Journal of Applied Analysis and Computation
Volume 9, Number 4, August 2019, 1359–1392
Website:http://jaac.ijournal.cn
DOI:10.11948/2156-907X.20180247

FAST SECOND-ORDER ACCURATE DIFFERENCE SCHEMES FOR TIME DISTRIBUTED-ORDER AND RIESZ SPACE FRACTIONAL DIFFUSION EQUATIONS

Huanyan Jian[1], Tingzhu Huang[1,†], Xile Zhao[1] and Yongliang Zhao[1]

Abstract The aim of this paper is to develop fast second-order accurate difference schemes for solving one- and two-dimensional time distributed-order and Riesz space fractional diffusion equations. We adopt the same measures for one- and two-dimensional problems as follows: we first transform the time distributed-order fractional diffusion problem into the multi-term time-space fractional diffusion problem with the composite trapezoid formula. Then, we propose a second-order accurate difference scheme based on the interpolation approximation on a special point to solve the resultant problem. Meanwhile, the unconditional stability and convergence of the new difference scheme in L_2-norm are proved. Furthermore, we find that the discretizations lead to a series of Toeplitz systems which can be efficiently solved by Krylov subspace methods with suitable circulant preconditioners. Finally, numerical results are presented to show the effectiveness of the proposed difference methods and demonstrate the fast convergence of our preconditioned Krylov subspace methods.

Keywords Distributed-order equation, multi-term fractional diffusion, Toeplitz matrix, circulant preconditioner, Krylov subspace method.

MSC(2010) 65N06, 65N12, 65F08, 65F10.

1. Introduction

Fractional diffusion equations (FDEs) have recently attracted considerable attention and interest due to its wide applications [33, 34]. Specifically, the time-fractional anomalous diffusion equation has become the focus of intensive investigations from both theoretical and practical perspectives [16, 24, 29].

Recently, the time-fractional anomalous diffusion equation with a single-term temporal derivative has been discussed and studied [23]. In [27, 28], the two-term time FDE was reported for describing processes that tend to be less anomalous. More generalized models were also developed as multi-term FDEs [2, 39], where several fractional derivatives were simultaneously involved. Although the single-term and multi-term FDEs are used extensively in many scientific fields, it is difficult for

†the corresponding author. Email address:tingzhuhuang@126.com (T. Huang)
[1]School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China

them to describe the non-Markovian processes for continuous time-scale distributions. Therefore, the time distributed-order FDEs [3,38,40,41] began to attract the attention of researchers. It can be considered as a generalization of the multi-term FDEs and has been found to be an important tool for modeling ultraslow diffusion processes and accelerating sub-diffusion [7,20].

Multiple numerical approaches [4,30,32] have emerged for solving the distributed-order FDEs, among which the finite difference method has grown popular [11,17]. The numerical method presented in [19] for solving the distributed-order FDE consists of: (a) approximation of the integral with a finite sum using a simple quadrature rule so that the distributed order FDE is converted into a multi-term FDE and (b) development of a numerical method to solve the resultant multi-term FDE. Such idea is essential for numerically solving the distributed-order FDEs and should be studied extensively. However, as far as we know, only a few algorithms have been developed to solve the distribution-order FDEs based on this idea. Ye et al. [37] proposed an implicit difference method for the time distributed-order and Riesz space FDEs on bounded domains and proved the difference method was unconditionally stable and convergent. An implicit numerical method of a new time distributed-order and two-sided space-fractional advection-dispersion equation was constructed by Hu et al. [17]. In [11], Gao et al. explored two alternating direction implicit difference schemes with the unconditional stability and convergence analysis for solving the 2D distributed-order FDEs. Bu et al. [4] introduced the finite difference method for a class of distributed-order time FDEs on bounded domains. In addition, most of these numerical approaches have no complete theoretical analysis of stability and convergence, especially for the time distribution-order and spatial FDEs [19,26].

In this paper, inspired by the above observations, we consider effective numerical methods for the following new time distributed-order and Riesz space FDEs (TDRFDEs):

$$D_t^{\omega(\alpha)}u(\mathbf{x},t) = Au(\mathbf{x},t) + f(\mathbf{x},t), \quad \mathbf{x}\in\Omega,\ 0<t\le T, \tag{1.1}$$

$$u(\mathbf{x},t)|_{\mathbf{x}\in\partial\Omega} = 0, \quad 0\le t\le T, \tag{1.2}$$

$$u(\mathbf{x},0) = \phi(\mathbf{x}), \quad \mathbf{x}\in\Omega, \tag{1.3}$$

where $\alpha\in(0,1]$, A is an operator and the function $f(\mathbf{x},t)$ is the source term with sufficient smoothness. In particular, if $\Omega=(x_L,x_R)\subset\mathbb{R}$, then

$$A = K\frac{\partial^\beta}{\partial|x|^\beta}, \quad K>0, \quad f(\mathbf{x},t)=f(x,t);$$

if $\Omega=(x_L,x_R)\times(y_L,y_R)\subset\mathbb{R}^2$, then

$$A = K_1\frac{\partial^\beta}{\partial|x|^\beta} + K_2\frac{\partial^\gamma}{\partial|y|^\gamma}, \quad K_1,K_2>0, \quad f(\mathbf{x},t)=f(x,y,t),$$

where $\beta,\gamma\in(1,2]$, and the $\frac{\partial^\beta}{\partial|x|^\beta}$ is the Riesz fractional derivative of order $\beta\in(1,2]$ defined as [18] ($\frac{\partial^\gamma}{\partial|y|^\gamma}$ is defined similarly)

$$\frac{\partial^\beta u(x,t)}{\partial|x|^\beta} = \begin{cases} -\frac{1}{2\cos(\beta\pi/2)\Gamma(2-\beta)}\frac{d^2}{dx^2}\int_{x_L}^{x_R}|x-\xi|^{1-\beta}u(\xi,t)d\xi, & 1<\beta<2,\\ \frac{\partial^2 u(x,t)}{\partial x^2}, & \beta=2,\end{cases}$$

where $\Gamma(\cdot)$ denotes the Gamma function. Moreover, the time distributed-order operator $D_t^{\omega(\alpha)}$ is defined by [31]

$$D_t^{\omega(\alpha)}u(\mathbf{x},t) = \int_0^1 \omega(\alpha) {}_0^C D_t^{\alpha} u(\mathbf{x},t) d\alpha,$$

where ${}_0^C D_t^{\alpha}$ denotes the Caputo fractional derivative [25] which is defined as follows:

$${}_0^C D_t^{\alpha} u(\mathbf{x},t) = \begin{cases} \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-\xi)^{-\alpha} \frac{\partial u}{\partial \xi}(\mathbf{x},\xi) d\xi, & 0 < \alpha < 1, \\ u_t(\mathbf{x},t), \quad \alpha = 1, \end{cases}$$

and the non-negative weight function $\omega(\alpha)$ satisfies that

$$0 \le \omega(\alpha),\ \omega(\alpha) \ne 0,\ \alpha \in [0,1],\ 0 < \int_0^1 \omega(\alpha) d\alpha < \infty.$$

Nonlocal behavior has been remarked as one of the main characteristics of the fractional differential operator. As a result, most numerical methods for FDEs produce dense matrices or even full coefficient matrices in 1D cases [13, 22]. Traditional methods, such as Gaussian elimination, need computational workload of $\mathcal{O}(M^3)$ and memory capacity of $\mathcal{O}(M^2)$, where M is the number of grid points [22]. Krylov subspace methods are studied and adopted to reduce the costs [12, 14]. The convergent speed of the Krylov subspace methods is dependent on the conditions of the discretized systems. To improve the performance of iterative methods, many preconditioners [15, 22, 42] are always designed according to the structure of the linear systems. For 1D cases, Wang et al. [36] made the important discovery that the resultant systems had Toeplitz coefficient matrices. By exploiting this structure, the memory requirement can be reduced from $\mathcal{O}(M^2)$ to $\mathcal{O}(M)$, and the fast Fourier transform (FFT) can be used to evaluate the matrix-vector product in $\mathcal{O}(M \log M)$ operations. Moreover, the coefficient matrices discretized from (1.1)-(1.3) should be symmetric positive definite Toeplitz matrices due to the existence of Riesz fractional derivatives. The circulant preconditioners [5, 6, 22] proved to be good choices to accelerate the convergence of Krylov subspace methods when solving the discretized linear systems. In high-dimensional cases, a nonsingular multilevel circulant preconditioner was proposed by Lei et al. [21], which efficiently accelerated the convergence rate of the Krylov subspace method. In [8], Chou et al. illustrated the efficiency of applying an approximate inverse preconditioner to the high dimensional FDEs when Krylov subspace methods are employed. They also showed that under certain conditions, the normalized preconditioned matrix is equal to the sum of an identity matrix, a matrix with small norm, and a matrix with low rank, such that the preconditioned Krylov subspace method converges superlinearly.

In this paper, we focus on establishing a fast numerical method and investigating the unconditional stability and convergence for solving the TDRFDEs (1.1)-(1.3). We first transform TDRFDEs (1.1)-(1.3) into multi-term time-space FDEs based on the composite trapezoid formula. Then we apply the interpolation approximation, as introduced by Gao et al. in [9], to approximate the time derivatives of the multi-term time-space FDEs at a special point. The global second-order numerical accuracy in time is independent to the order of fractional derivatives. To gather numerical solutions with high-order accuracy in space, the fractional centred difference formula [37] is used to discrete the space Riesz derivative. Therefore we develop a new difference scheme which converges with the second-order accuracy in time, space and distributed-order. On the other hand, by taking advantage of the Toeplitz structure of the resultant linear systems, we adopt the Krylov subspace method with efficient circulant preconditioners. It also proves that the eigenvalues of the preconditioned matrices are clustered around 1, and the convergence rate of our proposed iterative method is superlinear.

The rest of the paper is arranged as follows. In Section 2, we study the T-DRFDEs in the 1D case and present its corresponding difference scheme. The uniqueness, unconditional stability and convergence of the difference method are proved. Meanwhile, we design a preconditioned Krylov subspace method to solve the resultant Toeplitz linear system. In Section 3, the 2D TDRFDEs is discussed. We demonstrate that the difference scheme is uniquely solvable, unconditionally stable and convergent with the second order. We also adopt the preconditioned Krylov subspace method with suitable circulant preconditioners to handle the resulting system. Numerical experiments are carried out in Section 4 to illustrate the efficiency of our numerical approaches. Finally, the paper closes with conclusions and remarks in Section 5.

2. 1D problem

Consider the following 1D TDRFDEs:

$$\int_0^1 \omega(\alpha)\,{}_0^C D_t^\alpha u(x,t)d\alpha = K\frac{\partial^\beta u(x,t)}{\partial|x|^\beta} + f(x,t), \quad 0<x<L,\ 0<t\le T, \tag{2.1}$$

$$u(0,t)=0, \quad u(L,t)=0, \quad 0\le t\le T, \tag{2.2}$$

$$u(x,0)=\phi(x), \quad 0<x<L. \tag{2.3}$$

In this section, we show that the discretizations for the distributed-order integral term of (2.1) by the composite trapezoid formula lead to multi-term time-space FDE. We propose the second-order difference scheme based on the interpolation approximation on a special point to solve the multi-term equations. We also prove that the difference scheme is uniquely solvable, unconditionally stable and convergent with second-order accuracy in time, space and distributed-order integral variables. Moreover, we propose an efficient implementation based on Krylov subspace solver with suitable circulant preconditioners to solve the resultant Toeplitz linear system.

2.1. Numerical discretization of the (2.1)-(2.3)

We first discretize the integral interval $[0,1]$ by the grid $0=\alpha_0<\alpha_1<\cdots<\alpha_{2J}=1$ with $\Delta\alpha=\frac{1}{2J}$ and $\alpha_l = l\Delta\alpha,\ l=0,1,2,\cdots,2J$. The following lemma gives a complete description of the numerical approximation to the distributed-order integral term.

Lemma 2.1 (The composite trapezoid formula [10, 11]). *Let $z(\alpha)\in C^2([0,1])$, then we have*

$$\int_0^1 z(\alpha)d\alpha = \Delta\alpha\sum_{l=0}^{2J} d_l z(\alpha_l) - \frac{\Delta\alpha^2}{12}z^{(2)}(\eta), \quad \eta\in(0,1),$$

where

$$d_l = \begin{cases} \frac{1}{2}, & l = 0,\ 2J, \\ 1, & 1 \le l \le 2J-1. \end{cases}$$

Considering the left side of (2.1), let $z(\alpha) = \omega(\alpha){}_0^C D_t^{\alpha} u(x,t)$ and using Lemma 2.1, we can obtain

$$\int_0^1 \omega(\alpha){}_0^C D_t^{\alpha} u(x,t) d\alpha = \Delta\alpha \sum_{r=0}^{2J} d_r \omega(\alpha_r){}_0^C D_t^{\alpha_r} u(x,t) + \mathcal{O}(\Delta\alpha^2). \tag{2.4}$$

Let $m = 2J,\ \lambda_r = d_r \omega(\alpha_r)\Delta\alpha$. The problem (2.1)-(2.3) is now converted into the following multi-term time-space FDEs:

$$\sum_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(x,t) = K\frac{\partial^\beta u(x,t)}{\partial |x|^\beta} + f(x,t), \quad 0 < x < L,\ 0 < t \le T, \tag{2.5}$$

$$u(0,t) = 0, \quad u(L,t) = 0, \quad 0 \le t \le T, \tag{2.6}$$

$$u(x,0) = \phi(x), \quad 0 < x < L. \tag{2.7}$$

Next, we discrete the domain $[0,L]\times[0,T]$ with $x_i = ih$ $(0 \le i \le M)$ and $t_n = n\tau$ $(0 \le n \le N)$, where $h = \frac{L}{M}$ and $\tau = \frac{T}{N}$ are space and time step sizes respectively. Then we introduce the following preliminary lemma:

Lemma 2.2. *Suppose*

$$F(\sigma) = \sum_{r=0}^{m} \frac{\lambda_r}{\Gamma(3-\alpha_r)} \sigma^{1-\alpha_r} \left[\sigma - \left(1 - \frac{\alpha_r}{2}\right)\right] \tau^{2-\alpha_r}, \quad \sigma \ge 0.$$

Let $a = \min\limits_{0\le r\le m}\left\{1 - \frac{\alpha_r}{2}\right\}$, $\quad b = \max\limits_{0\le r\le m}\left\{1 - \frac{\alpha_r}{2}\right\}$, *we can obtain that the equation* $F(\sigma) = 0$ *has a unique positive root* $\sigma^* \in [a,b]$, *where*

$$a = 1 - \frac{1}{2}\max_{0\le r\le m}\{\alpha_r\} = 1 - \frac{\alpha_m}{2} = \frac{1}{2}, \quad b = 1 - \frac{1}{2}\min_{0\le r\le m}\{\alpha_r\} = 1 - \frac{\alpha_0}{2} = 1.$$

Proof. The proof is quite similar to Lemma 2.1 in [9] and thererfore is omitted. □

For convenience, we let $\sigma = \sigma^*$, which means that $\sigma \in [\frac{1}{2}, 1]$ satisfies $F(\sigma) = 0$.

Let $t_{n-1+\sigma} = (n-1+\sigma)\tau$, two lemmas are given below that will be useful in the discretizations of the multi-term time-space FDEs later.

