tau})_{L^2(\Omega)} + (x D_R^0 u_{h,\tau}, x D_R^0 v_{h,\tau})_{L^2(\Omega)} \right] dt.
\]

Let
\[
L_0(t) = \begin{cases} \frac{t_1 - t}{\tau_1}, & t \in I_1 \\ 0, & t \in \tilde{I}_n \setminus I_1 \end{cases}, \quad \tilde{L}_0(t) = \frac{1}{\Gamma(1 - \alpha)} \int_{t_0}^{t} \frac{dL_0(s)}{(t - s)^\alpha}, \quad \hat{L}_0(t) = \frac{1}{\Gamma(1 - \alpha)} \int_{t_0}^{t_1} \frac{dL_0(s)}{(t - s)^\alpha},
\]
\[
L_k(t) = \begin{cases} \frac{t_{k+1} - t}{\tau_{k+1}}, & t \in I_{k+1} \\ \frac{t - t_{k-1}}{\tau_k}, & t \in I_k \\ 0, & t \in \tilde{I}_n \setminus (I_k \cup I_{k+1}) \end{cases}, \quad \tilde{L}_k(t) = \frac{1}{\Gamma(1 - \alpha)} \int_{t_{k-1}}^{t_{k+1}} \frac{dL_k(s)}{(t - s)^\alpha}, \quad k = 1, \cdots, n - 1
\]
and
\[
L_n(t) = \begin{cases} \frac{t - t_{n-1}}{\tau_n}, & t \in I_n \\ 0, & t \in \tilde{I}_n \setminus I_n \end{cases}, \quad \tilde{L}_n(t) = \frac{1}{\Gamma(1 - \alpha)} \int_{t_{n-1}}^{t} \frac{dL_n(s)}{(t - s)^\alpha},
\]
Note that
\[
\mathcal{V}_n^* = \mathcal{S}_{\text{span}}\{ \phi_l(x) \times 1, \ l = 1, \cdots, M - 1 \}.
\]
where $\phi_l(x)$ is the shape function at $x_l \in \Omega_h$. Using

$$u_{h\tau}(x, t) = u_h^0(x)\mathcal{L}_0(t) + \sum_{k=1}^{n-1} u_h^k(x)\mathcal{L}_k(t) + u_h^n(x)\mathcal{L}_n(t),$$

we have

$$\begin{align*}
(C D^\alpha_t u_{h\tau}, \phi_l \times 1)_{Q_h} &= (u_h^0, \phi_l)_{L^2(\Omega)}(\mathcal{L}_0, 1)_{L^2(I_1)} + (u_h^1, \phi_l)_{L^2(\Omega)}(\mathcal{L}_1, 1)_{L^2(I_1)}, \\
(C D^\alpha_t u_{h\tau}, \phi_l \times 1)_{Q_n} &= (u_h^n, \phi_l)_{L^2(\Omega)}(\mathcal{L}_0, 1)_{L^2(I_n)} + \sum_{k=1}^{n-1} (u_h^k, \phi_l)_{L^2(\Omega)}(\mathcal{L}_k, 1)_{L^2(I_n)}, n > 1,
\end{align*}$$

(9)

$$\begin{align*}
\int_{t_{n-1}}^{t_n} 1 \times (x D_L^\beta u_{h\tau}, x D_R^\beta \phi_l)_{L^2(\Omega)}dt &= (x D_L^\beta u_h^0, x D_R^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_0, 1)_{L^2(I_n)} + \\
&\sum_{k=1}^{n-1} (x D_L^\beta u_h^k, x D_R^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_k, 1)_{L^2(I_n)} + (x D_L^\beta u_h^n, x D_R^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_n, 1)_{L^2(I_n)},
\end{align*}$$

(10)

$$\begin{align*}
\int_{t_{n-1}}^{t_n} 1 \times (x D_R^\beta u_{h\tau}, x D_L^\beta \phi_l)_{L^2(\Omega)}dt &= (x D_R^\beta u_h^0, x D_L^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_0, 1)_{L^2(I_n)} + \\
&\sum_{k=1}^{n-1} (x D_R^\beta u_h^k, x D_L^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_k, 1)_{L^2(I_n)} + (x D_R^\beta u_h^n, x D_L^\beta \phi_l)_{L^2(\Omega)}(\mathcal{L}_n, 1)_{L^2(I_n)}.
\end{align*}$$

(11)

Substituting (9)-(12) into (8), yields

$$C_{h\tau}^n U_{h\tau}^n = G_{h\tau}^n,$$

(13)

where the coefficient matrix

$$C_{h\tau}^n = M_h + \frac{\Gamma(3 - \alpha)}{2} \tau_n^\alpha A_h^\beta,$$

(14)

the right-hand side vector

$$G_{h\tau}^n = \Gamma(3 - \alpha) \tau_n^{\alpha - 1} F_h^n + \left[ M_h - \frac{\Gamma(3 - \alpha)}{2} \tau_n^\alpha A_h^\beta \right] U_{h\tau}^{n-1} - \sum_{k=1}^{n-1} \tau_n^{\alpha - 1} \times
\begin{align*}
(t_n - t_{k-1})^{2-\alpha} - (t_{n-1} - t_{k-1})^{2-\alpha} - (t_n - t_k)^{2-\alpha} + (t_{n-1} - t_k)^{2-\alpha} &M_h(U_{h\tau}^k - U_{h\tau}^{k-1}),
\end{align*}$$

the mass matrix

$$M_h = \frac{h}{6} \begin{pmatrix}
4 & 1 \\
1 & 4 & 1 \\
\vdots & \ddots & \ddots & \ddots \\
1 & 4 & 1 \\
1 & 4 & & & & \cdots & & \cdots \\
& & & & & & \cdots & & \cdots & \ddots & \ddots & \ddots & 1 & 4 \\
\end{pmatrix}_{(M-1) \times (M-1)},$$

(15)
Theorem 1. The stiffness matrix $A_h^\beta = (a_{i,j}^h)_{(M-1) \times (M-1)}$ with its entries

\[
\begin{align*}
da_{i,i}^h &= \frac{h^{1-2\beta}(2^{4-2\beta} - 8)}{2\cos(\beta\pi)\Gamma(4-2\beta)}, \quad i = 1, \ldots, M-1 \\
da_{j,j+1}^h &= \frac{h^{1-2\beta}(3^{3-2\beta} - 2^{5-2\beta} + 7)}{2\cos(\beta\pi)\Gamma(4-2\beta)}, \quad j = 1, \ldots, M-2 \\
da_{k,k+l}^h &= \frac{h^{1-2\beta}}{2\cos(\beta\pi)\Gamma(4-2\beta)} [(l+2)^{3-2\beta} \\
&- 4(l+1)^{3-2\beta} + 6l^{3-2\beta} - 4(l-1)^{3-2\beta} + (l-2)^{3-2\beta}], \quad k = 1, \ldots, M-l-1
d\end{align*}
\]

the vector

\[F_n^h = (f_1^n, f_2^n, \ldots, f_{M-1}^n)^T, \quad f_l^n = (f, \phi_l \times 1)Q_n, \quad l = 1, \ldots, M-1\]

and the fully FE approximations

\[U_{hr}^k = (u_1^k, u_2^k, \ldots, u_M^{k-1})^T, \quad u_j^k = \psi_0, I(x_j), \quad u_j^l = u_k^l(x_j), \quad k = 1, \ldots, n, \quad j = 1, \ldots, M-1.\]

Remark 3. \(A_h^\beta\) is reduced via dividing both sides of (18) by the factor \(\tau_n^{1-\alpha}/\Gamma(3-\alpha)\), in case of the severe loss in convergence of the fully discrete FE scheme.

Next, a number of characterizations are established regarding $A_h^\beta$ just defined by (16).

Theorem 1. The stiffness matrix $A_h^\beta$ is symmetric and satisfies

1. $a_{i,i}^h > 0$ for $i = 1, \ldots, M-1$;
2. $a_{i,j}^h < 0$ for $i \neq j, i, j = 1, \ldots, M-1$;
3. $\sum_{j=1}^{M-1} a_{i,j}^h > 0$ for $i = 1, \ldots, M-1$;
4. The following relation holds for the particular case when $h \leq 1/7$

\[
\sum_{j=1}^{M-1} a_{i,j}^h \geq \begin{cases} 
- \frac{h^{1-2\beta}(4 - 2^{3-2\beta})}{2\cos(\beta\pi)\Gamma(4-2\beta)}, & i = 1, M-1 \\
- \frac{2^{2\beta}h(2\beta - 1)}{\cos(\beta\pi)\Gamma(2-2\beta)}, & i = 2, \ldots, M-2
\end{cases}
\]

5. $A_h^\beta$ is an $M$-matrix.

