Citation: Shuyu Yue, Tao Wang. Fast higher-order numerical schemes for two-dimensional nonlinear Riesz space fractional diffusion equations. *Journal of Harbin Institute of Technology (New Series)*. DOI: 10.11916/j.issn.1005-9113.25041

Fast Higher-Order Numerical Schemes for Two-Dimensional Nonlinear Riesz Space Fractional Diffusion Equations

Shuyu Yue and Tao Wang *

(*School of Mathematical Sciences, University of Jinan, Jinan 250022, China*)

Abstract: Fast fourth-order linearized Alternating Direction Implicit (ADI) and Locally One-Dimensional (LOD) schemes are presented to solve two-dimensional nonlinear Riesz Space Fractional Diffusion Equations (RSFDEs). The proposed schemes employ the Crank-Nicolson method for temporal discretization, a quartic approximation for the Riesz derivative, and an explicit treatment for linearizing the nonlinear term. Through a rigorous discrete energy analysis, we derive explicit error estimates, proving that both schemes achieve second-order accuracy in time and fourth-order accuracy in space. The ADI and LOD methods are employed to decompose two-dimensional problems into a collection of Symmetric Positive Definite (SPD) Toeplitz systems, for which the fast sine transform can be utilized to mitigate computational complexity. The computational treatment of these subsystems is further conducted using a Preconditioned Conjugate Gradient (PCG) approach that leverages sine transforms. Theoretical analysis demonstrates that the preconditioned matrix can be decomposed into the identity matrix, a matrix of small norm, and a low-rank matrix. The effectiveness and efficiency of the scheme are validated through numerical examples.

Keywords: linearized scheme, nonlinear Riesz space-fractional diffusion equations, τ-preconditioner, preconditioned conjugated gradient method

CLC number: O241.82 **Document code**: A **Article ID**: 1005-9113(2025)00-0000-15

0 Introduction

Considerable attention has been focused on Space Fractional Differential Equations (SFDEs) in recent decades due to their extensive applications in diverse fields, such as fluid mechanics[1], physics[2], and finance[3-4]. However, the nonlocality inherent in fractional calculus means that closed-form solutions to these equations are rarely attainable, making the development of robust numerical methods essential.

Consequently, numerous numerical schemes have been proposed for solving fractional diffusion equations[5-13]. To date, two principal methodologies have been widely adopted for the numerical approximation of Riesz fractional derivatives. The first approach, introduced by Ortigueira[14], is based on the fractional centered difference scheme. This method was later applied by Çelik et al.[15] to Riesz Space Fractional Diffusion Equations (RSFDEs), demonstrating that it achieves second-order convergence. The second major methodology originates from the work of Meerschaert et al.[5], who proposed a shifted Grünwald-Letnikov approximation for the Riemann-Liouville (RL) derivative. This foundation subsequently facilitated the development of the Weighted Shifted Grünwald-Letnikov Difference (WSGD) scheme for solving RSFDEs[16]. Further advancing this line of research, Zhou et al.[17] constructed a quasi-compact discretization scheme with third-order accuracy. More recently, building on the WSGD approach, Hao et al.[18] introduced a higher-order scheme attaining fourth-order precision.

However, dense coefficient matrices are generated in the discretized systems because of the non-local nature of fractional derivatives, thereby imposing heavy computational and memory burdens on solving the resulting linear systems. With the

Received 2025-08-18.

Sponsored by the National Natural Science Foundation of China (Grant No. 11801216), Natural Science Foundation of Shandong Province (Grant No. ZR2017BA010).

* Corresponding author. Tao Wang, Ph.D, Associated Professor. Email: wtpeak@163.com.

development of numerical methods, several robust preconditioners have been proposed to tackle this issue in Toeplitz systems. For instance, Yu et al.[19] formulated Preconditioned Conjugate Gradient (PCG) methods based on T. Chan[20] and R. Chan's[21] circulant preconditioners to address the challenges of solving Riesz distributed-order nonlinear SFDEs. Meanwhile, utilizing the approximate inverse preconditioning technique based on the inverse of R. Chan's[21] circulant matrix, Tang et al.[22] developed an efficient second-order numerical method for solving SFDEs with variable coefficients. The preconditioning approach incorporating the τ-preconditioner has emerged as a promising method for solving SFDEs[23-25].