Lemma 2.3. *Suppose* $u(t) \in C^3([t_0, t_n])$, *consider the linear combination of multi-term fractional derivatives* $\sum\limits_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(t)$ *at the point* $t = t_{n-1+\sigma}$, *where* λ_r $(r = 0,1,2,\cdots,m) > 0$, $0 \le \alpha_0 < \alpha_1 < \cdots < \alpha_m \le 1$ *and at least one of* α_i*'s belongs to* $(0,1)$. *The second-order accurate interpolation approximation for the* $\sum\limits_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(t)$ *is as follows:*

$$\sum_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(t_{n-1+\sigma}) = \sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left[u(t_{n-k}) - u(t_{n-k-1})\right] + \mathcal{O}(\tau^{3-\alpha_m}),$$

where
$$\hat{c}_k^{(n)} = \sum_{r=0}^{m} \lambda_r \frac{\tau^{-\alpha_r}}{\Gamma(2-\alpha_r)} c_k^{(n,\alpha_r)},$$
in which $c_0^{(n,\alpha_r)} = a_0^{(\alpha_r)}$*, when* $n = 1$*;*

For $n \geq 2$*, we have*
$$c_k^{(n,\alpha_r)} = \begin{cases} a_0^{(\alpha_r)} + b_1^{(\alpha_r)}, & k = 0, \\ a_k^{(\alpha_r)} + b_{k+1}^{(\alpha_r)} - b_k^{(\alpha_r)}, & 1 \leq k \leq n-2, \\ a_k^{(\alpha_r)} - b_k^{(\alpha_r)}, & k = n-1, \end{cases}$$
where
$$a_0^{\alpha_r} = \sigma^{1-\alpha_r};\ a_l^{\alpha_r} = (l+\sigma)^{1-\alpha_r} - (l-1+\sigma)^{1-\alpha_r}\ (l \geq 1),$$
$$b_l^{\alpha_r} = \frac{1}{2-\alpha_r}\left[(l+\sigma)^{2-\alpha_r} - (l-1+\sigma)^{2-\alpha_r}\right] - \frac{1}{2}\left[(l+\sigma)^{1-\alpha_r} + (l-1+\sigma)^{1-\alpha_r}\right].$$
In particular, when $\alpha_r = 1$*, we have* $c_0^{(n,\alpha_r)} = 1,\ c_k^{(n,\alpha_r)} = 0\ (1 \leq k \leq n-1)$*; when* $\alpha_r = 0$*, we have* $c_0^{(n,\alpha_r)} = \sigma,\ c_k^{(n,\alpha_r)} = 1\ (1 \leq k \leq n-1)$*.*

Proof. For a rigorous proof of this lemma, the reader is referred to [9]. □

Lemma 2.4 ([37]). *Suppose that* $u(x) \in C^5[0,L]$ *satisfy the boundary condition* $u(0) = u(L) = 0$*. The fractional centred difference formula for approximating the Riesz derivatives when* $1 < \beta \leq 2$ *is as follows:*
$$\frac{\partial^\beta u(x_i)}{\partial |x|^\beta} = -h^{-\beta} \sum_{k=i-M}^{i} g_k^{(\beta)} u(x_{i-k}) + \mathcal{O}(h^2),$$
where
$$g_k^{(\beta)} = \frac{(-1)^k \Gamma(\beta+1)}{\Gamma(\beta/2 - k + 1)\Gamma(\beta/2 + k + 1)}.$$

Assume that $u(x,t) \in C^{(5,3)}([0,L] \times [0,T])$ is a solution to the problem (2.1)-(2.3). Consider the equation (2.5) at $(x_i, t_{n-1+\sigma})$, and we get
$$\sum_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(x_i, t_{n-1+\sigma}) = K \frac{\partial^\beta u(x_i, t_{n-1+\sigma})}{\partial |x|^\beta} + f(x_i, t_{n-1+\sigma}), \tag{2.8}$$
where $1 \leq i \leq M-1,\ 1 \leq n \leq N$. For simplicity, we define
$$U_i^n = u(x_i, t_n), \quad 0 \leq i \leq M,\ 0 \leq n \leq N;$$
$$f_i^{n-1+\sigma} = f(x_i, t_{n-1+\sigma}), \quad 0 \leq i \leq M,\ 1 \leq n \leq N.$$
Using Lemma 2.3, we have
$$\sum_{r=0}^{m} \lambda_r {}_0^C D_t^{\alpha_r} u(x_i, t_{n-1+\sigma}) = \sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left(U_i^{n-k} - U_i^{n-k-1}\right) + \mathcal{O}(\tau^{3-\alpha_m}). \tag{2.9}$$

By applying the second-order linear interpolation formula to the Riesz derivative on the right side of equation (2.8), we obtain that

$$\frac{\partial^\beta u(x_i,t_{n-1+\sigma})}{\partial|x|^\beta}=\sigma\frac{\partial^\beta u(x_i,t_n)}{\partial|x|^\beta}+(1-\sigma)\frac{\partial^\beta u(x_i,t_{n-1})}{\partial|x|^\beta}+\mathcal{O}(\tau^2). \tag{2.10}$$

Furthermore, based on Lemma 2.4, we have

$$\frac{\partial^\beta u(x_i,t_n)}{\partial|x|^\beta}=-h^{-\beta}\sum_{k=i-M}^{i}g_k^{(\beta)}U_{i-k}^n+\mathcal{O}(h^2). \tag{2.11}$$

Combine formulae (2.10) and (2.11), and we get

$$\frac{\partial^\beta u(x_i,t_{n-1+\sigma})}{\partial|x|^\beta}=-h^{-\beta}\sum_{k=i-M}^{i}g_k^{(\beta)}\left[\sigma U_{i-k}^n+(1-\sigma)U_{i-k}^{n-1}\right]+\mathcal{O}(h^2+\tau^2). \tag{2.12}$$

By substituting (2.9) and (2.12) into (2.8), we obtain

$$\begin{aligned}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(U_i^{n-k}-U_i^{n-k-1}\right)=&-Kh^{-\beta}\sum_{k=i-M}^{i}g_k^{(\beta)}\left[\sigma U_{i-k}^n+(1-\sigma)U_{i-k}^{n-1}\right]\\&+f_i^{n-1+\sigma}+R_i^n,\end{aligned} \tag{2.13}$$

where there exists a positive constant c_1 such that

$$|R_i^n|\le c_1\left(h^2+\tau^2+\Delta\alpha^2\right),\quad 1\le i\le M-1,\ 1\le n\le N. \tag{2.14}$$

Notice the initial-boundary conditions (2.6)-(2.7). We have

$$U_0^n=0,\quad U_M^n=0,\quad 0\le n\le N, \tag{2.15}$$

$$U_i^0=\phi(x_i),\quad 1\le i\le M-1. \tag{2.16}$$

Suppose u_i^k is the numerical approximation to $u(x_i,t_k)$. By omitting the local truncation error term R_i^n in (2.13) and replacing the exact solution U_i^n with u_i^k in (2.13), (2.15)-(2.16), we can construct the following difference scheme for the (2.1)-(2.3):

$$\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(u_i^{n-k}-u_i^{n-k-1}\right)=-Kh^{-\beta}\sum_{k=i-M}^{i}g_k^{(\beta)}\left[\sigma u_{i-k}^n+(1-\sigma)u_{i-k}^{n-1}\right]+f_i^{n-1+\sigma},$$

$$1\le i\le M-1,\ 1\le n\le N, \tag{2.17}$$

$$u_0^n=0,\quad u_M^n=0,\quad 0\le n\le N, \tag{2.18}$$

$$u_i^0=\phi(x_i),\quad 1\le i\le M-1. \tag{2.19}$$

2.2. Solvability, stability and convergence analysis

In this subsection, we analyze the unique solvability, unconditional stability and convergence of the difference scheme (2.17)-(2.19) obtained in Section 2.1. Meanwhile, we show that the convergence orders of the proposed difference scheme are two in space, time and distributed-order integral.

We define $V_h = \{v \mid v = (v_0, v_1, \cdots, v_{M-1}, v_M)^T,\ v_0 = 0,\ v_M = 0\}$. For all $v,\ w \in V_h$, the discrete inner product and the corresponding discrete L_2-norm are defined as follows:

$$(v, w) = h \sum_{i=1}^{M-1} v_i w_i, \quad \text{and} \quad \|v\| = \sqrt{(v, v)}.$$

Before introducing the properties on the solvability, unconditional stability and convergence, several useful lemmas are prepared below.

Lemma 2.5 ([37]). *Let $1 < \beta \leq 2$ and take $g_k^{(\beta)}$ as defined in Lemma 2.4. We have*

$$\begin{cases} g_0^{(\beta)} = \frac{\Gamma(\beta+1)}{\Gamma^2(\beta/2+1)} \geq 0, \quad g_{-k}^{(\beta)} = g_k^{(\beta)} \leq 0, \quad k = 1, 2, \cdots, \\ \sum\limits_{k=-\infty}^{\infty} g_k^{(\beta)} = 0, \quad -\sum\limits_{\substack{k=-M+i \\ k \neq 0}}^{i} g_k^{(\beta)} \leq g_0^{(\beta)}, \quad 1 \leq i \leq M-1, \\ g_k^{(\beta)} = \left(1 - \frac{\beta+1}{\beta/2+k}\right) g_{k-1}^{(\beta)}, \quad k \geq 1. \end{cases}$$

Lemma 2.6 ([9]). *Let $\hat{c}_k^{(n)} = \sum\limits_{r=0}^{m} \lambda_r \frac{\tau^{-\alpha_r}}{\Gamma(2-\alpha_r)} c_k^{(n,\alpha_r)}$, $k = 0, 1, \cdots, n-1$, as is defined in Lemma 2.3, it holds*

$$\hat{c}_0^{(n)} > \hat{c}_1^{(n)} > \cdots > \hat{c}_{n-2}^{(n)} > \hat{c}_{n-1}^{(n)} > \sum_{r=0}^{m} \lambda_r \frac{\tau^{-\alpha_r}}{\Gamma(2-\alpha_r)} \cdot \frac{1-\alpha_r}{2} (n-1+\sigma)^{-\alpha_r}.$$

Lemma 2.7 ([1]). *Let V represent the inner product space and $(\cdot, \cdot)$ denote the inner product with the induced norm $\|\cdot\|$. For $v^0,\ v^1, \cdots,\ v^n \in V$, when $n \geq 1$ we have*

$$\sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left(v^{n-k} - v^{n-k-1},\ \sigma v^n + (1-\sigma) v^{n-1}\right) \geq \frac{1}{2} \sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left(\|v^{n-k}\|^2 - \|v^{n-k-1}\|^2\right).$$

Lemma 2.8 ([35]). *For $1 < \beta \leq 2$ and any $v \in V_h$, it holds that*

$$-h^{-\beta} h \sum_{i=1}^{M-1} \left(\sum_{k=i-M}^{i} g_k^{(\beta)} v_{i-k} \right) v_i \leq -c_*^{(\beta)} (2L)^{-\beta} h \sum_{i=1}^{M-1} v_i^2,$$

where $c_*^{(\beta)} = \frac{2}{\beta} r_\beta$, with $r_\beta = e^{-2\frac{(4-\beta)(2-\beta)\beta}{(6+\beta)(4+\beta)(2+\beta)}} \cdot \frac{\Gamma(\beta+1)}{\Gamma^2(\beta/2+1)} \left(3 + \frac{\beta}{2}\right)^{\beta+1}$.

We first consider the unique solvability of the numerical method (2.17)-(2.19).

Theorem 2.1. *The difference scheme* (2.17)-(2.19) *is uniquely solvable.*

Proof. Let $u^n = (u_0^n, u_1^n, u_2^n, \cdots, u_{M-1}^n, u_M^n)^T$. According to (2.18) and (2.19), the value of u^0 is determined. Now suppose that $\{u^k \mid 0 \leq k \leq n-1\}$ has been determined. According to (2.17) and (2.18), we get a linear equation system with respect to u^n. Then we only need to prove that the corresponding homogeneous linear system

$$\hat{c}_0^{(n)} u_i^n = -K\sigma h^{-\beta} \sum_{k=i-M}^{i} g_k^{(\beta)} u_{i-k}^n, \quad 1 \leq i \leq M-1, \tag{2.20}$$

$$u_0^n = 0, \quad u_M^n = 0 \tag{2.21}$$

only has solution of 0.

We first rewrite the equation (2.20) as follows:

$$\left[\hat{c}_0^{(n)} + K\sigma h^{-\beta} g_0^{(\beta)}\right] u_i^n = K\sigma h^{-\beta} \sum_{\substack{k=i-M \\ k\neq 0}}^{i} \left(-g_k^{(\beta)}\right) u_{i-k}^n, \quad 1 \le i \le M-1. \tag{2.22}$$

Let $\|u^n\|_\infty = |u_{i_n}^n|$, where $i_n \in \{1, 2, \cdots, M-1\}$. Let us consider equation (2.22) with $i = i_n$ and take absolute values on both sides of the equation. Based on Lemma 2.5 and the fact that the coefficients $K > 0$, it can be seen that

$$\begin{aligned}\left[\hat{c}_0^{(n)} + K\sigma h^{-\beta} g_0^{(\beta)}\right] \|u^n\|_\infty \le & K\sigma h^{-\beta} \sum_{\substack{k=i_n-M \\ k\neq 0}}^{i_n} \left(-g_k^{(\beta)}\right) |u_{i_n-k}^n| \\ \le & K\sigma h^{-\beta} \sum_{\substack{k=i_n-M \\ k\neq 0}}^{i_n} \left(-g_k^{(\beta)}\right) \|u^n\|_\infty \\ \le & K\sigma h^{-\beta} g_0^{(\beta)} \|u^n\|_\infty.\end{aligned}$$

Therefore, $\|u^n\|_\infty = 0$ is derived, which indicates that the homogeneous linear equations (2.20)-(2.21) have a single solution of 0. □

We are now going to prove the unconditional stability of the difference scheme (2.17)-(2.19) with respect to the initial value and the inhomogeneous term $f(x,t)$. The correlation result is shown in the following theorem.

Theorem 2.2. *Let* $\{u_i^n \mid 0 \le i \le M,\ 0 \le n \le N\}$ *be the solution of the difference scheme* (2.17)-(2.19). *We have*

$$\|u^n\|^2 \le \|u^0\|^2 + \frac{(2L)^\beta}{Kc_*^{(\beta)} \sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \max_{1\le l\le n} \|f^{l-1+\sigma}\|^2, \quad 1 \le n \le N,$$

where $\|f^{l-1+\sigma}\|^2 = h \sum\limits_{i=1}^{M-1} \left(f_i^{l-1+\sigma}\right)^2$.