Proof. The symmetric property of $A_h^\beta$ is an obvious fact by (16). Since $\beta \in (1/2, 1)$, then $4 - 2\beta < 3$ and $\cos(\beta\pi) < 0$, which give immediately $a_{i,i}^h > 0, i = 1, \ldots, M-1$. This proves the first part of the theorem. The second part is an immediate consequence of the facts that on the interval $\beta \in (1/2, 1)$, $f(\beta) = 3^{3-2\beta} - 2^{5-2\beta} + 7$ is a strictly increasing function, and the bivariate function

\[f_\beta(l) = (l+2)^{3-2\beta} - 4(l+1)^{3-2\beta} + 6l^{3-2\beta} - 4(l-1)^{3-2\beta} + (l-2)^{3-2\beta} > 0, \quad 2 \leq l \leq M-2.\]
In fact, it is evident that
\[
1.5^{2\beta} > 1.5 > \frac{3^3 \ln 3}{2^5 \ln 2} \Rightarrow f'(\beta) = -2 \ln 3 \cdot 3^{3-2\beta} + 2 \ln 2 \cdot 2^{5-2\beta} > 0,
\]
and
\[
f_\beta(l) = h^{2\beta-3}(g(x_{i+2}) - 4g(x_{i+1}) + 6g(x_i) - 4g(x_{i-1}) + g(x_{i-2})) > 0
\]
using Taylor’s expansion with
\[
\left(\frac{l}{l+1}\right)^{2+2\beta} - \left(\frac{l}{l-1}\right)^{2+2\beta} > -\frac{30l}{2\beta+1} = l^{-1-2\beta} + \frac{2\beta + 1}{30} \[(l + 1)^{-2-2\beta} - (l - 1)^{-2-2\beta}] > 0,
\]
where \( g(x) = (x-a)^{3-2\beta} \) and \( x_l = lh + a \).

To prove the third part, use
\[
(xD_L^\beta \phi_i, xD_R^\beta \phi_j)_{L^2(\Omega)} = -(xD_L^{2\beta-1} \phi_i, d\phi_j/dx)_{L^2(\Omega)}
\]
and
\[
\tilde{\phi} := \sum_{j=1}^{M-1} \phi_j = 1 - \phi_0 - \phi_M
\]
to obtain the relation
\[
\sum_{j=1}^{M-1} \phi_i^h = -\left((xD_L^{2\beta-1} \phi, 1)_{\Omega_i} -(xD_L^{2\beta-1} \phi, 1)_{\Omega_{i+1}} + (xD_L^{2\beta-1} \phi_i, 1)_{\Omega_i} -(xD_L^{2\beta-1} \phi_i, 1)_{\Omega_Mh}\right) / 2h \cos(\beta\pi),
\]
where
\[
xD_L^{2\beta-1} \phi = \begin{cases} \frac{(x-a)^{2-2\beta}}{h\Gamma(3-2\beta)}, & a < x < x_1 \\ \frac{(x-a)^{2-2\beta} - (x-x_1)^{2-2\beta}}{h\Gamma(3-2\beta)}, & x_1 < x < x_{M-1} \\ \frac{(x-a)^{2-2\beta} - (x-x_1)^{2-2\beta} - (x-x_{M-1})^{2-2\beta}}{h\Gamma(3-2\beta)}, & x_{M-1} < x < x_M \end{cases},
\]
and
\[
xD_L^{2\beta-1} \phi_i(x) = \begin{cases} 0, & x < x_{i-1} \\ \frac{(x-x_{i-1})^{2-2\beta}}{h\Gamma(3-2\beta)}, & x_{i-1} < x < x_i \\ \frac{(x-x_{i-1})^{2-2\beta} - 2(x-x_i)^{2-2\beta}}{h\Gamma(3-2\beta)}, & x_i < x < x_{i+1} \\ \frac{(x-x_{i-1})^{2-2\beta} - 2(x-x_i)^{2-2\beta} + (x-x_{i+1})^{2-2\beta}}{h\Gamma(3-2\beta)}, & x > x_{i+1} \end{cases}.
\]
Assume that $\Omega = (0, 1)$ without loss of generality, one can easily derive

\[
\sum_{j=1}^{M-1} a_{i,j}^h = \frac{(4 - 2^3 - 2^\beta) h^3 - 2\beta - 1 + 3(1 - h) h^{3 - 2\beta} - 3(1 - 2h) h^{3 - 2\beta} + (1 - 3h) h^{3 - 2\beta}}{2 \cos(\beta \pi) h^2 \Gamma(4 - 2\beta)}, \quad i = 1, M - 1,
\]

\[
\sum_{j=1}^{M-1} a_{i,j}^h = \frac{3(ih) h^{3 - 2\beta} - 3[(i - 1)h]^3 - 2\beta + [(i - 2)h]^3 - 2\beta - [(i + 1)h]^3 - 2\beta}{2 \cos(\beta \pi) h^2 \Gamma(4 - 2\beta)} - \frac{3(1 - ih) h^{3 - 2\beta} - 3[1 - (i - 1)h] h^{3 - 2\beta} - 3[1 - (i + 1)h] h^{3 - 2\beta} + 1 - (i + 2)h] h^{3 - 2\beta}}{2 \cos(\beta \pi) h^2 \Gamma(4 - 2\beta)}, \quad i = 2, \ldots, M - 2
\]

and deduce $\sum_{j=1}^{M-1} a_{i,j}^h > 0$ by Taylor’s formula and $\beta \in (1/2, 1)$.

Another step to do in the proof is the result 4, which follows from

\[
h \leq 1/7 \Rightarrow 7 - 4(1 - \xi)^{-1 - 2\beta} > \frac{1}{\beta h} \Rightarrow -1 + 3(1 - h) h^{3 - 2\beta} - 3(1 - 2h) h^{3 - 2\beta} + (1 - 3h) h^{3 - 2\beta} > 0
\]

for all $\xi \in (0, 2h)$,

\[
(i - \frac{1}{i})^{3 + 2\beta} > \frac{\beta + 1}{21(\beta + 1) + 30i} \Rightarrow \frac{3(ih) h^{3 - 2\beta} - 3(ih - h) h^{3 - 2\beta} + (ih - 2h) h^{3 - 2\beta} - (ih + h) h^{3 - 2\beta}}{h^3(3 - 2\beta)(2 - 2\beta)(2\beta - 1)(ih)^{-2\beta}} > 1
\]

and

\[
\frac{3(1 - ih) h^{3 - 2\beta} - [1 - (i - 1)h] h^{3 - 2\beta} - 3[1 - (i + 1)h] h^{3 - 2\beta} + 1 - (i + 2)h] h^{3 - 2\beta}}{h^3(3 - 2\beta)(2 - 2\beta)(2\beta - 1)(1 - ih)^{-2\beta}} > 1
\]

for $i = 2, \ldots, M - 2$, together with the inequality $(ih)^{-2\beta} + (1 - ih)^{-2\beta} \geq 2^{1+2\beta}$.

Finally, according to properties 1 and 2, the result 5 will be proved by showing that $(A_{h}^{n})^{-1}$ is nonnegative, which can be easily proved by contradiction with property 3.

Observe from (13)-(16) that $M_h$ and $A_{h}^{n}$ are both symmetric Toeplitz matrices independent of any time terms. The under-mentioned corollaries are natural consequences of Theorem 1.

**Corollary 1.** The coefficient matrix $C_{h,n}^{a}$ is a symmetric Toeplitz matrix. Furthermore, it will be independent of time level $n$ if the temporal discretization is also uniform.