When it comes to modeling and simulating multidimensional systems, the numerical discretization of partial differential equations, which is fundamental to the process, proves to be computationally expensive. By reducing multidimensional partial differential equations to a sequence of one-dimensional problems, Alternating Direction Implicit (ADI) and Locally One-Dimensional (LOD) methods offer a simplified approach [26-30]. Recently, Tang et al.[31] employed the ADI method to decompose the discretized linear system into multiple subsystems, which yielded first-order accuracy in space. They then combined the Generalized Minium Residual (GMRES) method with the constructed Large Language Model-Augmented Semantic Digital Twins (LSDTs) τ-preconditioner to solve these subsystems. Chen et al.[32] developed a second-order LOD multigrid method for multidimensional RSFDEs. Combining the Backward Differentiation Formula (BDF) method with a fourth-order fractional compact operator, Hu and Cao[33] constructed a high-accuracy ADI scheme for 2D nonlinear RSFDEs. However, the development of a fast algorithm was not addressed. The scarcity of high-order fast algorithms based on the ADI and LOD methods for multidimensional RSFDEs has motivated our research in this direction.

In this study, we investigate the subsequent 2D nonlinear RSFDEs:

$$\begin{cases} \dfrac{\partial u}{\partial t} = K_\beta \dfrac{\partial^\beta u}{\partial |x|^\beta} + K_\gamma \dfrac{\partial^\gamma u}{\partial |y|^\gamma} + f(u) + p, \\ \quad (x,y) \in \Omega, t \in (0,T] \\ u(x,y,0) = \varphi(x,y), (x,y) \in \Omega \\ u(x,y,t) = 0, (x,y) \in \partial\Omega, t \in (0,T] \end{cases} \tag{1}$$

where p and φ are sufficiently smooth in their respective domains. $\beta, \gamma \in (1,2)$ and $\Omega = (0,L_1) \times (0,L_2)$, L_1, L_2 and T are positive constants. $K_\beta, K_\gamma > 0$, $f(u)$ is a continuous function and satisfies the Lipschitz condition. Moreover, $\dfrac{\partial^\beta u}{\partial |x|^\beta}$ and $\dfrac{\partial^\gamma u}{\partial |y|^\gamma}$ denote the Riesz fractional derivative[23, 25].

In this study, fourth-order linearized ADI and LOD schemes are proposed to solve 2D nonlinear problem (1). Both proposed schemes employ the Crank-Nicolson method for temporal discretization and a fourth-order approximation[18] for the Riesz derivative, along with explicit linearization of the nonlinear terms. Both schemes are proven to achieve second-order temporal accuracy and fourth-order spatial accuracy, as established through explicit error estimates derived from a discrete energy analysis. The ADI and LOD methods are employed to decompose two-dimensional problems into a collection of Symmetric Positive Definite (SPD) Toeplitz systems, for which the fast sine transform can be utilized to mitigate computational complexity. The solution of these subsystems is achieved through the application of the PCG method preconditioned by the τ-matrix approximation. In addition, theoretical analysis confirms that the τ-preconditioner matrix exhibits eigenvalues tightly clustered around unity, while numerical results substantiate the extended τ-preconditioner's remarkable efficacy in accelerating convergence for two-dimensional problems.

The paper is structured as follows. In Section 1, linearized fourth-order ADI and LOD schemes are developed for the nonlinear RSFDEs. The PCG method with a τ-preconditioner is presented in Section 2 to accelerate the solution of discrete linear systems, and the spectral characteristics of the preconditioned matrices are analyzed. In Section 3, numerical examples are conducted to demonstrate the convergence behavior and computational efficiency of the fast numerical schemes incorporating preconditioners. A summary of key findings is offered in Section 4.