Proof. Multiplying (2.17) by $h(\sigma u_i^n + (1-\sigma)u_i^{n-1})$ and summing up with i from 1 to $M-1$, we get

$$\begin{aligned}&\sum_{k=0}^{n-1} \hat{c}_k^{(n)} h \sum_{i=1}^{M-1} \left(u_i^{n-k} - u_i^{n-k-1}\right) \left[\sigma u_i^n + (1-\sigma)u_i^{n-1}\right] \\ =& -Kh^{-\beta} h \sum_{i=1}^{M-1} \sum_{k=i-M}^{i} g_k^{(\beta)} \left[\sigma u_{i-k}^n + (1-\sigma)u_{i-k}^{n-1}\right] \left[\sigma u_i^n + (1-\sigma)u_i^{n-1}\right] \\ &+ h \sum_{i=1}^{M-1} f_i^{n-1+\sigma} \left[\sigma u_i^n + (1-\sigma)u_i^{n-1}\right], \quad 1 \le n \le N.\end{aligned} \tag{2.23}$$

According to Lemma 2.7, it follows that

$$\sum_{k=0}^{n-1} \hat{c}_k^{(n)} h \sum_{i=1}^{M-1} \left(u_i^{n-k} - u_i^{n-k-1}\right) \left[\sigma u_i^n + (1-\sigma)u_i^{n-1}\right]$$

$$
\begin{aligned}
&=\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(u^{n-k}-u^{n-k-1},\sigma u^n+(1-\sigma)u^{n-1}\right)\\
&\geq\frac{1}{2}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(\|u^{n-k}\|^2-\|u^{n-k-1}\|^2\right). \qquad (2.24)
\end{aligned}
$$

Using Lemma 2.8, we obtain

$$
\begin{aligned}
&-Kh^{-\beta}h\sum_{i=1}^{M-1}\sum_{k=i-M}^{i}g_k^{(\beta)}\left[\sigma u_{i-k}^n+(1-\sigma)u_{i-k}^{n-1}\right]\left[\sigma u_i^n+(1-\sigma)u_i^{n-1}\right]\\
\leq&-Kc_*^{(\beta)}(2L)^{-\beta}\|\sigma u^n+(1-\sigma)u^{n-1}\|^2. \qquad (2.25)
\end{aligned}
$$

In addition, by exploiting Cauchy-Schwarz inequality, we can get

$$
\begin{aligned}
&h\sum_{i=1}^{M-1}f_i^{n-1+\sigma}\left[\sigma u_i^n+(1-\sigma)u_i^{n-1}\right]\\
\leq&\|f^{n-1+\sigma}\|\cdot\|\sigma u^n+(1-\sigma)u^{n-1}\|\\
\leq&Kc_*^{(\beta)}(2L)^{-\beta}\|\sigma u^n+(1-\sigma)u^{n-1}\|^2+\frac{(2L)^\beta}{4Kc_*^{(\beta)}}\|f^{n-1+\sigma}\|^2. \qquad (2.26)
\end{aligned}
$$

By substituting (2.24)-(2.26) into (2.23), we have

$$
\frac{1}{2}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(\|u^{n-k}\|^2-\|u^{n-k-1}\|^2\right)\leq\frac{(2L)^\beta}{4Kc_*^{(\beta)}}\|f^{n-1+\sigma}\|^2,\quad 1\leq n\leq N. \qquad (2.27)
$$

With the use of Lemma 2.6, we get

$$
\hat{c}_{n-1}^{(n)}\geq\sum_{r=0}^{m}\lambda_r\frac{\tau^{-\alpha_r}}{\Gamma(2-\alpha_r)}\cdot\frac{1-\alpha_r}{2}(n-1+\sigma)^{-\alpha_r}\geq\frac{1}{2}\sum_{r=0}^{m}\frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}. \qquad (2.28)
$$

Combine (2.27) and (2.28), and we arrives at the following inequality:

$$
\begin{aligned}
\hat{c}_0^{(n)}\|u^n\|^2\leq&\sum_{k=1}^{n-1}\left(\hat{c}_{k-1}^{(n)}-\hat{c}_k^{(n)}\right)\|u^{n-k}\|^2+\hat{c}_{n-1}^{(n)}\|u^0\|^2+\frac{(2L)^\beta}{2Kc_*^{(\beta)}}\parallel f^{n-1+\sigma}\parallel^2\\
\leq&\sum_{k=1}^{n-1}\left(\hat{c}_{k-1}^{(n)}-\hat{c}_k^{(n)}\right)\|u^{n-k}\|^2+\hat{c}_{n-1}^{(n)}\left(\|u^0\|^2+\frac{(2L)^\beta\|f^{n-1+\sigma}\|^2}{Kc_*^{(\beta)}\sum\limits_{r=0}^{m}\frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}\right),
\end{aligned}
$$

where $1\leq n\leq N$. By applying the mathematical induction method to the above inequality, we can get

$$
\|u^n\|^2\leq\|u^0\|^2+\frac{(2L)^\beta}{Kc_*^{(\beta)}\sum\limits_{r=0}^{m}\frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}\max_{1\leq l\leq n}\|f^{l-1+\sigma}\|^2,\quad 1\leq n\leq N.
$$

This completes the proof. $\square$

We have established the unconditional stability of the difference scheme (2.17)-(2.19), and now we further show its convergence.

Suppose that $\{U_i^n \mid 0 \le i \le M,\ 0 \le n \le N\}$ is the exact solution of the system (2.1)-(2.3) and $\{u_i^n \mid 0 \le i \le M,\ 0 \le n \le N\}$ is the numerical solution of the difference scheme (2.17)-(2.19). Let $e_i^n = U_i^n - u_i^n$ $(0 \le i \le M,\ 0 \le n \le N)$.

By subtracting (2.17)-(2.19) from (2.13), (2.15)-(2.16), respectively, we obtain the system of error equations as follows:

$$\begin{aligned}
&\sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left(e_i^{n-k} - e_i^{n-k-1}\right) = -Kh^{-\beta} \sum_{k=i-M}^{i} g_k^{(\beta)} \left[\sigma e_{i-k}^n + (1-\sigma) e_{i-k}^{n-1}\right] + R_i^n, \\
&\qquad\qquad\qquad\qquad 1 \le i \le M-1,\ 1 \le n \le N, \\
&e_0^n = 0, \quad e_M^n = 0, \quad 0 \le n \le N, \\
&e_i^0 = 0, \quad 1 \le i \le M-1.
\end{aligned}$$

By applying the conclusion of Theorem 2.2 and noticing (2.14), we have

$$\|e^n\|^2 \le \frac{(2L)^\beta \max\limits_{1\le l\le n} \|R^l\|^2}{Kc_*^{(\beta)} \sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \le \frac{(2L)^\beta}{Kc_*^{(\beta)} \sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \left[c_1 \left(h^2 + \tau^2 + \Delta\alpha^2\right)\right]^2 L,$$

where $1 \le n \le N$. Extract the square root on both sides of the equation above, we get

$$\|e^n\| \le c_1 \sqrt{\frac{2^\beta L^{\beta+1}}{Kc_*^{(\beta)} \sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}} \left(h^2 + \tau^2 + \Delta\alpha^2\right), \quad 1 \le n \le N.$$

Therefore, we can get the following theorem.

Theorem 2.3. *Suppose that the continuous problem* (2.1)-(2.3) *has a smooth solution* $u(x,t) \in C^{(5,3)}(\Omega \times [0,T])$, *and let* u_i^n *be the solution of the difference scheme* (2.17)-(2.19). *It holds that*

$$\|e^n\| \le c_1 \sqrt{\frac{2^\beta L^{\beta+1}}{Kc_*^{(\beta)} \sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}} \left(h^2 + \tau^2 + \Delta\alpha^2\right), \quad 1 \le n \le N.$$

2.3. Fast solution techniques with circulant preconditioner

We rewrite the proposed implicit difference scheme (2.17) as the following matrix form at the time level n:

$$A^n u^n = b^{n-1}, \quad n = 1, 2, \ldots, N, \tag{2.29}$$

where

$$A^n = \hat{c}_0^{(n)} I + \sigma K h^{-\beta} G, \tag{2.30}$$

and

$$b^{n-1} = -(1-\sigma)Kh^{-\beta}Gu^{n-1} + \sum_{k=1}^{n-1}(\hat{c}_{k-1}^{(n)} - \hat{c}_k^{(n)})u^{n-k} + \hat{c}_{n-1}^{(n)}u^0 + f^{n-1+\sigma}.$$

Here I is the identity matrix of order $M-1$ and

$$G = \begin{bmatrix} g_0^{(\beta)} & g_{-1}^{(\beta)} & g_{-2}^{(\beta)} & \cdots & g_{3-M}^{(\beta)} & g_{2-M}^{(\beta)} \\ g_1^{(\beta)} & g_0^{(\beta)} & g_{-1}^{(\beta)} & \cdots & g_{4-M}^{(\beta)} & g_{3-M}^{(\beta)} \\ g_2^{(\beta)} & g_1^{(\beta)} & g_0^{(\beta)} & \cdots & g_{5-M}^{(\beta)} & g_{4-M}^{(\beta)} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ g_{M-3}^{(\beta)} & g_{M-4}^{(\beta)} & g_{M-5}^{(\beta)} & \cdots & g_0^{(\beta)} & g_{-1}^{(\beta)} \\ g_{M-2}^{(\beta)} & g_{M-3}^{(\beta)} & g_{M-4}^{(\beta)} & \cdots & g_1^{(\beta)} & g_0^{(\beta)} \end{bmatrix}. \tag{2.31}$$

It is obvious that G is a symmetric Toeplitz matrix (see [22]). Therefore, it can be stored with only $M-1$ entries and the FFT can be used to carry out the matrix-vector product in only $\mathcal{O}((M-1)\log(M-1))$ operations.

The following lemma guarantees the invertibility of the matrix A^n defined in (2.30).

Lemma 2.9. *The coefficient matrix*

$$A^n = \hat{c}_0^{(n)}I + \sigma Kh^{-\beta}G$$

of the linear system (2.29) *is symmetric positive definite.*

Proof. Let a_{ij}^n be the (i,j) entry of the A^n. We notice Lemma 2.5 and $\hat{c}_0^{(n)} > 0$, thus

$$\begin{aligned} |a_{ii}^n| - \sum_{j=1,j\neq i}^{M-1}|a_{ij}^n| =& (\hat{c}_0^{(n)} + \sigma Kh^{-\beta}g_0^{(\beta)}) - \sigma Kh^{-\beta}(\sum_{j=-M+i+1,j\neq 0}^{i-1}|g_j^{(\beta)}|) \\ =& \hat{c}_0^{(n)} + \sigma Kh^{-\beta}\sum_{j=-M+i+1}^{i-1}g_j^{(\beta)} \\ >& \hat{c}_0^{(n)} > 0. \end{aligned}$$

This implies that A^n is a strictly diagonally dominant matrix. According to Lemma 2.5, it is easy to prove that A^n is symmetric and all of its main diagonal elements are positive. Hence, all its eigenvalues are positive. □

It is well-known that the conjugate gradient (CG) method is a popular and effective Krylov subspace method [22] for solving symmetric positive systems with Toeplitz coefficient matrix. Nevertheless, the drawback of the CG method is its slow convergence when the eigenvalues of the coefficient matrix A^n are not clustered [6]. To overtake this shortcoming, we use the preconditioned CG method (PCG) to solve such linear systems [22].

We propose a circulant preconditioner, which is generated from the famous R. Chan's circulant preconditioner [5] to solve the Toeplitz linear system (2.29). For a Toeplitz matrix $G_n \in \mathbb{C}^{n\times n}$ with form of (2.31), the R. Chan's circulant preconditioner R_n makes use of all the entries [5]. Its entries $r_{ij} = r_{i-j}$ are given by

$$r_k = \begin{cases} g_0, & k = 0, \\ g_k + g_{k-n}, & 0 < k < n \\ r_{k+n}, & 0 < -k < n. \end{cases}$$

Then the PCG method is employed to solve the following preconditioned system

$$(C^n)^{-1}A^n u^n = (C^n)^{-1}b^{n-1}, \quad n = 1, 2, \dots, N,$$

and the R. Chan's-based circulant preconditioner C^n takes the following form

$$C^n = \hat{c}_0^{(n)} I + \sigma K h^{-\beta} c(G).$$

More precisely, the first column of $c(G)$ is given by

$$\begin{pmatrix} g_0^{(\beta)} \\ g_1^{(\beta)} + g_{2-M}^{(\beta)} \\ g_2^{(\beta)} + g_{3-M}^{(\beta)} \\ \vdots \\ \vdots \\ g_{M-3}^{(\beta)} + g_{-2}^{(\beta)} \\ \\ g_{M-2}^{(\beta)} + g_{-1}^{(\beta)} \\ \end{pmatrix}.$$

Below we discuss the basic properties of the circulant preconditioner C^n.

Lemma 2.10. *The circulant preconditioner*

$$C^n = \hat{c}_0^{(n)} I + \sigma K h^{-\beta} c(G)$$

is a symmetric positive definite matrix.

Proof. As similar to Lemma 2.9, suppose c_{ij}^n be the (i,j) entry of C^n. Based on Lemma 2.5 and $\hat{c}_0^{(n)} > 0$ we get

$$\begin{aligned} |c_{ii}^n| - \sum_{j=1, j\neq i}^{M-1} |c_{ij}^n| =& (\hat{c}_0^{(n)} + \sigma K h^{-\beta} g_0^{(\beta)}) - \sigma K h^{-\beta} (\sum_{j=1}^{M-2} |g_j^{(\beta)} + g_{-j}^{(\beta)}|) \\ =& \hat{c}_0^{(n)} + \sigma K h^{-\beta} \sum_{j=2-M}^{M-2} g_j^{(\beta)} \\ >& \hat{c}_0^{(n)} > 0, \end{aligned}$$

which implies that C^n is a strictly diagonally dominant matrix. From Lemma 2.5, we can easily know that the main diagonal elements of C^n are positive and C^n is symmetric. Therefore, C^n is a symmetric positive definite matrix. □

Lemma 2.10 suggests that C^n is invertible. In addition, the eigenvalue distributions of preconditioned matrices $(C^n)^{-1}A^n$ are theoretically proven to be clustered around 1 [5]. The convergence rate of PCG is superlinear [6]. We will numerically demonstrate in Section 4 that the circulant preconditioner has good clustering eigenvalues. It is both numerically and theoretically guaranteed that the computational cost per iteration of PCG is $\mathcal{O}((M-1)\log(M-1))$ and the total cost at each time step is $\mathcal{O}((M-1)\log(M-1))$.

3. 2D problem

Consider the following 2D TDRFDEs:

$$D_t^{\omega(\alpha)}u(x,y,t) = K_1\frac{\partial^\beta u(x,y,t)}{\partial|x|^\beta} + K_2\frac{\partial^\gamma u(x,y,t)}{\partial|y|^\gamma} + f(x,y,t),$$
$$(x,y)\in\Omega,\ 0<t\le T, \tag{3.1}$$
$$u(x,y,t) = 0, \quad (x,y)\in\partial\Omega,\ 0\le t\le T, \tag{3.2}$$
$$u(x,y,0) = \phi(x,y), \quad (x,y)\in\Omega, \tag{3.3}$$

where $\Omega = (0,L_1)\times(0,L_2)$, $\partial\Omega$ is the boundary of Ω, $f(x,y,t)$ and $\phi(x,y)$ are given functions. Especially, $\phi(x,y)=0$ holds when $(x,y)\in\partial\Omega$.

In this section, we can directly extend the idea for solving the 1D problem (2.1)-(2.3) to handle the 2D problem (3.1)-(3.3). We propose a second-order difference scheme based on the interpolation approximation on a special point to solve the 2D TDRFDEs. The unique solvability, unconditional stability and convergence of the proposed difference scheme are also discussed. Furthermore, a multilevel circulant preconditioner is proposed to accelerate the convergence rate of the Krylov subspace method.

3.1. Numerical discretization for (3.1)-(3.3)

To derive the difference scheme of (3.1)-(3.3), we first divide the interval $[0,L_1]$ into M_1-subintervals with $h_1 = \frac{L_1}{M_1}$ and $x_i = ih_1$ $(0\le i\le M_1)$, and divide the interval $[0,L_2]$ into M_2-subintervals with $h_2=\frac{L_2}{M_2}$ and $y_j = jh_2$ $(0\le j\le M_2)$.