**Corollary 2.** The coefficient matrix $C_{h,T}^{a}$ is an $M$-matrix, if and only if

\[
\frac{\tau_n^a}{h^{2\beta}} > -\frac{2 \cos(\beta \pi) \Gamma(4 - 2\beta)}{3 \Gamma(3 - \alpha)(3^2 - 2\beta - 2^\alpha + 7)}.
\]

**Proof.** This result will follow from Theorem 1 if we can show that

\[
\frac{h}{6} + \frac{\Gamma(3 - \alpha)}{2} \frac{h^{1-2\beta}(3^2 - 2\beta - 2^\alpha + 7)}{2 \cos(\beta \pi) \Gamma(4 - 2\beta)} < 0,
\]

which is an immediate application of the condition (17).
2.3. Numerical experiments and the saturation error order

Example 1. Consider (13) with $\Omega = (0, 1)$, $T = 1$, $\psi_0(x) = 0$ and

$$f(x, t) = \frac{\Gamma(3 - \alpha)}{\Gamma(3 - 2\alpha)} t^{2-2\alpha} x^2 (1 - x)^2 + \frac{t^{2-\alpha}}{\cos(\beta \pi)} \left[ \frac{x^{2-2\beta} + (1-x)^{2-2\beta}}{\Gamma(3 - 2\beta)} - \frac{6x^{3-2\beta} + 6(1-x)^{3-2\beta}}{\Gamma(4 - 2\beta)} \right].$$

The exact solution is $u(x, t) = t^{2-\alpha} x^2 (1 - x)^2$. In the case of uniform temporal and spatial meshes, Tables 1 and 2 present errors $\|u\|_0 := \|u(\cdot), 1 - u_{h, T}(\cdot, 1)\|_{L^2(\Omega)}$ and convergence rates.

| $N$ | $\beta = 0.6$ | $\beta = 0.8$ |
|-----|----------------|----------------|
|     | $\alpha = 0.01$ | $\alpha = 0.01$ | $\alpha = 0.01$ |
|     | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ |
|     | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate |
| 8   | 8.94E-4 | - | 8.78E-4 | - | 8.88E-4 | - | 9.73E-4 | - | 9.56E-4 | - | 9.72E-4 | - |
| 16  | 2.02E-4 | 2.15 | 1.98E-4 | 2.15 | 2.00E-4 | 2.15 | 2.34E-4 | 2.06 | 2.29E-4 | 2.06 | 2.33E-4 | 2.06 |
| 64  | 1.01E-5 | 2.15 | 1.00E-5 | 2.14 | 1.01E-5 | 2.14 | 1.30E-5 | 2.09 | 1.27E-5 | 2.09 | 1.29E-5 | 2.09 |

| $N$ | $\beta = 0.6$ | $\beta = 0.8$ |
|-----|----------------|----------------|
|     | $\alpha = 0.10$ | $\alpha = 0.10$ | $\alpha = 0.10$ |
|     | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ |
|     | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate |
| 8   | 8.90E-4 | - | 8.80E-4 | - | 8.83E-4 | - | 9.72E-4 | - | 9.63E-4 | - | 9.64E-4 | - |
| 16  | 1.99E-4 | 2.16 | 1.97E-4 | 2.16 | 1.99E-4 | 2.15 | 2.31E-4 | 2.07 | 2.28E-4 | 2.08 | 2.31E-4 | 2.06 |
| 64  | 4.39E-5 | 2.18 | 4.40E-5 | 2.16 | 4.44E-5 | 2.16 | 5.39E-5 | 2.10 | 5.34E-5 | 2.09 | 5.44E-5 | 2.09 |
| 256 | 9.95E-6 | 2.14 | 1.00E-5 | 2.14 | 1.01E-5 | 2.14 | 1.25E-5 | 2.11 | 1.26E-5 | 2.08 | 1.28E-5 | 2.09 |

| $N$ | $\beta = 0.6$ | $\beta = 0.8$ |
|-----|----------------|----------------|
|     | $\alpha = 0.10$ | $\alpha = 0.10$ | $\alpha = 0.10$ |
|     | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ |
|     | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate |
| 16  | 3.64E-3 | - | 3.63E-3 | - | 3.63E-3 | - | 3.76E-3 | - | 3.75E-3 | - | 3.75E-3 | - |
| 64  | 8.94E-4 | 1.01 | 8.87E-4 | 1.02 | 8.88E-4 | 1.02 | 9.73E-4 | 0.97 | 9.69E-4 | 0.98 | 9.72E-4 | 0.98 |
| 256 | 2.02E-4 | 1.07 | 2.00E-4 | 1.08 | 2.00E-4 | 1.08 | 2.34E-4 | 1.03 | 2.32E-4 | 1.03 | 2.33E-4 | 1.03 |

| $N$ | $\beta = 0.6$ | $\beta = 0.8$ |
|-----|----------------|----------------|
|     | $\alpha = 0.01$ | $\alpha = 0.01$ | $\alpha = 0.01$ |
|     | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ | $\|u\|_0$ |
|     | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate | $\|u\|_0$ | rate |
| 16  | 3.64E-3 | - | 3.63E-3 | - | 3.63E-3 | - | 3.75E-3 | - | 3.75E-3 | - | 3.75E-3 | - |
| 64  | 8.90E-4 | 1.02 | 8.87E-4 | 1.02 | 8.86E-4 | 1.02 | 9.72E-4 | 0.98 | 9.71E-4 | 0.98 | 9.69E-4 | 0.98 |
| 256 | 2.01E-4 | 1.07 | 2.01E-4 | 1.07 | 2.00E-4 | 1.08 | 2.33E-4 | 1.03 | 2.33E-4 | 1.03 | 2.32E-4 | 1.03 |
From Tables 1 and 2, we can obtain that the fully FE solution \( u_{h\tau} \) achieves the saturation error order \( O(\tau^2 + h^2) \) under \( \| \cdot \|_0 \) norm.

Fig. 1 illustrates the comparisons of exact solutions and numerical solutions of \( \alpha = 0.2, 0.4 \) and \( \beta = 0.6, 0.8 \) with \( t = 1 \) and \( h = \tau = 1/32 \).

![Figure 1: Illustration for comparisons on exact solutions and numerical solutions with \( t = 1 \) and \( h = \tau = 1/32 \).]

3. Condition number estimation

This section is devoted to deriving the condition number estimation on the coefficient matrix of (13) in uniform temporal and and spatial discretizations.

**Theorem 2.** For the linear system (13), we have

\[
\kappa(C^n_{h\tau}) = 1 + \mathcal{O}(\tau^\alpha h^{-2\beta}).
\]  

(18)

**Proof.** Let \( C_\alpha = \Gamma(3 - \alpha)/2 \), we divide our proof in three steps. First, it is trivially true that \( C^n_{h\tau} \) is spectrally equivalent to the matrix \( I + C_\alpha \tau^\alpha M_h^{-\frac{1}{2}} A_h^\beta M_h^{-\frac{1}{2}} \), i.e.

\[
\kappa(C^n_{h\tau}) \simeq \kappa\left(I + C_\alpha \tau^\alpha M_h^{-\frac{1}{2}} A_h^\beta M_h^{-\frac{1}{2}}\right).
\]  

(19)

The next thing to do in the proof is to verify

\[
\lambda_{\text{min}}(M_h^{-\frac{1}{2}} A_h^\beta M_h^{-\frac{1}{2}}) \gtrsim 1, \quad \lambda_{\text{max}}(M_h^{-\frac{1}{2}} A_h^\beta M_h^{-\frac{1}{2}}) \lesssim h^{-2\beta},
\]  

(20)
which is equivalent to

\[(\vec{v}_h, \vec{v}_h) \lesssim (M_h^{-\frac{1}{2}} A_h^\beta M_h^{-\frac{1}{2}} \vec{v}_h, \vec{v}_h) \lesssim h^{-2\beta} (\vec{v}_h, \vec{v}_h), \quad \forall \vec{v}_h \in \mathbb{R}^{M-1}. \]  

Set \( \vec{u}_h = M_h^{-\frac{1}{2}} \vec{v}_h := (u_h^1, \cdots, u_h^{M-1})^T \), rewrite (21) as \((M_h \vec{u}_h, \vec{u}_h) \lesssim (A_h^\beta \vec{u}_h, \vec{u}_h) \lesssim h^{-2\beta} (M_h \vec{u}_h, \vec{u}_h)\).

It is sufficient to verify that \((M_h \vec{u}_h, \vec{u}_h) \simeq h(\vec{u}_h, \vec{u}_h)\). It follows by (13) and the Cauchy-Schwarz inequality that

\[ h \left( \frac{1}{3} \right) (\vec{u}_h, \vec{u}_h) \leq (M_h \vec{u}_h, \vec{u}_h) = h \left[ \frac{2}{3} \sum_{i=1}^{M-1} (u_i^h)^2 + \frac{1}{3} \sum_{i=1}^{M-2} u_i^h u_{i+1}^h \right] \leq h(\vec{u}_h, \vec{u}_h). \]