1 Linearized Numerical Scheme

1.1 Derivation of the Numerical Scheme

Take positive integers M_1, M_2 and N. Denote $\Delta x = L_1/(M_1+1)$, $\Delta y = L_2/(M_2+1)$ and $\Delta t = T/N$, Let $[[k_0, k_1]] = \{k \mid k_0 \leqslant k \leqslant k_1,$ and $k \in \mathbb{N}\}$, $x_i =$

$i\Delta x(i\in[[0,M_1+1]])$, $y_j=i\Delta y(j\in[[0,M_2+1]])$, $t_k=k\Delta t(k\in[[0,N]])$, $\bar{\Omega}_h\{(x_i,y_i)\mid i\in[[0,M_1+1]],j\in[[0,M_2+1]]\}$, $\Omega_h=\bar{\Omega}_h\cap\Omega$, $\partial\Omega_h=\bar{\Omega}_h\cap\partial\Omega$, $\omega=\{(i,j)\mid(x_i,y_j)\in\Omega_h\}$, $\partial\omega=\{(i,j)\mid(x_i,y_j)\in\partial\Omega_h\}$, $\bar{\omega}=\omega\cup\partial\omega$, $\Omega_{\Delta t}=\{t_k\mid 0\leqslant k\leqslant N\}$, $V_h=\{u\mid u=\{u_{i,j}\mid(i,j)\in\bar{\omega}\}\}$.

Given any discrete function $u\in V_h$, we define

$$u_{i,j}^{k+\frac{1}{2}}=\frac{1}{2}(u_{i,j}^{k+1}+u_{i,j}^{k}),\ \delta_t u_{i,j}^{k+\frac{1}{2}}=\frac{1}{\Delta t}(u_{i,j}^{k+1}-u_{i,j}^{k})$$

$$\mathcal{H}_x^{\beta}u_{i,j}=\begin{cases}d_2^{\beta}u_{i-1,j}+(1-2d_2^{\beta})u_{i,j}+d_2^{\beta}u_{i+1,j},\\ \quad i\in[[1,M_1]],j\in[[0,M_2+1]]\\ u_{i,j},i=0,\ M_1+1,j\in[[0,M_2+1]]\end{cases}$$

$$\mathcal{H}_y^{\gamma}u_{i,j}=\begin{cases}d_2^{\gamma}u_{i,j-1}+(1-2d_2^{\gamma})u_{i,j}+d_2^{\gamma}u_{i,j+1},\\ \quad j\in[[1,M_2]],i\in[[0,M_1+1]]\\ u_{i,j},j=0,\ M_2+1,\quad i\in[[0,M_1+1]]\end{cases}$$

where

$$d_2^{\beta}=\frac{-\beta^2+\beta+4}{24}\in\left(\frac{1}{12},\frac{1}{6}\right)$$

$$d_2^{\gamma}=\frac{-\gamma^2+\gamma+4}{24}\in\left(\frac{1}{12},\frac{1}{6}\right)$$

Define

$$U_{i,j}^{n}=u(x_i,y_j,t_n),p_{i,j}^{n}=p(x_i,y_j,t_n)$$

$$p_{i,j}^{n+\frac{1}{2}}=\frac{1}{2}(p_{i,j}^{n+1}+p_{i,j}^{n}),(i,j)\in\bar{\omega},n\in[[0,N-1]]$$

Let

$$\begin{cases}\hat{w}_0^{(\alpha)}=\dfrac{\alpha^2+3\alpha+2}{12}g_0^{(\alpha)}\\ \hat{w}_1^{(\alpha)}=\dfrac{\alpha^2+3\alpha+2}{12}g_1^{(\alpha)}+\dfrac{4-\alpha^2}{6}g_0^{(\alpha)}\\ \hat{w}_k^{(\alpha)}=\dfrac{\alpha^2+3\alpha+2}{12}g_k^{(\alpha)}+\dfrac{4-\alpha^2}{6}g_{k-1}^{(\alpha)}+\\ \quad\dfrac{\alpha^2-3\alpha+2}{12}g_{k-2}^{(\alpha)}\end{cases}$$

where $k\geqslant 2$, $g_k^{(\alpha)}=(-1)^k\begin{pmatrix}\alpha\\k\end{pmatrix}$.