Denote $\omega = \{(i,j)\mid 1\le i\le M_1-1,\ 1\le j\le M_2-1\}$, $\partial\omega = \{(i,j)\mid (x_i,y_j)\in\partial\Omega\}$, $\bar\omega = \omega\bigcup\partial\omega$. We define

$$U_{ij}^n = u(x_i,y_j,t_n), \quad (i,j)\in\bar\omega,\ 0\le n\le N;$$
$$f_{ij}^{n-1+\sigma} = f(x_i,y_j,t_{n-1+\sigma}), \quad (i,j)\in\bar\omega,\ 1\le n\le N.$$

Suppose $u(x,y,t)\in C^{(5,5,3)}(\Omega\times[0,T])$. Considering (3.1) at the point $(x_i,y_j,t_{n-1+\sigma})$, we have

$$D_t^{\omega(\alpha)}u(x_i,y_j,t_{n-1+\sigma}) = K_1\frac{\partial^\beta u(x_i,y_j,t_{n-1+\sigma})}{\partial|x|^\beta} + K_2\frac{\partial^\gamma u(x_i,y_j,t_{n-1+\sigma})}{\partial|y|^\gamma} + f_{ij}^{n-1+\sigma}, \tag{3.4}$$

where $(i,j)\in\omega,\ 1\le n\le N$. Using Lemma 2.1 and Lemma 2.3, we get

$$D_t^{\omega(\alpha)}u(x_i,y_j,t_{n-1+\sigma})=\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(U_{ij}^{n-k}-U_{ij}^{n-k-1}\right)+\mathcal{O}\left(\tau^{3-\alpha_m}+\Delta\alpha^2\right). \quad (3.5)$$

Moreover, by applying the second-order linear interpolation formula to the Riesz derivative on the right side of (3.4) and using Lemma 2.4, we obtain

$$\frac{\partial^\beta u(x_i,y_j,t_{n-1+\sigma})}{\partial|x|^\beta}=-h_1^{-\beta}\sum_{k=i-M_1}^{i}g_k^{(\beta)}\left[\sigma U_{i-k,j}^n+(1-\sigma)U_{i-k,j}^{n-1}\right]+\mathcal{O}(h_1^2+\tau^2), \quad (3.6)$$

and

$$\frac{\partial^\gamma u(x_i,y_j,t_{n-1+\sigma})}{\partial|y|^\gamma}=-h_2^{-\gamma}\sum_{k=j-M_2}^{j}g_k^{(\gamma)}\left[\sigma U_{i,j-k}^n+(1-\sigma)U_{i,j-k}^{n-1}\right]+\mathcal{O}(h_2^2+\tau^2). \quad (3.7)$$

By substituting (3.5)-(3.7) into (3.4), we can get

$$\begin{aligned}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(U_{ij}^{n-k}-U_{ij}^{n-k-1}\right)=&-K_1h_1^{-\beta}\sum_{k=i-M_1}^{i}g_k^{(\beta)}\left[\sigma U_{i-k,j}^n+(1-\sigma)U_{i-k,j}^{n-1}\right]\\&-K_2h_2^{-\gamma}\sum_{k=j-M_2}^{j}g_k^{(\gamma)}\left[\sigma U_{i,j-k}^n+(1-\sigma)U_{i,j-k}^{n-1}\right]\\&+f_{ij}^{n-1+\sigma}+S_{ij}^n,\quad (i,j)\in\omega,\ 1\le n\le N, \end{aligned}\quad (3.8)$$

where there exists a positive constant c_2 such that

$$\mid S_{ij}^n\mid\le c_2\left(h_1^2+h_2^2+\tau^2+\Delta\alpha^2\right),\quad (i,j)\in\omega,\ 1\le n\le N. \quad (3.9)$$

Notice the initial and boundary conditions (3.2)-(3.3), and we have

$$U_{ij}^n=0,\quad (i,j)\in\partial\omega,\ 0\le n\le N, \quad (3.10)$$
$$U_{ij}^0=\phi(x_i,y_j),\quad (i,j)\in\omega. \quad (3.11)$$

Thus, by neglecting the small term S_{ij}^n in (3.8) and replacing the exact solution U_{ij}^n with the numerical ones u_{ij}^k in (3.8) and (3.10)-(3.11), we can get the difference scheme for solving (3.1)-(3.3) as follows:

$$\begin{aligned}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(u_{ij}^{n-k}-u_{ij}^{n-k-1}\right)=&-K_1h_1^{-\beta}\sum_{k=i-M_1}^{i}g_k^{(\beta)}\left[\sigma u_{i-k,j}^n+(1-\sigma)u_{i-k,j}^{n-1}\right]\\&-K_2h_2^{-\gamma}\sum_{k=j-M_2}^{j}g_k^{(\gamma)}\left[\sigma u_{i,j-k}^n+(1-\sigma)u_{i,j-k}^{n-1}\right]\\&+f_{ij}^{n-1+\sigma},\quad (i,j)\in\omega,\ 1\le n\le N, \end{aligned}\quad (3.12)$$

$$u_{ij}^n=0,\quad (i,j)\in\partial\omega,\ 0\le n\le N, \quad (3.13)$$
$$u_{ij}^0=\phi(x_i,y_j),\quad (i,j)\in\omega. \quad (3.14)$$

3.2. Solvability, stability and convergence analysis

In this subsection, we show that the difference scheme (3.12)-(3.14) is uniquely solvable, unconditionally stable and convergent with the order of $\mathcal{O}(h_1^2 + h_2^2 + \tau^2 + \Delta\alpha^2)$.

Let $V_h = \{v \mid v = \{v_{ij} \mid (i,j) \in \bar{\omega}\}\}$, $\hat{V}_h = \{v \mid v \in V_h;\ v_{ij} = 0\ when\ (i,j) \in \partial\omega\}$. For any $v,\ w \in \hat{V}_h$, the discrete inner product and the corresponding discrete L_2-norms are defined as follows:

$$(v,w) = h_1 h_2 \sum_{i=1}^{M_1-1} \sum_{j=1}^{M_2-1} v_{ij} w_{ij}, \quad \text{and} \quad \|v\| = \sqrt{(v,v)}.$$

We now work towards showing the unique solvability of difference scheme (3.12)-(3.14). The desired result is reported by the following theorem.

Theorem 3.1. *The difference scheme* (3.12)-(3.14) *is uniquely solvable.*

Proof. Let $u^n = \{u_{ij}^n \mid (i,j) \in \bar{\omega}\}$. According to (3.13)-(3.14), the value of u^0 is determined. Now suppose that $\{u^k \mid 0 \le k \le n-1\}$ has been determined. According to (3.12) and (3.13), we get a linear equation system with respect to u^n. Then we only need to prove that the corresponding homogeneous linear system

$$\hat{c}_0^{(n)} u_{ij}^n = -K_1 \sigma h_1^{-\beta} \sum_{k=i-M_1}^{i} g_k^{(\beta)} u_{i-k,j}^n - K_2 \sigma h_2^{-\gamma} \sum_{k=j-M_2}^{j} g_k^{(\gamma)} u_{i,j-k}^n,\ (i,j) \in \omega, \tag{3.15}$$

$$u_{ij}^n = 0, \quad (i,j) \in \partial\omega \tag{3.16}$$

only has solution of 0.

We first rewrite the equation (3.15) as follows:

$$\begin{aligned}
&\left[\hat{c}_0^{(n)} + K_1 \sigma h_1^{-\beta} g_0^{(\beta)} + K_2 \sigma h_2^{-\gamma} g_0^{(\gamma)}\right] u_{ij}^n \\
=&K_1 \sigma h_1^{-\beta} \sum_{\substack{k=i-M_1 \\ k\neq 0}}^{i} \left(-g_k^{(\beta)}\right) u_{i-k,j}^n + K_2 \sigma h_2^{-\gamma} \sum_{\substack{k=j-M_2 \\ k\neq 0}}^{j} \left(-g_k^{(\gamma)}\right) u_{i,j-k}^n.
\end{aligned} \tag{3.17}$$

Let $\|u^n\|_\infty =\mid u_{i_n,j_n}^n \mid$, where $(i_n, j_n) \in \omega$. We consider the equation (3.17) with $(i,j) = (i_n, j_n)$ and take absolute values on both sides of the equation. Noticing that the coefficients $K_1 > 0,\ K_2 > 0$, based on Lemma 2.5 and using triangle inequality, we have

$$\begin{aligned}
&\left[\hat{c}_0^{(n)} + K_1 \sigma h_1^{-\beta} g_0^{(\beta)} + K_2 \sigma h_2^{-\gamma} g_0^{(\gamma)}\right] \|u^n\|_\infty \\
=&K_1 \sigma h_1^{-\beta} \sum_{\substack{k=i_n-M_1 \\ k\neq 0}}^{i_n} \left(-g_k^{(\beta)}\right) |u_{i_n-k,j_n}^n| + K_2 \sigma h_2^{-\gamma} \sum_{\substack{k=j_n-M_2 \\ k\neq 0}}^{j_n} \left(-g_k^{(\gamma)}\right) |u_{i_n,j_n-k}^n| \\
\le& \left[K_1 \sigma h_1^{-\beta} g_0^{(\beta)} + K_2 \sigma h_2^{-\gamma} g_0^{(\gamma)}\right] \|u^n\|_\infty.
\end{aligned}$$

Therefore, we get $\|u^n\|_\infty = 0$, which indicates that the homogeneous linear equations (3.15)-(3.16) only have solution 0. According to the mathematical induction, the difference scheme (3.12)-(3.14) is uniquely solvable. □

We will now discuss the unconditional stability of the difference scheme (3.12)-(3.14).

Theorem 3.2. *Let* $\{u_{ij}^n \mid (i,j) \in \bar{\omega},\ 0 \le n \le N\}$ *be the solution of the difference scheme* (3.12)-(3.14). *We have*

$$\|u^n\|^2 \le \|u^0\|^2 + \frac{1}{4}\left[\frac{(2L_1)^{\beta}}{K_1 c_*^{(\beta)}} + \frac{(2L_2)^{\gamma}}{K_2 c_*^{(\gamma)}}\right] \frac{1}{\sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \max_{1\le l\le n} \|f^{l-1+\sigma}\|^2, 1 \le n \le N,$$

where $\|f^{l-1+\sigma}\|^2 = h_1 h_2 \sum\limits_{i=1}^{M_1-1}\sum\limits_{j=1}^{M_2-1}\left(f_{ij}^{l-1+\sigma}\right)^2.$

Proof. By multiplying (3.12) by $h_1h_2[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}]$ and summing up (i,j) with respect to ω, we get

$$\begin{aligned}
&\sum_{k=0}^{n-1} \hat{c}_k^{(n)} h_1 h_2 \sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1}\left(u_{ij}^{n-k} - u_{ij}^{n-k-1}\right)\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right] \\
=& - K_1 h_2 \sum_{j=1}^{M_2-1}\Big\{h_1^{-\beta} h_1 \sum_{i=1}^{M_1-1}\sum_{k=i-M_1}^{i} g_k^{(\beta)}\left[\sigma u_{i-k,j}^n + (1-\sigma)u_{i-k,j}^{n-1}\right]\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right]\Big\} \\
&- K_2 h_1 \sum_{i=1}^{M_1-1}\Big\{h_2^{-\gamma} h_2 \sum_{j=1}^{M_2-1}\sum_{k=j-M_2}^{j} g_k^{(\gamma)}\left[\sigma u_{i,j-k}^n + (1-\sigma)u_{i,j-k}^{n-1}\right]\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right]\Big\} \\
&+ h_1 h_2 \sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1} f_{ij}^{n-1+\sigma}\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right].
\end{aligned} \tag{3.18}$$

According to Lemma 2.7, it follows that

$$\begin{aligned}
&\sum_{k=0}^{n-1} \hat{c}_k^{(n)} h_1 h_2 \sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1}\left(u_{ij}^{n-k} - u_{ij}^{n-k-1}\right)\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right] \\
\ge& \frac{1}{2}\sum_{k=0}^{n-1} \hat{c}_k^{(n)}\left(\|u^{n-k}\|^2 - \|u^{n-k-1}\|^2\right).
\end{aligned} \tag{3.19}$$

Using Lemma 2.8, we obtain

$$\begin{aligned}
&- h_1^{-\beta} h_1 \sum_{i=1}^{M_1-1}\sum_{k=i-M_1}^{i} g_k^{(\beta)}\left[\sigma u_{i-k,j}^n + (1-\sigma)u_{i-k,j}^{n-1}\right]\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right] \\
\le& - c_*^{(\beta)}(2L_1)^{-\beta} h_1 \sum_{i=1}^{M_1-1}\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right]^2
\end{aligned} \tag{3.20}$$

and

$$\begin{aligned}
&- h_2^{-\gamma} h_2 \sum_{j=1}^{M_2-1}\sum_{k=j-M_2}^{j} g_k^{(\gamma)}\left[\sigma u_{i,j-k}^n + (1-\sigma)u_{i,j-k}^{n-1}\right]\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right] \\
\le& - c_*^{(\gamma)}(2L_2)^{-\gamma} h_2 \sum_{j=1}^{M_2-1}\left[\sigma u_{ij}^n + (1-\sigma)u_{ij}^{n-1}\right]^2.
\end{aligned} \tag{3.21}$$

By substituting (3.19)-(3.21) into (3.18), we get

$$
\begin{aligned}
&\frac{1}{2}\sum_{k=0}^{n-1}\hat{c}_k^{(n)}\left(\|u^{n-k}\|^2-\|u^{n-k-1}\|^2\right)\\
\leq& -K_1c_*^{(\beta)}(2L_1)^{-\beta}h_1h_2\sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1}\left[\sigma u_{ij}^n+(1-\sigma)u_{ij}^{n-1}\right]^2\\
&-K_2c_*^{(\gamma)}(2L_2)^{-\gamma}h_1h_2\sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1}\left[\sigma u_{ij}^n+(1-\sigma)u_{ij}^{n-1}\right]^2\\
&+h_1h_2\sum_{i=1}^{M_1-1}\sum_{j=1}^{M_2-1}f_{ij}^{n-1+\sigma}\left[\sigma u_{ij}^n+(1-\sigma)u_{ij}^{n-1}\right]\\
\leq& -K_1c_*^{(\beta)}(2L_1)^{-\beta}\|\sigma u^n+(1-\sigma)u^{n-1}\|^2\\
&-K_2c_*^{(\gamma)}(2L_2)^{-\gamma}\|\sigma u^n+(1-\sigma)u^{n-1}\|^2\\
&+\|f^{n-1+\sigma}\|\cdot\|\sigma u^n+(1-\sigma)u^{n-1}\|\\
\leq&\frac{1}{16}\left[\frac{(2L_1)^\beta}{K_1c_*^{(\beta)}}+\frac{(2L_2)^\gamma}{K_2c_*^{(\gamma)}}\right]\|f^{n-1+\sigma}\|^2,\quad 1\leq n\leq N.
\end{aligned}
\tag{3.22}
$$

With the use of Lemma 2.6, we have

$$
\hat{c}_{n-1}^{(n)}\geq\sum_{r=0}^{m}\lambda_r\frac{\tau^{-\alpha_r}}{\Gamma(2-\alpha_r)}\cdot\frac{1-\alpha_r}{2}(n-1+\sigma)^{-\alpha_r}\geq\frac{1}{2}\sum_{r=0}^{m}\frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}. \tag{3.23}
$$

By combining (3.22) and (3.23), we arrive at the following inequality:

$$
\begin{aligned}
&\hat{c}_0^{(n)}\|u^n\|^2\\
\leq&\sum_{k=1}^{n-1}\left(\hat{c}_{k-1}^{(n)}-\hat{c}_k^{(n)}\right)\|u^{n-k}\|^2+\hat{c}_{n-1}^{(n)}\|u^0\|^2+\left[\frac{(2L_1)^\beta}{K_1c_*^{(\beta)}}+\frac{(2L_2)^\gamma}{K_2c_*^{(\gamma)}}\right]\frac{\|f^{n-1+\sigma}\|^2}{8}\\
\leq&\sum_{k=1}^{n-1}\left(\hat{c}_{k-1}^{(n)}-\hat{c}_k^{(n)}\right)\|u^{n-k}\|^2\\
&+\hat{c}_{n-1}^{(n)}\left\{\|u^0\|^2+\frac{1}{4}\left[\frac{(2L_1)^\beta}{K_1c_*^{(\beta)}}+\frac{(2L_2)^\gamma}{K_2c_*^{(\gamma)}}\right]\frac{1}{\sum\limits_{r=0}^{m}\frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}\|f^{n-1+\sigma}\|^2\right\},
\end{aligned}
$$

where $1\leq n\leq N$. Applying the mathematical induction method to the above inequality, we can get the conclusion of Theorem 3.2. This completes the proof. □

Now we will prove that the proposed difference scheme (3.12)-(3.14) is unconditionally convergent in L_2-norm with the quadratic-order accuracy in all variables.