Thus (20) will follow if we can show that \((M_h \vec{u}_h, \vec{u}_h) \lesssim (A_h^\beta \vec{u}_h, \vec{u}_h) \lesssim h^{1-2\beta}(\vec{u}_h, \vec{u}_h)\). We start by showing the second inequality. Utilizing Theorem 1 and the Cauchy-Schwarz inequality, we arrive at

\[
\begin{align*}
(A_h^\beta \vec{u}_h, \vec{u}_h) & \leq \sum_{i=1}^{M-1} \left( a_{i,i}^h (u_i^h)^2 - \frac{1}{2} \sum_{j=1}^{M-1} a_{i,j}^h (u_i^h)^2 + (u_j^h)^2 \right) \\
& = \sum_{i=1}^{M-1} a_{i,i}^h (u_i^h)^2 - \sum_{j=i+1}^{M-1} \sum_{i=1}^{M-1} a_{i,j}^h (u_i^h)^2 - \sum_{j=i+1}^{M-1} \sum_{j=1}^{M-1} a_{i,j}^h (u_j^h)^2 \\
& = \sum_{i=1}^{M-1} (u_i^h)^2 \left[ a_{i,i}^h - \sum_{j=i+1}^{M-1} \left( \sum_{j=1}^{M-1} a_{i,j}^h \right) \right] \\
& \leq 2 a_{1,1}^h (\vec{u}_h, \vec{u}_h) \leq \frac{(2^{4-2\beta} - 8)}{\cos(\beta \pi) \Gamma(4-2\beta)} h^{1-2\beta}(\vec{u}_h, \vec{u}_h),
\end{align*}
\]

which proves the second inequality. To prove the left inequality, set \( u_h := \Phi_h \vec{u}_h \), we rewrite it as

\[
\frac{1}{\cos(\beta \pi)} (x D_L^\beta u_h, x D_R^\beta u_h)_{L^2(\Omega)} = (A_h^\beta \vec{u}_h, \vec{u}_h) \gtrsim h(\vec{u}_h, \vec{u}_h) \simeq (M_h \vec{u}_h, \vec{u}_h) = (u_h, u_h)_{L^2(\Omega)}
\]

which can be deduced by (5) - (6), where \( \Phi_h = (\phi_1, \cdots, \phi_{M-1}) \).

Finally, we have to show that

\[
\kappa \left( I + C \tau^\alpha M^{-\frac{1}{2}} A_h^\beta M^{-\frac{1}{2}} \right) = \frac{\lambda_{\text{max}} \left( I + C \tau^\alpha M^{-\frac{1}{2}} A_h^\beta M^{-\frac{1}{2}} \right)}{\lambda_{\text{min}} \left( I + C \tau^\alpha M^{-\frac{1}{2}} A_h^\beta M^{-\frac{1}{2}} \right)} \\
= \frac{1 + C \tau^\alpha \lambda_{\text{max}} (M^{-\frac{1}{2}} A_h^\beta M^{-\frac{1}{2}})}{1 + C \tau^\alpha \lambda_{\text{min}} (M^{-\frac{1}{2}} A_h^\beta M^{-\frac{1}{2}})} \leq 1 + C \frac{(2^{4-2\beta} - 8)}{\cos(\beta \pi) \Gamma(4-2\beta)} \tau^\alpha h^{-2\beta}.
\]

This completes the proof based on the spectral equivalence relation (19).
Remark 4. The estimation \( (18) \) is compatible with the correlative result \( 1 + \mathcal{O}(\tau h^{-2}) \) of integer order parabolic differential equations.

An important particular case of Theorem 2 is singled out in the following corollary.

Corollary 3. Let \( \tau \) be proportional to \( h^{\mu} \) with \( \mu \alpha \geq 2 \beta \). Then

\[
\kappa(C_{n\tau}) = \mathcal{O}(1).
\] (22)

In what follows, we examine the correctness of (18) concerning Example 1 with typical \( \alpha \) and \( \beta \) for three specific cases: \( \tau = h \), \( \tau = h^{2} \) and \( \tau \) is fixed (doesn’t change along with \( h \)). In under-mentioned tables, \( \lambda_{\text{min}} \) and \( \lambda_{\text{max}} \) respectively indicate the smallest and largest eigenvalues, \( \kappa \) represents the condition number and ratio is the quotient of the condition number in fine grid divided by that in coarse grid.

| \( \alpha \) | \( M \) | \( \beta = 0.6 \) | \( \beta = 0.8 \) |
|---|---|---|---|
| \( \lambda_{\text{min}} \) | \( \lambda_{\text{max}} \) | \( \kappa \) | ratio | \( \lambda_{\text{min}} \) | \( \lambda_{\text{max}} \) | \( \kappa \) | ratio |
| 0.99 | 8 | 1.45E-1 | 1.97E-1 | 1.36E+0 | - | 1.63E-1 | 5.47E-1 | 3.35E+0 | - |
| 16 | 6.81E-2 | 1.09E-1 | 1.60E+0 | 1.18 | 7.26E-2 | 4.11E-1 | 5.66E+0 | 1.69 |
| 32 | 3.27E-2 | 6.03E-2 | 1.84E+0 | 1.15 | 3.93E-2 | 3.09E-1 | 9.12E+0 | 1.61 |
| 64 | 1.60E-2 | 3.34E-2 | 2.09E+0 | 1.13 | 1.63E-2 | 2.33E-1 | 1.43E+1 | 1.57 |
| 0.5 | 8 | 2.09E-1 | 5.41E-1 | 2.59E+0 | - | 2.74E-1 | 1.89E+0 | 6.89E+0 | - |
| 16 | 9.31E-2 | 4.25E-1 | 4.56E+0 | 1.76 | 1.16E-1 | 2.03E+0 | 1.74E+1 | 2.53 |
| 32 | 4.22E-2 | 3.37E-1 | 8.00E+0 | 1.75 | 5.05E-2 | 2.17E+0 | 4.30E+1 | 2.47 |
| 64 | 1.95E-2 | 2.70E-1 | 1.39E+1 | 1.73 | 2.24E-2 | 2.32E+0 | 1.04E+2 | 2.41 |
| 0.01 | 8 | 4.81E-1 | 2.08E+0 | 4.32E+0 | - | 7.50E-1 | 7.65E+0 | 1.02E+1 | - |
| 16 | 2.42E-1 | 2.35E+0 | 9.69E+0 | 2.24 | 3.77E-1 | 1.17E+1 | 3.99E+1 | 3.03 |
| 32 | 1.21E-1 | 2.66E+0 | 2.21E+1 | 2.28 | 1.88E-1 | 1.76E+1 | 9.36E+1 | 3.03 |
| 64 | 6.00E-2 | 3.03E+0 | 5.06E+1 | 2.29 | 9.35E-2 | 2.65E+1 | 2.83E+2 | 3.03 |

It is observed from Tables 3-5 that numerical results are in good agreement with our theoretical estimation.

4. AMG’s convergence analysis and an adaptive AMG method

Within the section, involving FFT to perform Toeplitz matrix-vector multiplications, we introduce the so-called Ruge-Stüber or classical AMG method \[37\] with low algorithmic complexity, fulfill its theoretical investigation, and then propose an adaptive AMG method through Corollary 3.
Table 4: The smallest and largest eigenvalues and condition numbers with $\tau = h^2$.

| $\alpha$ | $M$ | $\beta = 0.6$ | $\beta = 0.8$ |
|----------|-----|---------------|---------------|
|          |     | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio |
| $0.5$    | 8   | 1.52E-1       | 2.33E-1       | 1.53E+0    | -      | 1.76E-1       | 6.96E-1       | 3.97E+0    | -      |
|          | 16  | 6.98E-2       | 1.30E-1       | 1.86E+0    | 1.21   | 7.57E-2       | 5.23E-1       | 6.91E+0    | 1.74   |
|          | 32  | 3.31E-2       | 7.20E-2       | 2.17E+0    | 1.17   | 3.46E-2       | 3.92E-1       | 1.13E+1    | 1.64   |
|          | 64  | 1.61E-2       | 4.01E-2       | 2.49E+0    | 1.15   | 1.65E-2       | 2.95E-1       | 1.79E+1    | 1.58   |
| $0.01$   | 8   | 4.74E-1       | 2.04E+0       | 4.30E+0    | -      | 7.37E-1       | 7.49E+0       | 1.02E+1    | -      |
|          | 16  | 2.37E-1       | 2.28E+0       | 9.62E+0    | 2.24   | 3.69E-1       | 1.13E+1       | 3.08E+1    | 3.02   |
|          | 32  | 1.18E-1       | 2.57E+0       | 2.19E+1    | 2.27   | 1.83E-1       | 1.70E+1       | 9.30E+1    | 3.03   |
|          | 64  | 5.82E-2       | 2.91E+0       | 5.00E+1    | 2.29   | 9.03E-2       | 2.54E+1       | 2.81E+2    | 3.03   |

Table 5: The smallest and largest eigenvalues and the condition number with $\tau = 1/32$.

| $\alpha$ | $M$ | $\beta = 0.6$ | $\beta = 0.8$ |
|----------|-----|---------------|---------------|
|          |     | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio |
| $0.99$   | 64  | 1.64E-2       | 5.83E-2       | 3.56E+0    | 1.93   | 1.69E-2       | 4.58E-1       | 2.70E+1    | 2.96   |
|          | 128 | 8.19E-3       | 6.24E-2       | 7.62E+0    | 2.14   | 8.48E-3       | 6.89E-1       | 8.13E+1    | 3.01   |
|          | 256 | 4.10E-3       | 6.97E-2       | 1.70E+1    | 2.23   | 4.24E-3       | 1.04E+0       | 2.46E+2    | 3.02   |
|          | 512 | 2.05E-3       | 7.91E-2       | 3.86E+1    | 2.27   | 2.12E-3       | 1.58E+0       | 7.45E+2    | 3.03   |
| $0.5$    | 64  | 2.11E-2       | 3.80E-1       | 1.80E+1    | 2.25   | 2.52E-2       | 3.28E+0       | 1.30E+2    | 3.02   |
|          | 128 | 1.06E-2       | 4.33E-1       | 4.10E+1    | 2.28   | 1.26E-2       | 4.97E+0       | 3.94E+2    | 3.03   |
|          | 256 | 5.27E-3       | 4.95E-1       | 9.39E+1    | 2.29   | 6.31E-3       | 7.53E+0       | 1.19E+3    | 3.03   |
|          | 512 | 2.64E-3       | 5.68E-1       | 2.15E+2    | 2.29   | 3.16E-3       | 1.14E+1       | 3.62E+3    | 3.03   |
| $0.01$   | 64  | 6.03E-2       | 3.05E-0       | 5.07E+1    | 2.29   | 9.40E-2       | 2.67E+1       | 2.84E+2    | 3.03   |
|          | 128 | 3.01E-2       | 3.50E+0       | 1.16E+2    | 2.30   | 4.70E-2       | 4.05E+1       | 8.61E+2    | 3.03   |
|          | 256 | 1.51E-2       | 4.02E+0       | 2.67E+2    | 2.30   | 2.35E-2       | 6.13E+1       | 2.61E+3    | 3.03   |
|          | 512 | 7.53E-3       | 4.62E-0       | 6.14E+2    | 2.30   | 1.17E-2       | 9.29E+1       | 7.91E+3    | 3.03   |

| $\beta$ | $M$ | $\alpha = 0.01$ | $\alpha = 0.99$ |
|----------|-----|----------------|----------------|
|          |     | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio | $\lambda_{\text{min}}$ | $\lambda_{\text{max}}$ | $\kappa$ | ratio |
| $0.999$  | 64  | 1.63E-1       | 2.42E+2       | 1.49E+3    | 4.00   | 1.81E-2       | 4.11E+0       | 2.27E+2    | 3.98   |
|          | 128 | 8.14E-2       | 4.84E+2       | 5.95E+3    | 4.00   | 9.06E-3       | 8.21E+0       | 9.07E+2    | 3.99   |
|          | 256 | 4.07E-2       | 9.66E+2       | 2.38E+4    | 3.99   | 4.53E-3       | 1.64E+1       | 3.62E+3    | 3.99   |
|          | 512 | 2.03E-2       | 1.93E+3       | 9.49E+4    | 3.99   | 2.27E-3       | 3.28E+1       | 1.45E+4    | 3.99   |
Algorithm 1. The classical AMG method for the linear system (13).

Step 1 Perform the Setup phase to the coefficient matrix $C_{h^r}$.

1.1 Set the strength-of-connection tolerance $\theta$;

1.2 Build the ingredients required by a hierarchy of levels, coarsest to finest, including the grid transfer operator $P$.

Step 2 Invoke the classical $V(\varrho_1,\varrho_2)$-cycle to solve (13) until convergence. Below is the description of two-grid $V(\varrho_1,\varrho_2)$-cycle.

2.1 Do $\varrho_1$ pre-smoothing steps on (13);

2.2 Compute and restrict the residual: $r^c = P^T(G_{h^r} - C_{h^r}U_{h^r});$

2.3 Solve the residual equation on coarse level: $(P^T C_{h^r})e^c = r^c$;

2.4 Interpolation and correction: $U_{h^r} = U_{h^r} + Pe^c$;

2.5 Do $\varrho_2$ post-smoothing steps on (13).

Remark 5. In pre- and post-smoothing processes, damped-Jacobi iterative methods are favorable choices, which can maintain the low computational cost $O(M \log M)$ calculated by FFT.

For theoretical investigations, we rewrite (13) and the grid transfer operator $P$ in block form regarding a given C/F splitting

$$C_{h^r}U_{h^r} = \begin{pmatrix} A_{FF} & A_{FC} \\ A_{CF} & A_{CC} \end{pmatrix} \begin{pmatrix} u_F \\ u_C \end{pmatrix} = \begin{pmatrix} f_F \\ f_C \end{pmatrix} = G_{h^r}, \quad P = \begin{pmatrix} I_{FC} \\ I_{CC} \end{pmatrix},$$

and introduce the following inner products

$$(u_F, v_F)_0,F = (D_{FF}u_F, v_F), \quad (u, v)_1 = (C_{h^r}u, v), \quad (u, v)_2 = (D_{h^r}^{-1}C_{h^r}u, C_{h^r}v)$$

with their associated norms $\| \cdot \|_{0,F} = \sqrt{(\cdot, \cdot)_0,F}$ and $\| \cdot \|_i = \sqrt{(\cdot, \cdot)_i}$ ($i = 1, 2$), where $I_{CC}$ is the identity operator, $D_{FF} = \text{diag}(A_{FF})$ and $D_{h^r} = \text{diag}(C_{h^r})$.

For simplicity, we here denote $C_{h^r} = (c_{ij})_{(M-1)\times(M-1)}$, and only consider the two-grid $V(0,1)$-cycle, whose iteration matrix has the form

$$M_{h,H} = S[I - P(P^T C_{h^r}P)^{-1}P^T C_{h^r}],$$

where $S$ is a relaxation operator usually chosen as damped-Jacobi or Gauss-Seidel iterative method.

Combining Corollary 2 and the two-level convergence theory in the work [38], leads to the following lemmas and theorem.
Lemma 4. Under the condition (17), for all $e_h \in \mathbb{R}^{M-1}$, damped-Jacobi and Gauss-Seidel relaxations satisfy the smoothing property

$$\|S e_h\|_2^2 \leq \|e_h\|_2^2 - \sigma_1 \|e_h\|_2^2$$

(23)

with $\sigma_1$ independent of $e_h$ and step sizes $h$ and $\tau_n$.

Proof. On the strength of Theorem A.3.1 and A.3.2 in [38], we produce that damped-Jacobi relaxation with parameter $0 < \omega < 2/\eta$ satisfies (23) with

$$\sigma_1 = \omega(2 - \omega \eta),$$

and Gauss-Seidel relaxation satisfies (23) with

$$\sigma_1 = \frac{1}{(1 + \gamma_-(1 + \gamma_+))},$$

both independent of $e_h$, where

$$\eta \geq \rho(D^{-1}_{h\tau}C^n_{h\tau}), \quad \gamma_- = \max_i \left\{ \frac{1}{w_i c_{ii}} \sum_{j<i} w_j |c_{ij}| \right\}, \quad \gamma_+ = \max_i \left\{ \frac{1}{w_i c_{ii}} \sum_{j>i} w_j |c_{ij}| \right\},$$

and $w = (w_i)$ is an arbitrary positive vector with $C^n_{h\tau} w$ being also positive.

By exploiting (14)-(16), the assumption (17) and Corollary 2, we conclude that $C^n_{h\tau}$ is strictly diagonally dominant. Recall that $C^n_{h\tau} w$ is a positive vector, yield $\gamma_- < 1$, $\gamma_+ < 1$ and

$$\rho(D^{-1}_{h\tau}C^n_{h\tau}) \leq \|D^{-1}_{h\tau}C^n_{h\tau}\|_w = \max_i \left\{ \frac{1}{w_i} \sum_{j} w_j |c_{ij}| \right\} < 2,$$

(24)

which implicitly mean that $\eta$, $\gamma_-$ and $\gamma_+$ can be chosen to be independent of $h$ and $\tau_n$, and complete the proof.

Remark 6. The inequality (24) implies that there exists $\epsilon > 0$ such that $\rho(D^{-1}_{h\tau}C^n_{h\tau}) = 2 - 3\epsilon$. Then $\eta = 2 - 2\epsilon > \rho(D^{-1}_{h\tau}C^n_{h\tau})$ and hence the upper bound of parameter $\omega$: $2/\eta = 1/(1 - \epsilon) > 1$, which suggests that Jacobi relaxation with $\omega = 1$ is available in such a case.

Remark 7. For all symmetric M-matrices, $\sigma_1 \leq 1/\eta < 1$ holds for damped-Jacobi relaxation, while $\sigma_1 \in (1/4,1)$ for Gauss-Seidel relaxation.