Suppose that $\{U_{ij}^n \mid (i,j)\in\bar{\omega},\ 0\leq n\leq N\}$ is the exact solution of the system (3.1)-(3.3) and $\{u_{ij}^n \mid (i,j)\in\bar{\omega},\ 0\leq n\leq N\}$ is the numerical solution of the difference scheme (3.12)-(3.14). Let $e_{ij}^n=U_{ij}^n-u_{ij}^n$ $((i,j)\in\bar{\omega},\ 0\leq n\leq N)$.

By subtracting (3.12)-(3.14) from (3.8), (3.10)-(3.11), respectively, we can get the following error equations:

$$\begin{aligned}
\sum_{k=0}^{n-1} \hat{c}_k^{(n)} \left(e_{ij}^{n-k} - e_{ij}^{n-k-1}\right) = & -K_1 h_1^{-\beta} \sum_{k=i-M_1}^{i} g_k^{(\beta)} \left[\sigma e_{i-k,j}^n + (1-\sigma) e_{i-k,j}^{n-1}\right] \\
& - K_2 h_2^{-\gamma} \sum_{k=j-M_2}^{j} g_k^{(\gamma)} \left[\sigma e_{i,j-k}^n + (1-\sigma) e_{i,j-k}^{n-1}\right] \\
& + S_{ij}^n, \quad (i,j) \in \omega,\ 1 \le n \le N,
\end{aligned}$$

$$e_{ij}^n = 0, \quad (i,j) \in \partial\omega,\ 0 \le n \le N,$$

$$e_{ij}^0 = 0, \quad (i,j) \in \omega.$$

Applying the conclusion of Theorem 3.2 and noticing (3.9), we have

$$\begin{aligned}
\|e^n\|^2 \le & \frac{1}{4} \left[\frac{(2L_1)^\beta}{K_1 c_*^{(\beta)}} + \frac{(2L_2)^\gamma}{K_2 c_*^{(\gamma)}}\right] \frac{1}{\sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \max_{1 \le l \le n} \|S^l\|^2 \\
\le & \frac{1}{4} \left[\frac{(2L_1)^\beta}{K_1 c_*^{(\beta)}} + \frac{(2L_2)^\gamma}{K_2 c_*^{(\gamma)}}\right] \frac{1}{\sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}} \left[c_2 \left(h_1^2 + h_2^2 + \tau^2 + \Delta\alpha^2\right)\right]^2 L_1 L_2.
\end{aligned}$$

By extracting the square root on both sides of the above equation, we acquire

$$\|e^n\| \le \frac{c_2}{2} \sqrt{\left[\frac{(2L_1)^\beta}{K_1 c_*^{(\beta)}} + \frac{(2L_2)^\gamma}{K_2 c_*^{(\gamma)}}\right] \frac{L_1 L_2}{\sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}} \left(h_1^2 + h_2^2 + \tau^2 + \Delta\alpha^2\right),$$

where $1 \le n \le N$. Now, the following result can be arrived.

Theorem 3.3. *Suppose that the continuous problem* (3.1)-(3.3) *has a smooth solution* $u(x,y,t) \in C^{(5,5,3)}(\Omega \times [0,T])$, *and let* u_{ij}^n *be the solution of the difference scheme* (3.12)-(3.14). *it holds that*

$$\|e^n\| \le \frac{c_2}{2} \sqrt{\left[\frac{(2L_1)^\beta}{K_1 c_*^{(\beta)}} + \frac{(2L_2)^\gamma}{K_2 c_*^{(\gamma)}}\right] \frac{L_1 L_2}{\sum\limits_{r=0}^{m} \frac{\lambda_r}{T^{\alpha_r}\Gamma(1-\alpha_r)}}} \left(h_1^2 + h_2^2 + \tau^2 + \Delta\alpha^2\right), 1 \le n \le N.$$

3.3. Fast solution techniques with circulant preconditioner

Let

$$\mathbf{u}^n = (u_{1,1}^n, \cdots, u_{M_1-1,1}^n, u_{1,2}^n, \cdots, u_{M_1-1,2}^n, u_{1,M_2-1}^n, \cdots, u_{M_1-1,M_2-1}^n)^T,$$

$$\mathbf{f}^n = (f_{1,1}^n, \cdots, f_{M_1-1,1}^n, f_{1,2}^n, \cdots, f_{M_1-1,2}^n, f_{1,M_2-1}^n, \cdots, f_{M_1-1,M_2-1}^n)^T.$$

Then the implicit difference scheme (3.12) can be rewritten in the matrix form

$$M^n \mathbf{u}^n = p^{n-1}, \quad n = 1, 2, \ldots, N, \tag{3.24}$$

in which

$$M^n = \hat{c}_0^{(n)} I_3 + \sigma K_1 h_1^{-\beta} I_2 \otimes G_\beta + \sigma K_2 h_2^{-\gamma} G_\gamma \otimes I_1, \tag{3.25}$$

and

$$\begin{aligned} p^{n-1} = & -(1-\sigma)[K_1 h_1^{-\beta} I_2 \otimes G_\beta + K_2 h_2^{-\gamma} G_\gamma \otimes I_1]\mathbf{u}^{n-1} \\ & + \sum_{k=1}^{n-1} (\hat{c}_{k-1}^{(n)} - \hat{c}_k^{(n)})\mathbf{u}^{n-k} + \hat{c}_{n-1}^{(n)} u^0 + \mathbf{f}^{n-1+\sigma}, \end{aligned}$$

where $\otimes$ denotes the Kronecker product, I_1, I_2 and I_3 are identity matrices with orders of M_1-1, M_2-1 and $(M_1-1)(M_2-1)$, respectively. $G_\beta \in \mathbb{R}^{(M_1-1)\times(M_1-1)}$ and $G_\gamma \in \mathbb{R}^{(M_2-1)\times(M_2-1)}$ are Toeplitz matrices and have forms as (2.31).

The following lemma guarantees the invertibility of the coefficient matrix M^n.

Lemma 3.1. *The coefficient matrix*

$$M^n = \hat{c}_0^{(n)} I_3 + \sigma K_1 h_1^{-\beta} I_2 \otimes G_\beta + \sigma K_2 h_2^{-\gamma} G_\gamma \otimes I_1$$

of the linear system (3.24) *is symmetric positive definite.*

Proof. According to Lemma 2.5 and the definitions of the matrices G_β and G_γ, one can prove that G_β and G_γ are symmetric positive definite matrices. Therefore, the matrices $I_2 \otimes G_\beta$ and $G_\gamma \otimes I_1$ are also symmetric positive definite. Given that $\hat{c}_0^{(n)} > 0$ and $K > 0$, it is easy to show that the matrix M^n, which is defined by (3.25), is also a symmetric positive definite matrix. □

We also use the CG method for solving the linear system (3.24). In order to improve the performance and reliability of the CG method, the preconditioning techniques are exploited. We refer to the coefficient matrix M^n as a block Toeplitz matrix with Toeplitz blocks (BTTB) [5], Therefore the following level-2 circulant preconditioner which is a block circulant matrix with circulant blocks (BCCB) is considered:

$$C_2^n = \hat{c}_0^{(n)} I_3 + \sigma K_1 h_1^{-\beta} I_2 \otimes c(G_\beta) + \sigma K_2 h_2^{-\gamma} c(G_\gamma) \otimes I_1.$$

Similarly, we discuss the properties of the circulant preconditioner C_2^n in the following.

Lemma 3.2. *The level-2 circulant preconditioner*

$$C_2^n = \hat{c}_0^{(n)} I_3 + \sigma K_1 h_1^{-\beta} I_2 \otimes c(G_\beta) + \sigma K_2 h_2^{-\gamma} c(G_\gamma) \otimes I_1$$

is a symmetric positive definite matrix.

Proof. According to the proof of Lemma 2.10, it is easy to see that $c(G_\beta)$ and $c(G_\gamma)$ are symmetric positive definite matrices. Then, as similar to the Lemma 3.1, we can prove that the level-2 circulant preconditioner C_2^n is a symmetric positive definite matrix. □

According to Lemma 3.2, we can know that the preconditioner C_2^n is nonsingular. Theoretically, for the BCCB matrix C_2^n, the spectrum of $(C_2^n)^{-1}M^n$ is clustered around 1 except for at most $\mathcal{O}(M_1-1)+\mathcal{O}(M_2-1)$ outlying eigenvalues [5]. When the PCG method is used to solve (3.24), the convergence rate will be fast. In Section 4, we will also present numerical examples to demonstrate the usefulness of the proposed circulant preconditioner C_2^n. Thus, the total complexity of the PCG method with preconditioner C_2^n for solving the (3.24) remains $\mathcal{O}((M_1-1)(M_2-1)\log(M_1-1)(M_2-1))$.

4. Numerical example

In this section, we carry out numerical examples to demonstrate the second-order accuracy of the proposed difference schemes and the computational efficiency of the preconditioned Krylov subspace methods. At each time level, we employ the Cholesky method, the CG method and the PCG method for solving the resultant linear systems, respectively. The initial guess for all method is chosen as the zero vector and the stopping criterion is $\|r^{(k)}\|_2/\|r^{(0)}\|_2 < 10^{-12}$, where $r^{(k)}$ is the residual vector after k iterations. Number of iterations required for convergence and CPU time of each method are reported. All numerical experiments are performed in MATLAB (R2016a) on a desktop with 16GB RAM, Inter (R) Core (TM) i7-8700K CPU @3.70GHz.

In Tables 4 and 8, “CPU(s)” denotes the total CPU time in seconds to solve the linear systems, and “Iter” denotes the average number of iterations over 10 runs. For the PCG method, we also report the Strang-based circulant preconditioner [22] S^n and the T. Chan's-based circulant preconditioner [5] T^n. In all tables, “Chol” denotes the Cholesky method, “PCG(S)” is the PCG with the Strang-based preconditioner, “PCG(T)” is the PCG with the T. Chan's-based circulant preconditioner, and “PCG(C)” is the PCG with the proposed circulant preconditioner. Among them, the circulant preconditioner S^n is shown below, and the circulant preconditioner T^n takes the same form except that we replace the S with T.

$$S^n = \hat{c}_0^{(n)} I + \sigma K h^{-\beta} s(G)$$

and

$$S_2^n = \hat{c}_0^{(n)} I_3 + \sigma K_1 h_1^{-\beta} I_2 \otimes s(G_\beta) + \sigma K_2 h_2^{-\gamma} s(G_\gamma) \otimes I_1,$$

where $s(\cdot)$ denotes the Strang circulant preconditioner for the Toeplitz matrix. More precisely, the first column of the criculant matrix $s(G_\beta)$ is given by

$$\begin{pmatrix} g_0^{(\beta)} \\ g_1^{(\beta)} \\ g_2^{(\beta)} \\ \vdots \\ g_{\lfloor \frac{M}{2} \rfloor - 1}^{(\beta)} \\ g_{\lfloor \frac{M}{2} \rfloor + 1 - M}^{(\beta)} \\ \vdots \\ g_{-2}^{(\beta)} \\ \\ g_{-1}^{(\beta)} \\ \\ \end{pmatrix}.$$

Example 4.1. Consider the following 1D time distributed-order and Riesz space fractional diffusion problem:

$$\begin{cases} \int_0^1 \Gamma(5-\alpha){}_0^C D_t^\alpha u(x,t)d\alpha = \dfrac{\partial^\beta u(x,t)}{\partial \mid x \mid^\beta} + f(x,t), \quad 0 < x < 1,\ 0 < t \le T, \\ u(0,t) = 0,\ u(1,t) = 0, \quad 0 \le t \le T, \\ u(x,0) = 0, \quad 0 < x < 1, \end{cases}$$

with $f(x,t) = f_0(x,t) - ct^4\left[f_1(x,t) - 3f_2(x,t) + 3f_3(x,t) - f_4(x,t)\right]$, where $c = -\frac{1}{2\cos(\beta\pi/2)}$, $f_0(x,t) = 24t^3(t-1)x^3(1-x)^3/\ln t$, and

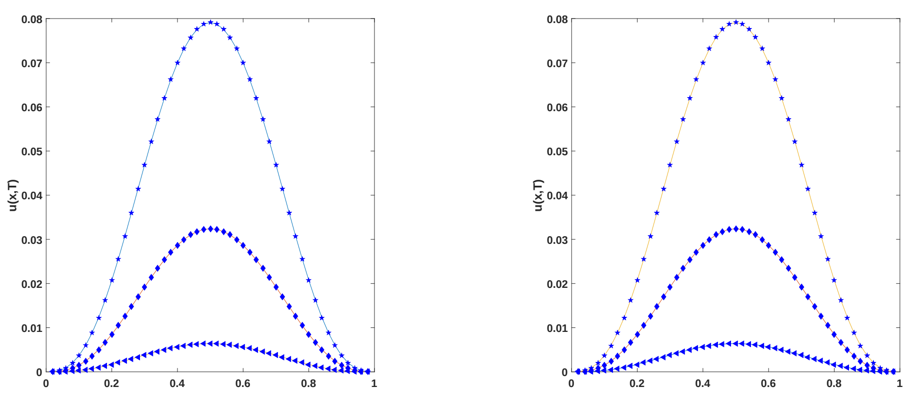
$$\begin{aligned} f_1(x,t) &= \Gamma(4)/\Gamma(4-\beta)[x^{3-\beta} + (1-x)^{3-\beta}], \\ f_2(x,t) &= \Gamma(5)/\Gamma(5-\beta)[x^{4-\beta} + (1-x)^{4-\beta}], \\ f_3(x,t) &= \Gamma(6)/\Gamma(6-\beta)[x^{5-\beta} + (1-x)^{5-\beta}], \\ f_4(x,t) &= \Gamma(7)/\Gamma(7-\beta)[x^{6-\beta} + (1-x)^{6-\beta}]. \end{aligned}$$

The exact solution of this example is given by $u(x,t) = t^4x^3(1-x)^3$.

Let $e(h,\tau,\Delta\alpha) = \max\limits_{\substack{0\le i\le M \\ 0\le n\le N}} |u(x_i,t_n,\Delta\alpha) - u_i^n|$, where $u(x_i,t_n,\Delta\alpha)$ and u_i^n are the exact solution and numerical solution with the step sizes h, τ and $\Delta\alpha$, respectively. We define the convergence orders as

$$rate_h = \log_2 \frac{e(h,\tau,\Delta\alpha)}{e(h/2,\tau,\Delta\alpha)}, rate_\tau = \log_2 \frac{e(h,\tau,\Delta\alpha)}{e(h,\tau/2,\Delta\alpha)}, rate_{\Delta\alpha} = \log_2 \frac{e(h,\tau,\Delta\alpha)}{e(h,\tau,\Delta\alpha/2)}.$$

We take $J = 50$, $M = 50$, $N = 50$. Fig. 1 shows a comparison between the exact solutions and numerical solutions of the difference scheme (2.17)-(2.19) when solving Example 4.1 with different β and T. The good agreement between numerical solutions with the exact solutions can be clearly seen.