Lemma 5. Under the condition (17) and a given C/F splitting, for all $e_h = (e^T_F, e^T_C)^T \in \mathbb{R}^{M-1}$, the direct interpolation $I_{FC}$ satisfies

$$\|e_F - I_{FC} e_C\|_{0,F}^2 \leq \sigma_2 \|e_h\|_2^2$$

(25)

with $\sigma_2$ independent of $e_h$, $h$ and $\tau_n$. 16
Proof. According to Theorem A.4.3 in [38], \( I_{FC} \) satisfies (25) with \( \sigma_2 \) of the form regarding a given C/F splitting

\[
\sigma_2 \geq \max_{i \in F} \left\{ \frac{\sum_{j \in N_i, c_{ij}}}{\sum_{j \in C_i, c_{ij}}} \right\}
\]

independent of \( e_h \), where \( N_i = \{ j \neq i : c_{ij} \neq 0 \} \), \( C_i \) is the subset of \( N_i \) whose values will be used to interpolate at F-point \( i \). As a result of (26) and the fact that \( c_{ij} (j \in N_i) \) are all negative, the following relation holds:

\[
\sigma_2 > \frac{1}{\pi}
\]

indicating that \( \sigma_2 \) is independent of \( h \) and \( \tau_n \) by plugging (17), and thus prove the theorem.

**Theorem 3.** Let any C/F splitting be given. Under the condition (17), there exist positive constants \( \sigma_1 \) and \( \sigma_2 \) independent of \( h \) and \( \tau_n \) and satisfying \( \sigma_2 > 1 > \sigma_1 \), such that a uniform two-grid convergence is achieved as follows

\[
\| M_{h,H} \|_1 \leq \sqrt{1 - \sigma_1/\sigma_2}.
\]

**Proof.** The proof of this result is straightforward and is based on Theorem A.4.1 and A.4.2 in [38], Lemma 4 and 5.

Now observe from Theorem 3 that, despite the independence of \( h \) and \( \tau_n \), \( \sigma_2 \) relies ruinously on \( \theta \) in Step 1.1 of Algorithm 1 due to the fact that \( C_{h,\tau}^n \) is nearly dense leading to a quite complicated adjacency graph. In addition, it is found that \( \sigma_2(\theta) \) may be much larger than 1 as \( \theta \) approaches zero, with that comes a sharp pullback in convergence rate. Hence, an appropriate \( \theta \) is a critical component of Algorithm 1 to handle fractional diffusion equations.

We now turn to reveal a reference formula on \( \theta \). Note the heuristic that the distribution of ratios of off-diagonal elements relative to the maximum absolute off-diagonal element (namely the minor diagonal element for \( C_{h,\tau}^n \)) plays a major role in the choice of \( \theta \). Since \( C_{h,\tau}^n \) is a symmetric Toeplitz matrix from Corollary 1 its first row involving all off-diagonal elements of \( C_{h,\tau}^n \) is deserved to be the representative row. Taking \( \beta = 0.8 \) as an example, Fig. 2 shows the distribution of the ratios \( c_{1j}/c_{12} \) (\( j \geq 2 \)), which reminds us of the attenuation in off-diagonal elements, states

\[
\frac{c_{13}}{c_{12}} \approx 0.160426, \quad \frac{c_{1j}}{c_{12}} < \frac{c_{14}}{c_{12}} \approx 0.034394, \quad j = 5, 6, \cdots, M - 1,
\]
and suggests that $c_{1j} \ (j \geq 4)$ should be viewed as weak couplings (wouldn’t be used for interpolation) because they are less than 5% of $c_{12}$. Besides, for a better complexity and higher efficiency, only the nearest neighbors are potentially used to limit the interpolation matrix on each grid level to at most 3 coefficients per row, although $c_{13}$ reaches around 16% of $c_{12}$. It thus appears that the strength-of-connection tolerance $\theta$ should be of the form

$$\theta = \frac{c_{13}}{c_{12}} + \epsilon_0,$$

where $\epsilon_0$ is some small number, which can be chosen to be $10^{-5}$ in one-dimensional realistic problems.

As is known, Algorithm 1 is much more expensive for well-conditioned problems than basic iterative techniques, such as conjugate gradient (CG) or (plain) Jacobi iterative method. For the purpose of solving (13) in an optimal way, an adaptive AMG method is proposed below by combining Algorithm 1, the reference formula (27) and the condition number estimation (22) in Corollary 3 as the clear distinction to adaptively pick an appropriate solver.

**Algorithm 2.** An adaptive AMG method $S_{ad}$ for the linear system (13).

**Step 1** If the condition (22) is unsatisfied, then goto Step 2, else set $S_{ad}$ as the CG or Jacobi iterative method;

**Step 2** Set $S_{ad}$ as the classical AMG method described in Algorithm 1, with $\theta$ chosen via the reference formula (27).

Figure 2: Distribution of ratios $c_{1j}/c_{12}, \ j = 2, 3, \cdots, M - 1$. 
5. Performance evaluation

Let us illustrate the effectiveness of Algorithms 1 and 2. Numerical experiments are performed in a 64 bit Fedora 18 platform, double precision arithmetic on Intel Xeon (W5590) with 24.0 GB RAM, 3.33 GHz, with an -O2 optimization parameter. In the following tables, dashed entries (-) indicate the solutions either diverge or fail to converge after 1000 iterations, $I_t$s is the number of iterations until the stopping criterion $10^{-12}$ is reached, $T_c$ represents the CPU time including both Setup and Solve phases with second as its unit, $C_g$ and $C_o$ respectively denote grid and operator complexities, which are defined as sums of the number of degrees of freedom and nonzero elements on all grid levels divided by those of the finest grid level, and used as measures for memory requirements, arithmetic operations and the execution time in Setup and Solve phases.

Example 2. **Comparisons of the classical AMG over CG and Jacobi iterative methods for the case when \((22)\) is satisfied with two different fractional orders.**

| $M$  | $\alpha = \beta = 0.6$ | $\alpha = \beta = 0.8$ |
|------|------------------|------------------|
|      | Jacobi | CG | AMG | Jacobi | CG | AMG |
| $I_t$s | $T_c$ | $I_t$s | $T_c$ | $I_t$s | $T_c$ | $I_t$s | $T_c$ | $I_t$s | $T_c$ |
| 32    | 18     | 1.78E-4 | 9     | 1.09E-4 | 4     | 5.22E-4 | 22    | 2.03E-4 | 11     | 1.27E-4 | 5     | 2.84E-4 |
| 64    | 18     | 3.90E-4 | 11    | 2.19E-4 | 4     | 7.79E-4 | 23    | 4.84E-4 | 13     | 1.99E-4 | 5     | 6.33E-4 |
| 128   | 19     | 1.31E-3 | 11    | 4.60E-4 | 4     | 2.52E-3 | 23    | 1.54E-3 | 13     | 5.26E-4 | 5     | 1.89E-3 |
| 256   | 19     | 4.69E-3 | 11    | 1.57E-3 | 4     | 9.73E-3 | 23    | 5.66E-3 | 13     | 1.82E-3 | 5     | 7.06E-3 |
| 512   | 19     | 2.61E-2 | 11    | 8.03E-3 | 4     | 5.49E-2 | 23    | 3.12E-2 | 13     | 9.56E-3 | 5     | 4.57E-2 |
| 1024  | 19     | 1.95E-1 | 11    | 6.04E-2 | 4     | 1.73E-1 | 23    | 2.36E-1 | 12     | 6.53E-2 | 5     | 1.32E-1 |
| 2048  | 19     | 3.98E-1 | 11    | 1.22E-1 | 4     | 9.39E-1 | 23    | 9.11E-1 | 12     | 1.32E-1 | 5     | 7.49E-1 |
| 4096  | 19     | 3.03    | 11    | 9.25E-1 | 4     | 2.80    | 23    | 3.65    | 12     | 1.01    | 5     | 2.98    |

As expected, the results in Table 6 show that Jacobi, CG and AMG methods are robust with respect to the mesh size and fractional order, which indicates indirectly the correctness of \((18)\). In addition, CG method runs 3.28 and 3.03 times faster than Jacobi and AMG methods for $M = 4096$ and $\alpha = \beta = 0.6$, respectively.

Example 3. **Comparisons between the classical AMG method and CG method for the case when \((22)\) is unsatisfied.**
As shown in Table 7, AMG method converges robustly regarding to the mesh size and may be weakly dependent of $\beta$, while the number of iterations of CG method is quite unstable, and sometimes CG method even break down. Furthermore AMG method runs 12.24 times faster than CG method for $M = 4096$ and $\beta = 0.8$.

Table 8 shows the results of $\tau = h$. Despite the advantage in computational cost and robustness over CG method, AMG method is nearly independent of $\alpha$ and $\beta$ in this circumstance. Meanwhile, by an investigation in terms of number of iterations in Tables 7 and 8, CG method converges faster because of the improvement in condition number from $O(h^{-2\beta})$ to $O(h^{\alpha-2\beta})$.

**Example 4.** Comparisons of $S_{ad}$ over the classical AMG and CG methods when the $i$-th time step size $\tau_i$ is chosen to be

$$
\tau_i = \begin{cases} 
h^2, & i = 1, \cdots, K_1 \\
1/32, & i = K_1 + 1, \cdots, K_1 + K_2
\end{cases}
$$

We can observe from Table 9 that $S_{ad}$ and AMG methods are fairly robust as to the mesh size, roughly 10 and 6 on the average. Yet the average number of iterations of CG method varies from 85 to 142. Moreover $S_{ad}$ has a considerable advantage over others in CPU time, runs 1.72 and 6.09 times faster than AMG and CG methods for $M = 2048$ and $K_2 = 100$. 

---

Table 7: Number of iterations and wall time for the case $\tau = 1/32$.

| $M$ | $\beta = 0.6$ | $\beta = 0.8$ | $\beta = 0.99$ |
|-----|---------------|---------------|----------------|
|     | CG | AMG | CG | AMG | CG | AMG |
|-----|----|-----|----|-----|----|-----|
|     | $H_t$ | $T_c$ | $H_t$ | $T_c$ | $H_t$ | $T_c$ | $H_t$ | $T_c$ | $H_t$ | $T_c$ |
| 512 | 97  | 0.119 | 8  | 0.042 | 180 | 0.209 | 8  | 0.069 | 256 | 0.314 |
| 1024| 147 | 0.715 | 8  | 0.169 | 314 | 1.627 | 8  | 0.301 | 512 | 2.537 |
| 2048| 223 | 2.230 | 8  | 0.677 | 546 | 6.037 | 8  | 0.772 | >1000 | - |
| 4096| 337 | 13.378 | 8  | 2.735 | 948 | 13.378 | 8  | 3.143 | >1000 | - |

Table 8: Number of iterations and wall time for the case $\tau = h$.

| $M$ | $\alpha = 0.2, \beta = 0.6$ | $\alpha = 0.6, \beta = 0.8$ |
|-----|----------------------------|----------------------------|
|     | CG | AMG | CG | AMG |
|-----|----|-----|----|-----|
|     | $H_t$ | $T_c$ | $H_t$ | $T_c$ | $H_t$ | $T_c$ | $H_t$ | $T_c$ |
| 128 | 40  | 1.686E-3 | 8  | 2.868E-3 | 57  | 2.310E-3 | 7  | 2.873E-3 |
| 256 | 62  | 9.090E-3 | 8  | 1.571E-2 | 99  | 1.622E-2 | 7  | 1.081E-2 |
| 512 | 95  | 1.523E-1 | 8  | 6.976E-2 | 171 | 2.626E-1 | 7  | 4.808E-2 |
| 1024| 145 | 5.166E-1 | 8  | 2.916E-1 | 291 | 1.7895   | 7  | 2.063E-1 |
Table 9: Comparisons among $S_{ad}$, CG and AMG.

| $K_2$ | $M = 1024, K_1 = K_2$ | $M = 2048, K_1 = K_2$ |
|-------|----------------|----------------|
|       | $S_{ad}$ | CG | AMG | $S_{ad}$ | CG | AMG |
|       | Its | $T_c$ | Its | $T_c$ | Its | $T_c$ | Its | $T_c$ | Its | $T_c$ | Its | $T_c$ |
| 25    | 459 | 2.09 | 7994 | 9.99 | 303 | 2.91 | 458 | 8.65 | 13681 | 71.76 | 304 | 12.03 |
| 50    | 909 | 4.11 | 15900 | 19.65 | 603 | 5.79 | 895 | 17.25 | 27196 | 141.81 | 604 | 23.79 |
| 75    | 1337 | 6.54 | 23748 | 30.74 | 903 | 8.60 | 1320 | 25.59 | 40637 | 212.48 | 904 | 35.76 |
| 100   | 1760 | 8.54 | 31573 | 40.12 | 1201 | 11.92 | 1745 | 36.07 | 54038 | 292.03 | 1204 | 50.66 |

Example 5. Analyze effects of the strength-of-connection tolerance $\theta$ on the performance of the classical AMG method.

Table 10: Effect of $\theta$ on the classical AMG when $M = 512$.

| $\theta$ | $\beta = 0.8$ | $\beta = 0.99$ |
|-----------|----------------|----------------|
|           | Its | $T_c$ | $C_{g}$ | $C_{o}$ | Its | $T_c$ | $C_{g}$ | $C_{o}$ |
| 0.0001    | 293 | 1.952 | 1.037 | 1.001 | 103 | 6.784E-1 | 1.170 | 1.021 |
| 0.001     | 83  | 5.618E-1 | 1.098 | 1.008 | 60  | 4.189E-1 | 1.498 | 1.124 |
| 0.00684   | 31  | 1.797E-1 | 1.202 | 1.029 | 32  | 1.374E-1 | 1.652 | 1.147 |
| 0.00685   | 31  | 1.789E-1 | 1.202 | 1.029 | 3  | 3.255E-1 | 1.975 | 1.331 |
| 0.01      | 23  | 1.620E-1 | 1.247 | 1.041 | 3  | 3.301E-2 | 1.975 | 1.331 |
| 0.1       | 13  | 6.299E-2 | 1.489 | 1.124 | 3  | 4.356E-2 | 1.975 | 1.331 |
| 0.16042   | 13  | 6.037E-2 | 1.489 | 1.124 | 3  | 4.118E-2 | 1.975 | 1.331 |
| 0.16043   | 7   | 4.699E-2 | 1.975 | 1.331 | 3  | 3.158E-2 | 1.975 | 1.331 |
| 0.25      | 7   | 4.736E-2 | 1.975 | 1.331 | 3  | 4.353E-2 | 1.975 | 1.331 |
| 0.5       | 7   | 4.710E-2 | 1.975 | 1.331 | 3  | 4.354E-2 | 1.975 | 1.331 |

It is seen from Tables 10 and 11 that there is a unique threshold $\theta_0$ independent of $h$ which guarantees the robustness of the classical AMG method, and makes number of iterations of the classical AMG monotonically decreasing when $\theta < \theta_0$, or even the classical AMG possibly diverge when $\theta$ is small enough, e.g., $\theta_0 = 0.16043$ and $\theta_0 = 0.00685$ for cases $\beta = 0.8$ and $\beta = 0.99$. By
Table 11: Effect of $\theta$ on the classical AMG when $M = 2048$.

| $\theta$ | $\beta = 0.8$ |        |        |        |        | $\beta = 0.99$ |        |        |        |
|----------|---------------|--------|--------|--------|--------|---------------|--------|--------|--------|
|          | $R_t$  | $T_C$ | $C_g$ | $C_o$  |        | $R_t$  | $T_C$ | $C_g$ | $C_o$  |
| 0.0001   | 335   | 23.62 | 1.038 | 1.001  | 123   | 6.310 | 1.170 | 1.021  |
| 0.001    | 107   | 9.024 | 1.100 | 1.068  | 102   | 5.597 | 1.497 | 1.125  |
| 0.00684  | 33    | 3.096 | 1.207 | 1.029  | 33    | 2.055 | 1.662 | 1.148  |
| 0.00685  | 33    | 3.129 | 1.207 | 1.029  | 3     | 7.314E-1 | 1.993 | 1.333  |
| 0.01     | 26    | 2.551 | 1.249 | 1.042  | 3     | 5.369E-1 | 1.993 | 1.333  |
| 0.1      | 15    | 1.691 | 1.497 | 1.125  | 3     | 5.448E-1 | 1.993 | 1.333  |
| 0.1603   | 15    | 1.690 | 1.497 | 1.125  | 3     | 5.372E-1 | 1.993 | 1.333  |
| 0.1604   | 8     | 1.217 | 1.993 | 1.333  | 3     | 5.294E-1 | 1.993 | 1.333  |
| 0.2      | 8     | 1.219 | 1.993 | 1.333  | 3     | 5.294E-1 | 1.993 | 1.333  |
| 0.25     | 8     | 1.217 | 1.993 | 1.333  | 3     | 5.299E-1 | 1.993 | 1.333  |

Direct calculations, we have $c_{13}/c_{12} \approx 0.160426$ and $c_{13}/c_{12} \approx 0.006846$. Utilizing the relation $\epsilon_0 = 10^{-5}$, the corresponding values of $\theta$ are respectively larger than those of $\theta_0$. This confirms the reasonability of the reference formula (27).