(a) $\beta = 1.3$ (b) $\beta = 1.8$

Figure 1. Exact solutions (lines) and numerical solutions (symbols) of Example 4.1: (**a**) $\beta = 1.3$ at $T = 1.5$ (stars), 1.2 (rhombus), 0.8 (triangles); (**b**) $\beta = 1.8$ at $T = 1.5$ (stars), 1.2 (rhombus), 0.8 (triangles).

Some numerical results of the maximum errors as well as the spatial convergence orders for Example 4.1 with $\beta = 1.2,\ 1.5$ and 1.8 when $T = 1.5,\ J = 50,\ N = 1000$ are recorded in Table 1. The second-order convergence of the difference scheme (2.17)-(2.19) in space can be obtained, and the results are in good agreement with what we expect.

Table 1. Maximum errors and spatial convergence orders of difference scheme (2.17)-(2.19) for Example 4.1 with $T = 1.5$; $J = 50$; $N = 1000$.

	$\beta = 1.2$		$\beta = 1.5$		$\beta = 1.8$	
M	$e(h, \tau, \Delta\alpha)$	$rate_h$	$e(h, \tau, \Delta\alpha)$	$rate_h$	$e(h, \tau, \Delta\alpha)$	$rate_h$
32	3.423357e-05	-	6.057253e-05	-	9.145302e-05	-
64	8.665990e-06	1.9820	1.538522e-05	1.9771	2.313195e-05	1.9831
128	2.165532e-06	2.0006	3.847714e-06	1.9995	5.789457e-06	1.9984
256	5.410059e-07	2.0010	9.617429e-07	2.0003	1.449718e-06	1.9977
512	1.353449e-07	1.9990	2.404824e-07	1.9997	3.637954e-07	1.9946

When $T = 1.5,\ J = 50,\ M = 1000$, Table 2 provides numerical results of the maximum errors and the temporal convergence orders for Example 4.1 with different β. Form Table 2, we can see that the temporal convergence order of the difference scheme (2.17)-(2.19) is 2, which is consistent with the theoretical analysis.

Table 2. Maximum errors and temporal convergence orders of difference scheme (2.17)-(2.19) for Example 4.1 with $T = 1.5$; $J = 50$; $M = 1000$.

	$\beta = 1.2$		$\beta = 1.5$		$\beta = 1.8$	
N	$e(h, \tau, \Delta\alpha)$	$rate_\tau$	$e(h, \tau, \Delta\alpha)$	$rate_\tau$	$e(h, \tau, \Delta\alpha)$	$rate_\tau$
8	5.113611e-04	-	6.193316e-04	-	7.401543e-04	-
16	1.294286e-04	1.9822	1.590357e-04	1.9614	1.921523e-04	1.9456
32	3.228485e-05	2.0032	4.014123e-05	1.9862	4.892436e-05	1.9736
64	8.019176e-06	2.0093	1.004941e-05	1.9980	1.231641e-05	1.9900
128	2.005007e-06	1.9998	2.508026e-06	2.0025	3.067275e-06	2.0056

Table 3 gives the maximum errors and distributed-order integral convergence rate for Example 4.1 with $\beta = 1.2,\ 1.5$ and 1.8 respectively at $T = 1.5,\ M = 2000,\ N = 2000$ and various values of J. The desirable second-order convergence of the difference scheme (2.17)-(2.19) is verified. According to the results listed in these three tables, the convergence accuracy of the difference scheme (2.17)-(2.19) of $\mathcal{O}(h^2 + \tau^2 + \Delta\alpha^2)$ can be observed.

Table 3. Maximum errors and distributed-order integral convergence orders of difference scheme (2.17)-(2.19) for Example 4.1 with $T = 1.5$; $M = 2000$; $N = 2000$.

	$\beta = 1.2$		$\beta = 1.5$		$\beta = 1.8$	
J	$e(h, \tau, \Delta\alpha)$	$rate_{\Delta\alpha}$	$e(h, \tau, \Delta\alpha)$	$rate_{\Delta\alpha}$	$e(h, \tau, \Delta\alpha)$	$rate_{\Delta\alpha}$
2	3.774623e-05	-	3.492953e-05	-	3.152437e-05	-
4	9.457965e-06	1.9967	8.747785e-06	1.9975	7.889510e-06	1.9985
8	2.366072e-06	1.9990	2.185762e-06	2.0008	1.967605e-06	2.0035
16	5.917841e-07	1.9994	5.441518e-07	2.0061	4.862444e-07	2.0167
32	1.480919e-07	1.9986	1.336461e-07	2.0256	1.158196e-07	2.0698

From Table 4, one can see that the CPU times of the PCG methods are much less than that of the Cholesky method and the CG method. We also see that the PCG methods exhibit excellent performance in terms of iteration steps, and the number of iteration steps barely increases as M and N increase rapidly. The performance of the R. Chan's-based circulant preconditioner is best among all.

Table 4. Comparisons on Example 4.1 between the Cholesky method, the CG method, and the PCG method with different circulant preconditioners, where $\beta = 1.2$, 1.5 and 1.8, $J = 50$ and $T = 1.5$.

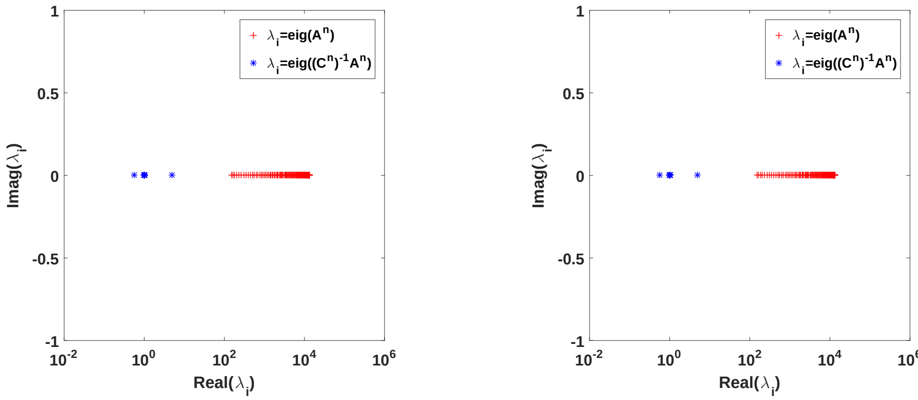
			Chol	CG		PCG(S)		PCG(T)		PCG(C)	
β	M	N	CPU(s)	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter
1.2	2^6	2^5	0.01	0.01	21.0	0.01	7.0	0.01	7.0	**0.01**	**6.0**
	2^7	2^6	0.03	0.04	26.9	0.03	7.0	0.03	7.0	**0.02**	**6.0**
	2^8	2^7	0.10	0.15	31.0	0.08	6.0	0.09	7.0	**0.08**	**6.0**
	2^9	2^8	0.92	0.70	35.9	0.32	6.0	0.34	7.0	**0.31**	**6.0**
	2^{10}	2^9	7.71	1.75	40.0	0.74	6.0	0.80	7.0	**0.73**	**6.0**
	2^{11}	2^{10}	83.32	11.29	44.0	4.12	6.0	4.10	6.0	**4.09**	**6.0**
1.5	2^6	2^5	0.01	0.02	31.0	0.01	8.0	0.01	9.0	**0.01**	**6.0**
	2^7	2^6	0.03	0.06	49.0	0.03	7.0	0.03	9.0	**0.02**	**6.0**
	2^8	2^7	0.10	0.27	73.0	0.09	7.0	0.10	9.0	**0.08**	**6.0**
	2^9	2^8	0.92	1.50	99.0	0.32	6.0	0.36	8.0	**0.34**	**7.0**
	2^{10}	2^9	7.69	4.67	126.0	0.79	7.0	0.85	8.0	**0.79**	**7.0**
	2^{11}	2^{10}	83.13	35.23	158.0	4.46	7.0	4.80	8.0	**4.45**	**7.0**
1.8	2^6	2^5	0.01	0.02	32.0	0.01	8.0	0.01	10.0	**0.01**	**6.0**
	2^7	2^6	0.03	0.08	64.0	0.03	8.0	0.03	11.9	**0.02**	**6.0**
	2^8	2^7	0.10	0.40	120.0	0.09	7.5	0.11	13.0	**0.08**	**6.0**
	2^9	2^8	0.91	3.22	215.0	0.35	7.5	0.49	14.0	**0.32**	**6.0**
	2^{10}	2^9	7.66	12.63	364.0	0.74	6.0	1.16	14.0	**0.84**	**7.8**
	2^{11}	2^{10}	83.05	117.39	552.0	4.79	8.0	6.82	14.0	**4.77**	**8.0**

Table 5 reports the memory usage of the above methods for Example 4.1. As seen from Table 5, the PCG methods and the CG method have similar performances in terms of the memory requirement, and they are considerably better than the Cholesky method. Because the direct method (the Cholesky method) needs to store dense coefficient matrices, while the iterative methods (the PCG methods and the CG method) do not need to store any dense matrices.

Table 5. Memory comparisons on Example 4.1 between the Cholesky method, the CG method, and the PCG methods, where $\beta = 1.2$, $J = 50$ and $T = 1.5$.

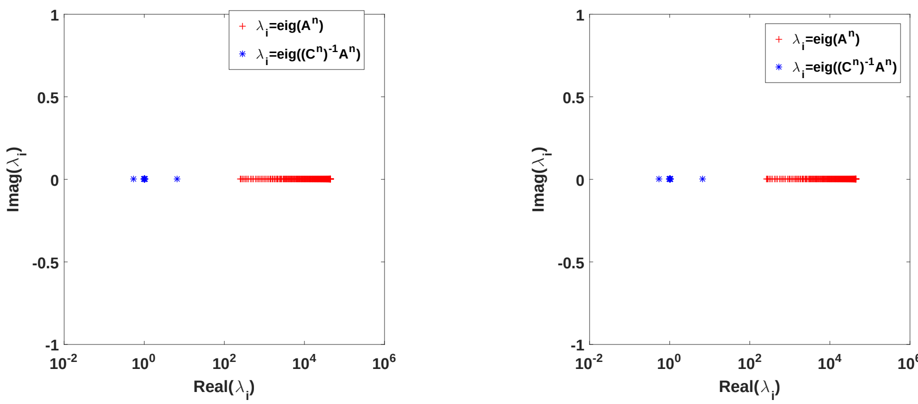
M	N	Chol	CG	PCG(S)	PCG(T)	PCG(C)
2^9	2^8	4.65 MB	2.15 MB	2.15 MB	2.16 MB	2.16 MB
2^{10}	2^9	25.46 MB	9.16 MB	9.17 MB	9.17 MB	9.17 MB
2^{11}	2^{10}	114.30 MB	38.73 MB	38.75 MB	38.75 MB	38.76 MB

The spectrum of the original matrix A^n and the preconditioned matrix $(C^n)^{-1}A^n$ are plotted in Figs. 2-3. We can see that the eigenvalues of the preconditioned matrix $(C^n)^{-1}A^n$ lie within a small interval around 1, expect for few outliers, yet all the eigenvalues are well separated away from 0. This confirms that the circulant preconditioner have nice clustering properties.



(a) n=0 (b) n = 1

Figure 2. Spectrum of original matrice (red) and R. Chan's-based preconditioned matrice (blue) for Example 4.1 at time level (**a**) $n = 0$ and (**b**) $n = 1$, respectively, when $M = N = 128$, $J = 50$, $\beta = 1.8$, and $T = 1.5$.



(a) n=0 (b) n = 1

Figure 3. Spectrum of original matrice (red) and R. Chan's-based preconditioned matrice (blue) for Example 4.1 at time level (**a**) $n = 0$ and (**b**) $n = 1$, respectively, when $M = N = 256$, $J = 50$, $\beta = 1.8$, and $T = 1.5$.

Example 4.2. Consider the following 1D time distributed-order and Riesz space fractional diffusion problem:

$$\begin{cases} \int_0^1 \Gamma(5-\alpha)_0^C D_t^{\alpha} u(x,t)d\alpha = \frac{\partial^{\beta} u(x,t)}{\partial |x|^{\beta}} + f(x,t), \quad 0 < x < 1,\ 0 < t \leq T, \\ u(0,t) = 0,\ u(1,t) = 0, \quad 0 \leq t \leq T, \\ u(x,0) = 0, \quad 0 < x < 1, \end{cases}$$

where $f(x,t) = x^2 t^2$.

Since we do not know the exact solution, we treat the calculated solution for a very fine spatial mesh as the exact solution. The fine mesh is $2^{12} \times 2^{12}, (M \times N)$.

Table 6. Comparisons on Example 4.2 between the Cholesky method, the CG method, and the PCG method with different circulant preconditioners, where $\beta = 1.2$, 1.5 and 1.8, $J = 50$ and $T = 1.5$.

			Chol	CG		PCG(S)		PCG(T)		PCG(C)	
β	M	N	CPU(s)	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter
1.2	2^6	2^4	0.00	0.01	31.2	0.00	9.0	0.00	9.0	**0.00**	**8.0**
	2^7	2^5	0.01	0.02	39.1	0.01	9.0	0.01	9.0	**0.01**	**8.0**
	2^8	2^6	0.02	0.08	46.4	0.02	9.0	0.02	9.0	**0.02**	**8.0**
	2^9	2^7	0.41	0.40	53.6	0.13	10.0	0.13	9.0	**0.13**	**9.0**
	2^{10}	2^8	3.75	1.09	60.8	0.33	9.6	0.31	9.0	**0.31**	**9.0**
	2^{11}	2^9	40.75	7.75	68.0	1.99	9.0	2.00	9.0	**1.98**	**9.0**
1.5	2^6	2^4	0.00	0.01	51.1	0.00	10.9	0.00	11.4	**0.00**	**8.0**
	2^7	2^5	0.01	0.04	78.9	0.01	9.7	0.01	12.4	**0.01**	**8.4**
	2^8	2^6	0.02	0.18	113.0	0.03	10.0	0.03	13.0	**0.02**	**9.5**
	2^9	2^7	0.44	1.10	151.7	0.14	10.0	0.17	13.0	**0.14**	**9.9**
	2^{10}	2^8	3.82	3.32	195.0	0.34	10.0	0.41	12.7	**0.33**	**10.0**
	2^{11}	2^9	40.58	26.25	244.7	2.34	11.0	2.50	12.0	**2.32**	**11.0**
1.8	2^6	2^4	0.00	0.01	63.0	0.00	10.0	0.00	15.0	**0.00**	**9.0**
	2^7	2^5	0.01	0.06	120.0	0.01	10.9	0.02	17.0	**0.01**	**9.0**
	2^8	2^6	0.02	0.34	217.3	0.03	11.0	0.04	18.0	**0.03**	**10.0**
	2^9	2^7	0.43	2.68	378.1	0.14	10.2	0.24	20.0	**0.13**	**10.0**
	2^{10}	2^8	3.81	10.09	603.4	0.35	10.0	0.61	20.2	**0.33**	**10.0**
	2^{11}	2^9	40.90	92.28	875.0	2.17	10.0	4.00	20.7	**2.15**	**10.0**

Table 7. Comparisons on Example 4.2 between the Cholesky method, the CG method, and the PCG method with different circulant preconditioners, where $\beta = 1.2$, 1.5 and 1.8, $J = 50$ and $T = 6$.