6. Conclusion

In this paper, we propose the variational formulation for a class of time-space Caputo-Riesz fractional diffusion equations, prove that the resulting matrix is a symmetric Toeplitz matrix, an M-matrix by appending a very weak constraint and its condition number is bounded by $1 + O(\tau^\alpha h^{-2\beta})$, introduce the classical AMG method and prove rigorously that its convergence rate is independent of time and space step sizes, provide explicitly a reference formula of the strength-of-connection tolerance to guarantee the robustness and predictable behavior of AMG method in all cases, and develop an adaptive AMG method via our condition number estimation to decrease the computation cost. Numerical results are all in conformity with the theoretical results, and verify the reasonability of the reference formula and the considerable advantage of the proposed adaptive AMG algorithm over other traditional iterative methods, e.g. Jacobi, CG and the classical AMG methods.

Acknowledgments

This work is under auspices of National Natural Science Foundation of China (11571293, 11601460, 11601462) and the General Project of Hunan Provincial Education Department of China (16C1540, 17C1527).
References

[1] F. Liu, P. Zhuang, V. Anh, I. Turner, K. Burrage, Stability and convergence of the difference methods for the space-time fractional advection-diffusion equation, Appl. Math. Comput. 191 (1) (2007) 12-20.

[2] Z. Q. Ding, A. G. Xiao, M. Li, Weighted finite difference methods for a class of space fractional partial differential equations with variable coefficients, J. Comput. Appl. Math. 233 (8) (2010) 1905-1914.

[3] G. H. Gao, Z. Z. Sun, A compact finite difference scheme for the fractional sub-diffusion equations, J. Comput. Phys. 230 (3) (2011) 586-595.

[4] S. P. Yang, A. G. Xiao, X. Y. Pan, Dependence analysis of the solutions on the parameters of fractional delay differential equations, Adv. Appl. Math. Mech. 3 (5) (2011) 586-597.

[5] X. N. Cao, J. L. Fu, H. Huang, Numerical method for the time fractional Fokker-Planck equation, Adv. Appl. Math. Mech. 4 (6) (2012) 848-863.

[6] H. Wang, T. S. Basu, A fast finite difference method for two-dimensional space-fractional diffusion equations, SIAM J. Sci. Comput. 34 (5) (2012) A2444-A2458.

[7] M. H. Chen, W. H. Deng, Y. J. Wu, Superlinearly convergent algorithms for the two-dimensional space-time Caputo-Riesz fractional diffusion equation, Appl. Numer. Math. 70 (2013) 22-41.

[8] D. L. Wang, A. G. Xiao, H. L. Liu, Dissipativity and stability analysis for fractional functional differential equations, Fract. Calc. Appl. Anal. 18 (6) (2015) 1399-1422.

[9] D. L. Wang, A. G. Xiao, W. Yang, Maximum-norm error analysis of a difference scheme for the space fractional CNLS, Appl. Math. Comput. 257 (2015) 241-251.

[10] W. Yang, D. L. Wang, L. Yang, A stable numerical method for space fractional Landau-Lifshitz equations, Appl. Math. Lett. 61 (2016) 149-155.

[11] V. J. Ervin, J. P. Roop, Variational formulation for the stationary fractional advection dispersion equation, Numer. Methods Partial Differ. Equ. 22 (3) (2006) 558-576.
[12] H. Zhang, F. Liu, V. Anh, Galerkin finite element approximation of symmetric space-fractional partial differential equations, Appl. Math. Comput. 217 (6) (2010) 2534-2545.

[13] K. Burrage, N. Hale, D. Kay, An efficient implicit FEM scheme for fractional-in-space reaction-diffusion equations, SIAM J. Sci. Comput. 34 (4) (2012) A2145-A2172.

[14] W. P. Bu, Y. F. Tang, J. Y. Yang, Galerkin finite element method for two-dimensional Riesz space fractional diffusion equations, J. Comput. Phys. 276 (2014) 26-38.

[15] K. Mustapha, B. Abdallah, K. M. Furati, A discontinuous Petrov-Galerkin method for time-fractional diffusion equations, SIAM J. Numer. Anal. 52 (5) (2014) 2512-2529.

[16] L. B. Feng, P. Zhuang, F. Liu, I. Turner, Y. T. Gu, Finite element method for space-time fractional diffusion equation, Numer. Algor. 72 (3) (2016) 749-767.

[17] W. P. Bu, A. G. Xiao, W. Zeng, Finite difference/finite element methods for distributed-order time fractional diffusion equations, J. Sci. Comput. 72 (1) (2017) 422-441.

[18] F. Liu, P. Zhuang, I. Turner, K. Burrage, V. Anh, A new fractional finite volume method for solving the fractional diffusion equation, Appl. Math. Model. 38 (15-16) (2014) 3871-3878.

[19] Y. M. Lin, C. J. Xu, Finite difference/spectral approximations for the time-fractional diffusion equation, J. Comput. Phys. 225 (2) (2007) 1533-1552.

[20] M. Zayernouri, W. R. Cao, Z. Q. Zhang, G. E. Karniadakis, Spectral and discontinuous spectral element methods for fractional delay equations, SIAM J. Sci. Comput. 36 (6) (2014) B904-B929.

[21] Y. Yang, Y. P. Chen, Y. Q. Huang, Spectral-collocation method for fractional Fredholm integro-differential equations, J. Korean Math. Soc. 51 (1) (2014) 203-224.

[22] Y. Yang, Y. P. Chen, Y. Q. Huang, Convergence analysis of the Jacobi spectral-collocation method for fractional integro-differential equations. Acta Math. Sci. 34B (3) (2014) 673-690.

[23] Y. Yang, Jacobi spectral Galerkin methods for fractional integro-differential equations, Calcolo 52 (4) (2015) 519-542.

[24] Y. Yang, Jacobi spectral Galerkin methods for Volterra integral equations with weakly singular kernel, Bull. Korean Math. Soc. 53 (1) (2016) 247-262.
[25] S. P. Yang, A. G. Xiao, An efficient numerical method for fractional differential equations with two Caputo derivatives, J. Comput. Math. 34 (2) (2016) 113-134.

[26] S. L. Lei, H. W. Sun, A circulant preconditioner for fractional diffusion equations, J. Comput. Phys. 242 (2013) 715-725.

[27] T. Moroney, Q. Q. Yang, A banded preconditioner for the two-sided, nonlinear space-fractional diffusion equation, Comput. Math. Appl. 66 (5) (2013) 659-667.

[28] J. H. Jia, H. Wang, Fast finite difference methods for space-fractional diffusion equations with fractional derivative boundary conditions, J. Comput. Phys. 293 (2015) 359-369.

[29] X. M. Gu, T. Z. Huang, X. L. Zhao, H. B. Li, L. Li, Strang-type preconditioners for solving fractional diffusion equations by boundary value methods, J. Comput. Appl. Math. 277 (2015) 73-86.

[30] M. Donatelli, M. Mazza, S. Serra-Capizzano, Spectral analysis and structure preserving preconditioners for fractional diffusion equations, J. Comput. Phys. 307 (2016) 262-279.

[31] H. K. Pang, H. W. Sun, Multigrid method for fractional diffusion equations, J. Comput. Phys. 231 (2012) 693-703.

[32] W. P. Bu, X. T. Liu, Y. F. Tang, J. Y. Yang, Finite element multigrid method for multi-term time fractional advection diffusion equations, Int. J. Model. Simul. Sci. Comput. 6 (2015) 1540001.

[33] Y. J. Jiang, X. J. Xu, Multigrid methods for space fractional partial differential equations, J. Comput. Phys. 302 (2015) 374-392.

[34] L. Chen, R. H. Nochetto, E. Otárola, A. J. Salgado, Multilevel methods for nonuniformly elliptic operators and fractional diffusion, Math. Comput. 85 (302) (2016) 2583-2607.

[35] X. Zhao, X. Z. Hu, W. Cai, G. E. Karniadakis, Adaptive finite element method for fractional differential equations using hierarchical matrices, arXiv: 1603.01358v2.

[36] J. Xu, Iterative methods by space decomposition and subspace correction, SIAM Rev. 34 (4) (1992) 581-613.

[37] J. W. Ruge, K. Stüben, Algebraic multigrid, in Multigrid Methods, Front. Appl. Math. 3 (1987) 73-130.
[38] K. Stüben, An introduction to algebraic multigrid, in Multigrid, U. Trottenberg, C. W. Oosterlee and A. Schüller, eds., Academic Press, Singapore, 2001.