			Chol	CG		PCG(S)		PCG(T)		PCG(C)	
β	M	N	CPU(s)	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter
1.2	2^6	2^4	0.00	0.01	39.1	0.00	10.0	0.00	10.0	**0.00**	**8.0**
	2^7	2^5	0.01	0.03	54.1	0.01	9.7	0.01	11.0	**0.01**	**9.0**
	2^8	2^6	0.02	0.11	70.5	0.02	10.0	0.02	11.0	**0.02**	**9.0**
	2^9	2^7	0.38	0.64	87.7	0.12	10.0	0.13	11.0	**0.12**	**9.0**
	2^{10}	2^8	3.62	1.81	104.7	0.29	11.0	0.30	11.0	**0.28**	**10.0**
	2^{11}	2^9	39.97	12.95	121.4	2.12	11.0	2.01	10.3	**1.99**	**10.1**
1.5	2^6	2^4	0.00	0.01	55.8	0.00	11.0	0.00	13.2	**0.00**	**8.0**
	2^7	2^5	0.01	0.05	92.1	0.01	11.9	0.01	14.1	**0.01**	**10.0**
	2^8	2^6	0.02	0.22	146.4	0.02	10.6	0.03	15.0	**0.02**	**10.0**
	2^9	2^7	0.38	1.60	223.1	0.12	10.0	0.18	15.9	**0.13**	**11.0**
	2^{10}	2^8	3.59	5.33	319.8	0.29	11.0	0.38	15.1	**0.29**	**11.0**
	2^{11}	2^9	39.80	44.90	429.5	2.11	11.0	2.73	15.0	**2.10**	**11.0**
1.8	2^6	2^4	0.00	0.01	63.0	0.00	11.1	0.00	16.0	**0.00**	**9.0**
	2^7	2^5	0.01	0.06	120.0	0.01	11.0	0.01	18.0	**0.01**	**9.8**
	2^8	2^6	0.02	0.33	229.1	0.02	12.0	0.04	20.6	**0.02**	**10.0**
	2^9	2^7	0.37	2.98	420.5	0.13	11.0	0.25	24.0	**0.12**	**10.0**
	2^{10}	2^8	3.59	12.50	758.6	0.29	11.0	0.64	27.3	**0.29**	**11.0**
	2^{11}	2^9	39.55	136.81	1321.1	2.26	12.0	5.07	30.0	**2.11**	**11.0**

From Tables 6-7, we see that the PCG methods exhibit excellent performance both in terms of CPU time and iteration steps, and the proposed PCG(C) method is the best one.

Table 8 reports the memory usage of the above methods for Example 4.2. As seen from Table 8, the PCG methods and the CG method have similar performances in terms of the memory requirement, and they are considerably better than the Cholesky method.

Table 8. Memory comparisons on Example 4.2 between the Cholesky method, the CG method, and the PCG methods, where $\beta = 1.5$, $J = 50$ and $T = 6$.

M	N	Chol	CG	PCG(S)	PCG(T)	PCG(C)
2^9	2^7	3.06 MB	0.51 MB	0.51 MB	0.51 MB	0.51 MB
2^{10}	2^8	18.36 MB	1.95 MB	1.96 MB	1.96 MB	1.96 MB
2^{11}	2^9	83.43 MB	7.63 MB	7.66 MB	7.66 MB	7.66 MB

Example 4.3. Consider the following 2D time distributed-order and Riesz space fractional diffusion problem:

$$\begin{cases} \int_0^1 \Gamma(5-\alpha)\,{}_0^C D_t^\alpha u(x,y,t)d\alpha = \frac{\partial^\beta u(x,y,t)}{\partial|x|^\beta} + \frac{\partial^\gamma u(x,y,t)}{\partial|y|^\gamma} + f(x,y,t), \\ \qquad\qquad\qquad\qquad (x,y)\in\Omega,\ 0<t\le T, \\ u(x,y,t)=0, \qquad (x,y)\in\partial\Omega,\ 0\le t\le T, \\ u(x,y,0)=0, \quad (x,y)\in\Omega, \end{cases}$$

with $\Omega = (0,1)\times(0,1)$, and

$$\begin{aligned} &f(x,y,t)\\ =&f_0(x,y,t) - c_1 t^4 y^3(1-y)^3[f_1(x,y,t) - 3f_2(x,y,t) + 3f_3(x,y,t) - f_4(x,y,t)]\\ &- c_2 t^4 x^3(1-x)^3[g_1(x,y,t) - 3g_2(x,y,t) + 3g_3(x,y,t) - g_4(x,y,t)], \end{aligned}$$

where $c_1 = -\frac{1}{2\cos(\beta\pi/2)}$, $c_2 = -\frac{1}{2\cos(\gamma\pi/2)}$, and

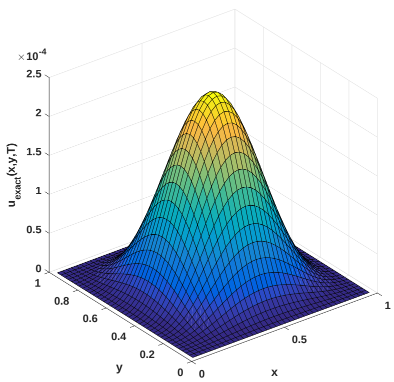
$$\begin{aligned} f_0(x,y,t) &= 24t^3(t-1)x^3(1-x)^3y^3(1-y)^3/\ln t,\\ f_1(x,y,t) &= \Gamma(4)/\Gamma(4-\beta)[x^{3-\beta} + (1-x)^{3-\beta}],\\ f_2(x,y,t) &= \Gamma(5)/\Gamma(5-\beta)[x^{4-\beta} + (1-x)^{4-\beta}],\\ f_3(x,y,t) &= \Gamma(6)/\Gamma(6-\beta)[x^{5-\beta} + (1-x)^{5-\beta}],\\ f_4(x,y,t) &= \Gamma(7)/\Gamma(7-\beta)[x^{6-\beta} + (1-x)^{6-\beta}],\\ g_1(x,y,t) &= \Gamma(4)/\Gamma(4-\gamma)[y^{3-\gamma} + (1-y)^{3-\gamma}],\\ g_2(x,y,t) &= \Gamma(5)/\Gamma(5-\gamma)[y^{4-\gamma} + (1-y)^{4-\gamma}],\\ g_3(x,y,t) &= \Gamma(6)/\Gamma(6-\gamma)[y^{5-\gamma} + (1-y)^{5-\gamma}],\\ g_4(x,y,t) &= \Gamma(7)/\Gamma(7-\gamma)[y^{6-\gamma} + (1-y)^{6-\gamma}]. \end{aligned}$$

The exact solution of the example is $u(x,t) = t^4x^3(1-x)^3y^3(1-y)^3$.

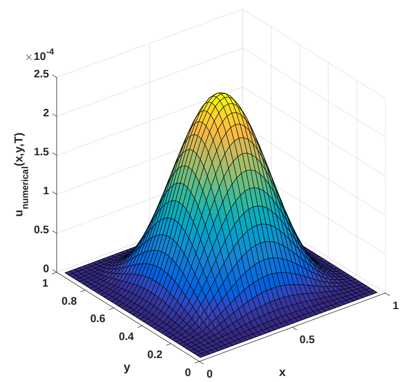
For simplicity, take $h_1 = h_2 = \widetilde{h}$, and $M_1 = M_2 = \widetilde{M}$. Let $e(\widetilde{h},\tau,\Delta\alpha) = \max\limits_{\substack{0\le i\le M_1,\ 0\le j\le M_2\\ 0\le n\le N}} |u(x_i,y_j,t_n,\Delta\alpha) - u_{ij}^n|$, where $u(x_i,y_j,t_n,\Delta\alpha)$ and u_{ij}^n represent the exact solution and numerical solution with the step sizes $\widetilde{h}$, τ and $\Delta\alpha$, respectively. The convergence orders are defined as

$$\widetilde{rate}_h = \log_2 \frac{e(\widetilde{h}, \tau, \Delta\alpha)}{e(\widetilde{h}/2, \tau, \Delta\alpha)}, \widetilde{rate}_\tau = \log_2 \frac{e(\widetilde{h}, \tau, \Delta\alpha)}{e(\widetilde{h}, \tau/2, \Delta\alpha)}, \widetilde{rate}_{\Delta\alpha} = \log_2 \frac{e(\widetilde{h}, \tau, \Delta\alpha)}{e(\widetilde{h}, \tau, \Delta\alpha/2)}.$$

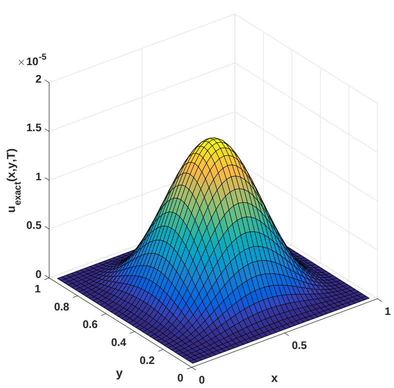
Fig. 4 exhibits the solution surface of Example 4.3 with $J = 50$, $\widetilde{M}$ =40, $N = 10$ at $T = 1$, $\beta = \gamma = 1.8$ and $T = 0.5$, $\beta = \gamma = 1.3$, respectively. It can be seen that the numerical solutions are in good conformity with the exact solutions.



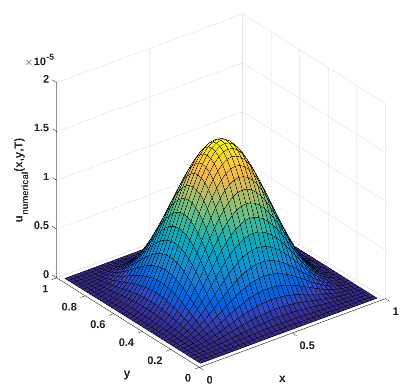
(a) $T = 1$, $\beta = \gamma = 1.8$



(b) $T = 1$, $\beta = \gamma = 1.8$



(c) $T = 0.5$, $\beta = \gamma = 1.3$



(d) $T = 0.5$, $\beta = \gamma = 1.3$

Figure 4. The solution surfaces obtained from Example 4.3 at $J = 50$, $\widetilde{M}$ =40 and $N = 10$: (**a**) the exact solution with $T = 1$, $\beta = \gamma = 1.8$; (**b**) the numerical solution with $T = 1$, $\beta = \gamma = 1.8$ by the scheme (3.12)-(3.14); (**c**) the exact solution with $T = 0.5$, $\beta = \gamma = 1.3$; (**d**) the numerical solution with $T = 0.5$, $\beta = \gamma = 1.3$ by the scheme (3.12)-(3.14);

When $T = 1.5$, $J = 50$ and $N = 2000$, Table 9 lists the maximum errors and convergence orders in spatial of the difference scheme with $\beta = \gamma = 1.2$, 1.5 and 1.8, respectively. From the numerical results we can conclude that the difference scheme (3.12)-(3.14) has the second-order convergence in spatial directions.

Table 9. Maximum errors and spatial convergence orders of difference scheme (3.12)-(3.14) for Example 4.3 with $T = 1.5$; $J = 50$; $N = 2000$.

	$\beta=\gamma=1.2$		$\beta=\gamma=1.5$		$\beta=\gamma=1.8$	
$\widetilde{M}$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_h$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_h$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_h$
8	1.287397e-05	-	2.196133e-05	-	3.333244e-05	-
16	3.193383e-06	2.0113	5.413665e-06	2.0203	8.121444e-06	2.0371
32	7.961929e-07	2.0039	1.348003e-06	2.0058	2.016181e-06	2.0101
64	1.982329e-07	2.0059	3.360192e-07	2.0042	5.025245e-07	2.0044
128	4.880024e-08	2.0222	8.327729e-08	2.0125	1.249136e-07	2.0083

When taking the fixed $T = 1.5$, $J = 50$, $\widetilde{M} = 100$, the maximum errors and convergence orders in temporal of the difference scheme (3.12)-(3.14) with $\beta = \gamma = 1.2$, 1.5 and 1.8 are listed in Table 10, respectively. From the numerical results in Table 10 we can clearly see that the convergence order in temporal of the difference scheme (3.12)-(3.14) is also nearly 2, which is in accord with the theoretical analysis.

Table 10. Maximum errors and temporal convergence orders of difference scheme (3.12)-(3.14) for Example 4.3 with $T = 1.5$; $J = 50$; $\widetilde{M} = 300$.

	$\beta=\gamma=1.2$		$\beta=\gamma=1.5$		$\beta=\gamma=1.8$	
N	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_\tau$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_\tau$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_\tau$
4	4.143174e-05	-	4.904054e-05	-	5.654620e-05	-
8	1.114619e-05	1.8942	1.334121e-05	1.8781	1.550841e-05	1.8664
16	2.888455e-06	1.9482	3.479612e-06	1.9389	4.067734e-06	1.9308
32	7.263603e-07	1.9915	8.810213e-07	1.9817	1.031666e-06	1.9792
64	1.767839e-07	2.0387	2.124202e-07	2.0523	2.460280e-07	2.0681

The numerical accuracy of scheme (3.12)-(3.14) for Example 4.3 in distributed-order integral variable is investigated. When $T = 1.5$, $J = 50$, $\widetilde{M} = 100$, Table 11 displays the computational results using the difference scheme (3.12)-(3.14) with $\beta = \gamma = 1.2$, 1.5 and 1.8, respectively. One can draw the conclusion that the convergence accuracy in distributed-order integral variable is $\mathcal{O}(\Delta\alpha^2)$. Namely, the numerical convergence order of the difference scheme (3.12)-(3.14) is $\mathcal{O}(h_1^2 + h_2^2 + \tau^2 + \Delta\alpha^2)$.

Table 11. Maximum errors and distributed-order integral convergence orders of difference scheme (3.12)-(3.14) for Example 4.3 with $T = 1.5$; $\widetilde{M} = 800$; $N = 2000$.

	$\beta=\gamma=1.2$		$\beta=\gamma=1.5$		$\beta=\gamma=1.8$	
J	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_{\Delta\alpha}$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_{\Delta\alpha}$	$e(\widetilde{h},\tau,\Delta\alpha)$	$\widetilde{rate}_{\Delta\alpha}$
1	2.037150e-06	-	1.777609e-06	-	1.489954e-06	-
2	5.120092e-07	1.9923	4.455139e-07	1.9964	3.719972e-07	2.0019
4	1.274115e-07	2.0067	1.101005e-07	2.0166	9.090557e-08	2.0329
8	3.105365e-08	2.0367	2.609633e-08	2.0769	2.053103e-08	2.1466

From Table 12, we can observe that the CPU time of the PCG method with circulant preconditioners is much less than that of the Cholesky method and the CG method. We also see that the number of iteration steps of the PCG method barely increases as the number of the spatial grid points increases. The performance of the R. Chan's-based circulant preconditioner is best amongst all.

Table 12. Comparisons on Example 4.3 between the Cholesky method, the CG method, and the PCG method with different circulant preconditioners, where $\beta = \gamma = 1.2,\ 1.5,\ 1.8$, $J = 50$ and $T = 1.5$.

			Chol	CG		PCG(S)		PCG(T)		PCG(C)	
$\beta{=}\gamma$	$\widetilde{M}$	N	CPU(s)	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter	CPU(s)	Iter
1.2	2^3	2^3	0.01	0.01	10.0	0.01	10.0	0.01	8.0	**0.01**	**7.0**
	2^4	2^4	0.02	0.03	15.0	0.03	9.0	0.03	8.1	**0.03**	**9.0**
	2^5	2^5	0.29	0.29	19.0	0.22	10.0	0.21	9.0	**0.21**	**9.0**
	2^6	2^6	16.55	1.42	23.0	1.06	10.0	1.01	9.0	**1.06**	**10.0**
	2^7	2^7	2029.77	16.32	27.0	10.06	10.0	9.52	9.0	**10.04**	**10.0**
1.5	2^3	2^3	0.01	0.01	10.0	0.01	12.0	0.01	8.0	**0.01**	**8.0**
	2^4	2^4	0.03	0.04	20.0	0.03	12.1	0.03	11.0	**0.03**	**10.0**
	2^5	2^5	0.30	0.42	31.9	0.27	13.0	0.25	12.0	**0.24**	**12.0**
	2^6	2^6	16.52	2.25	46.0	1.21	13.0	1.20	13.0	**1.20**	**13.0**
	2^7	2^7	2030.88	31.66	63.0	12.17	14.0	11.84	13.4	**12.16**	**14.0**
1.8	2^3	2^3	0.01	0.01	10.0	0.01	14.0	0.01	9.0	**0.01**	**8.0**
	2^4	2^4	0.03	0.05	25.9	0.04	15.0	0.04	13.0	**0.03**	**11.0**
	2^5	2^5	0.30	0.56	45.0	0.30	16.1	0.30	16.0	**0.26**	**13.0**
	2^6	2^6	16.33	3.45	79.0	1.43	17.9	1.49	19.0	**1.29**	**15.0**
	2^7	2^7	1999.64	59.59	128.0	14.78	18.9	16.94	23.0	**13.74**	**16.9**

Table 13 reports the memory usage of the above methods for Example 4.3. From Table 13, we can see that the PCG methods and the CG method have similar performances in terms of the memory requirement, and they are considerably better than the Cholesky method. Because the direct method needs to store dense coefficient matrices, while the iterative methods do not need to store any dense matrices.

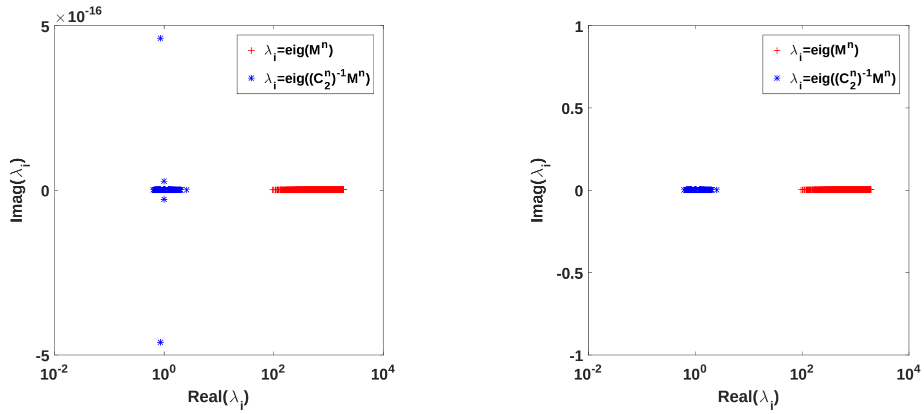
Table 13. Memory comparisons on Example 4.3 between the Cholesky method, the CG method, and the PCG methods, where $\beta = 1.8$, $J = 50$ and $T = 1.5$.

M	N	Chol	CG	PCG(S)	PCG(T)	PCG(C)
2^5	2^5	7.91 MB	0.14 MB	0.16 MB	0.15 MB	0.15 MB
2^6	2^6	129.35 MB	1.41 MB	1.51 MB	1.43 MB	1.43 MB
2^7	2^7	2045.17 MB	12.29 MB	12.52 MB	12.36 MB	12.36 MB

The spectrum of the matrix M^n and the preconditioned matrix $(C_2^n)^{-1}M^n$ are plotted in Figs. 5-6. These two figures also confirm that the circulant preconditioner have nice clustering properties. It shows that the eigenvalues of the preconditioned matrix are well grouped around 1 expect for few outliers. The vast majority of the eigenvalues are well separated away from 0.

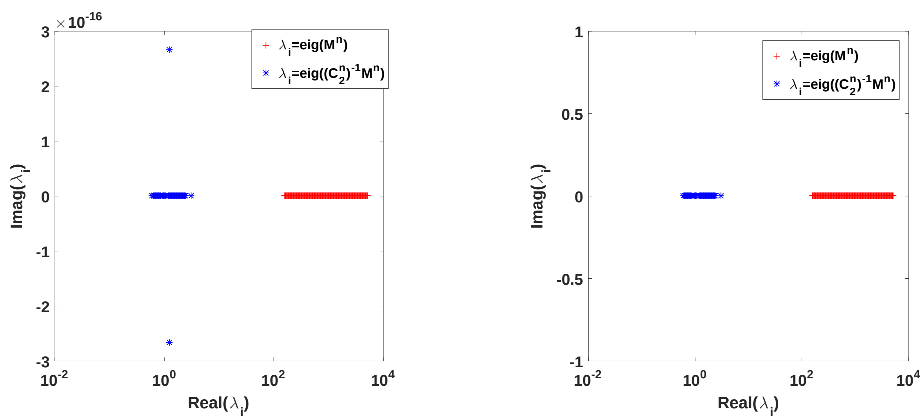
5. Conclusion

In this paper, efficient second-order difference schemes are proposed for one- and two-dimensional TDRFDEs. We first discretize the time distributed-order integral



(a) $n = 0$ (b) $n = 1$

Figure 5. Spectrum of original matrice (red) and R. Chan's-based preconditioned matrice (blue) for Example 4.3 at time level (**a**) $n = 0$ and (**b**) $n = 1$, respectively, when $\widetilde{M} = N = 64$, $J = 50$, $\beta = \gamma = 1.5$, and $T = 1.5$.



(a) $n = 0$ (b) $n = 1$

Figure 6. Spectrum of original matrice (red) and R. Chan's-based preconditioned matrice (blue) for Example 4.3 at time level (**a**) $n = 0$ and (**b**) $n = 1$, respectively, when $\widetilde{M} = N = 128$, $J = 50$, $\beta = \gamma = 1.5$, and $T = 1.5$.

term by using composite trapezoid formula and transform the TDRFDEs into the multi-term time-space FDEs. Then we solve the multi-term time-space FDEs with the second-order accurate interpolation approximation on a special point. We prove that the proposed difference schemes are uniquely solvable, unconditionally stable and convergent in the mesh L_2-norm with second-order accuracy in time, space and distributed-order integral variables. Moreover, we have proposed an efficient implementation of the proposed scheme based on the PCG method with R. Chan's-based circulant preconditioner, which only requires $\mathcal{O}((M-1)\log(M-1))$ computational complexity and $\mathcal{O}((M-1)$ storage cost. Numerical experiments confirm the theoretical results and show the effectiveness of the proposed preconditioned method. In future work, we will focus on the development of the effective numerical methods for solving high-dimensional time distributed-order fractional diffusion-wave equations.

Acknowledgements. We are grateful to the referees and editors for their insightful suggestions and comments that improved the presentation of this paper.

References

[1] A. A. Alikhanov, *A new difference scheme for the time fractional diffusion equation*, J. Comput. Phys., 2015, 280, 424–438.

[2] A. H. Bhrawy and M. A. Zaky, *A method based on the Jacobi tau approximation for solving multi-term time-space fractional partial differential equations*, J. Comput. Phys., 2015, 281, 876–895.

[3] A. H. Bhrawy and M. A. Zaky, *Numerical simulation of multi-dimensional distributed-order generalized Schrödinger equations*, Nonlinear Dyn., 2017, 89, 1415–1432.

[4] W. Bu, A. Xiao and W. Zeng, *Finite difference/finite element methods for distributed-order time fractional diffusion equations*, J. Sci. Comput., 2017, 72(3), 422–441.

[5] R. Chan and X. Q. Jin, *An Introduction to Iterative Toeplitz Solvers*, SIAM, PA, 2007.

[6] R. Chan and G. Strang, *Toeplitz equations by conjugate gradients with circulant preconditioner*, SIAM J. Sci. Stat. Comput., 1989, 10(1), 104–119.

[7] A. Chechkin, R. Gorenflo and I. Sokolov, *Retarding subdiffusion and accelerating superdiffusion governed by distributed-order fractional diffusion equations*, Phys. Rev. E, 2002, 66(4), 046129.

[8] L. K. Chou and S. L. Lei, *Fast ADI method for high dimensional fractional diffusion equations in conservative form with preconditioned strategy*, Comput. Math. Appl., 2017, 73(3), 385–403.

[9] G. H. Gao, A. A. Alikhanov and Z. Z. Sun, *The temporal second order difference schemes based on the interpolation approximation for solving the time multi-term and distributed-order fractional sub-diffusion equations*, J. Sci. Comput., 2017, 73, 93–121.

[10] G. H. Gao, H. W. Sun and Z. Z. Sun, *Some high-order difference schemes for the distributed-order differential equations*, J. Comput. Phys., 2015, 298, 337–359.

[11] G. H. Gao and Z. Z. Sun, *Two alternating direction implicit difference schemes with the extrapolation method for the two-dimensional distributed-order differential equations*, Comput. Math. Appl., 2015, 69(9), 926–948.

[12] X. M. Gu, T. Z. Huang, B. Carpentieri et al., *A hybridized iterative algorithm of the BiCORSTAB and GPBiCOR methods for solving non-Hermitian linear systems*, Comput. Math. Appl., 2015, 70(12), 3019–3031.

[13] X. M. Gu, T. Z. Huang, C. C. Ji et al., *Fast iterative method with a second order implicit difference scheme for time-space fractional convection-diffusion equations*, J. Sci. Comput., 2017, 72, 957–985.

[14] X. M. Gu, T. Z. Huang, H. B. Li et al., *On k-step CSCS-based polynomial preconditioners for Toeplitz linear systems with application to fractional diffusion equations*, Appl. Math. Lett., 2015, 42, 53–58.

[15] X. M. Gu, T. Z. Huang, X. L. Zhao et al., *Strang-type preconditioners for solving fractional diffusion equations by boundary value methods*, J. Comput. Appl. Math., 2015, 277, 73–86.

[16] B. Henry, T. Langlands and S. Wearne, *Anomalous diffusion with linear reaction dynamics: from continuous time random walks to fractional reaction-diffusion equations*, Phys. Rev. E, 2006, 74(3), 031116.

[17] X. Hu, F. Liu, I. Turner and V. Anh, *An implicit numerical method of a new time distributed-order and two-sided space-fractional advection-dispersion equation*, Numer. Algor., 2016, 72(2), 393–407.

[18] H. Jiang, F. Liu, I. Turner and K. Burrage, *Analytical solutions for the multi-term time-space Caputo-Riesz fractional advection-diffusion equations on a finite domain*, J. Math. Anal. Appl., 2012, 389(2), 1117–1127.

[19] J. T. Katsikadelis, *Numerical solution of distributed order fractional differential equations*, J. Comput. Phys., 2014, 259, 11–22.

[20] A. N. Kochubei, *Distributed order calculus and equations of ultraslow diffusion*, J. Math. Anal. Appl., 2008, 340(1), 252–281.

[21] S. L. Lei, X. Chen and X. Zhang, *Multilevel circulant preconditioner for high-dimensional fractional diffusion equations*, East Asian J. Appl. Math., 2016, 6(2), 109–130.

[22] S. L. Lei and H. W. Sun, *A circulant preconditioner for fractional diffusion equations*, J. Comput. Phys., 2013, 242, 715–725.

[23] C. Li and H. Ding, *Higher order finite difference method for the reaction and anomalous-diffusion equation*, Appl. Math. Model., 2014, 38(15), 3802–3821.

[24] H. Li, X. Wu, J. Zhang et al., *Numerical solution of the time-fractional sub-diffusion equation on an unbounded domain in two-dimensional space*, East. Asia. J. Appl. Math., 2017, 7(3), 439–454.

[25] Z. Li, Y. Luchko and M. Yamamoto, *Analyticity of solutions to a distributed order time-fractional diffusion equation and its application to an inverse problem*, Comput. Math. Appl., 2017, 73(6), 1041–1052.

[26] F. Liu, M. M. Meerschaert, R. J. McGough et al., *Numerical methods for solving the multi-term time-fractional wave-diffusion equation*, Fract. Calc. Appl. Anal., 2013, 16(1), 9–25.

[27] F. Liu, C. Yang and K. Burrage, *Numerical method and analytical technique of the modified anomalous subdiffusion equation with a nonlinear source term*, J. Comput. Appl. Math., 2009, 231(1), 160–176.

[28] Q. Liu, F. Liu, I. Turner and V. Anh, *Finite element approximation for a modified anomalous subdiffusion equation*, Appl. Math. Model., 2011, 35(8), 4103–4116.

[29] R. L. Magin, O. Abdullah, D. Baleanu and X. J. Zhou, *Anomalous diffusion expressed through fractional order differential operators in the Bloch-Torrey equation*, J. Magn. Reson., 2008, 190(2), 255–270.

[30] S. Mashayekhi and M. Razzaghi, *Numerical solution of distributed order fractional differential equations by hybrid functions*, J. Comput. Phys., 2016, 315, 169–181.

[31] M. M. Meerschaert, E. Nane and P. Vellaisamy, *Distributed-order fractional diffusions on bounded domains*, J. Math. Anal. Appl., 2011, 379(1), 216–228.

[32] D. Rui, Z. P. Hao and Z. Z. Sun, *Lubich second-order methods for distributed-order time-fractional differential equations with smooth solutions*, East. Asia. J. Appl. Math., 2016, 6(2), 131–151.

[33] H. Scher and E. W. Montroll, *Anomalous transit-time dispersion in amorphous solids*, Phys. Rev. B, 1975, 12(6), 2455.

[34] W. Schneider and W. Wyss, *Fractional diffusion and wave equations*, J. Math. Phys., 1989, 30(1), 134–144.

[35] Z.-Z. Sun and G.-H. Gao, *Finite Difference Methods for the Fractional Differential Equations*, Science Press, Beijing, 2015. (in Chinese).

[36] K. Wang and H. Wang, *A fast characteristic finite difference method for fractional advection-diffusion equations*, Adv. Water Resour., 2011, 34(7), 810–816.

[37] H. Ye, F. Liu, V. Anh and I. Turner, *Numerical analysis for the time distributed-order and Riesz space fractional diffusions on bounded domains*, IMA J. Appl. Math., 2015, 80(3), 825–838.

[38] M. A. Zaky, *A Legendre collocation method for distributed-order fractional optimal control problems*, Nonlinear Dyn., 2018, 91, 2667–2681.

[39] M. A. Zaky, *A Legendre spectral quadrature tau method for the multi-term time-fractional diffusion equations*, Comp. Appl. Math., 2018, 37, 3525–3538.

[40] M. A. Zaky, E. H. Doha and J. A. T. Machado, *A spectral numerical method for solving distributed-order fractional initial value problems*, J. Comput. Nonlinear Dynam., 2018, 13(10), 101007.

[41] M. A. Zaky and J. A. T. Machado, *On the formulation and numerical simulation of distributed-order fractional optimal control problems*, Commun. Nonlinear Sci. Numer. Simulat., 2017, 52, 177–189.

[42] X. L. Zhao, T. Z. Huang, S. L. Wu and Y. F. Jing, *DCT- and DST-based splitting methods for Toeplitz systems*, Int. J. Comput. Math., 2012, 89(5), 691–700.