Lemma 1[18] Assuming $\alpha\in(1,2)$, the coefficients $\hat{w}_k^{(\alpha)}$ satisfy

$$\begin{cases}\hat{w}_0^{(\alpha)}>0,\hat{w}_1^{(\alpha)}\leqslant 0,\hat{w}_k^{(\alpha)}\geqslant 0,k\geqslant 3\\ \displaystyle\sum_{k=0}^{\infty}\hat{w}_k^{(\alpha)}=0,\sum_{k=0}^{m}\hat{w}_k^{(\alpha)}\leqslant 0,m\geqslant 2\\ \hat{w}_0^{(\alpha)}+\hat{w}_2^{(\alpha)}\geqslant 0\end{cases}$$

Let

$$W_x^{\beta}u=\frac{\partial^{\beta}u}{\partial|x|^{\beta}},\ W_y^{\gamma}u=\frac{\partial^{\gamma}u}{\partial|y|^{\gamma}}$$

Define the following difference operator:

$$\delta_x^{\beta}u_{i,j}^{n}=-\frac{1}{2\cos(\beta\pi/2)}\Delta x^{-\beta}\cdot\left(\sum_{k=0}^{i+1}\hat{w}_k^{(\beta)}u_{i-k+1,j}^{n}+\sum_{k=0}^{M_1-i+2}\hat{w}_k^{(\beta)}u_{i+k+1,j}^{n}\right)$$

$$\delta_y^{\gamma}u_{i,j}^{n}=-\frac{1}{2\cos(\gamma\pi/2)}\Delta y^{-\gamma}\cdot\left(\sum_{k=0}^{j+1}\hat{w}_k^{(\gamma)}u_{i,j-k+1}^{n}+\sum_{k=0}^{M_2-j+2}\hat{w}_k^{(\gamma)}u_{i,j+k+1}^{n}\right)$$

Lemma 2[34] For any $u(x,y,\cdot)\in C^3[0,T]$, we have

$$\frac{1}{2}\left(\frac{\partial u(x_i,y_j,t_{n+1})}{\partial t}+\frac{\partial u(x_i,y_j,t_n)}{\partial t}\right)=\delta_t U_{i,j}^{n+\frac{1}{2}}+O(\Delta t^2)$$

Denote

$$\ell^{4+\alpha}(\mathbb{R})=\{f\mid f\in L^1(\mathbb{R}),\int_{-\infty}^{\infty}(1+|k|)^{4+\alpha}|F(k)|\mathrm{d}k<\infty\},\ \alpha=\beta,\gamma$$

where $F(k)$ is represented as the Fourier transformation of $f(t)$.

Lemma 3[18] Let $u(\cdot,y,t)\in\ell^{4+\beta}(\mathbb{R})$, and $u(x,\cdot,t)\in\ell^{4+\gamma}(\mathbb{R})$, we have

$$\mathcal{H}_x^{\beta}W_x^{\beta}u(x_i,y_j,t_n)=\delta_x^{\beta}U_{i,j}^{n}+O(\Delta x^4)$$

$$\mathcal{H}_y^{\gamma}W_y^{\gamma}u(x_i,y_j,t_n)=\delta_y^{\gamma}U_{i,j}^{n}+O(\Delta y^4)$$

Lemma 4 Let $f(u)\in C^2(R)$, then we have

$$\begin{cases}\dfrac{f(U^{n+1})+f(U^n)}{2}=\dfrac{3f(U^n)-f(U^{n-1})}{2}+\\ \quad O(\Delta t^2),n\geqslant 1\\ \dfrac{f(U^1)+f(U^0)}{2}=f(U^0)+O(\Delta t)\end{cases}$$

Proof: By employing a Taylor series expansion, the correctness of the formula can be directly validated.

Building upon the foregoing lemmas, we now proceed to derive high-order numerical schemes for solving problem (1), under the assumption that the solution $u(x,y,t)$ adheres to the conditions stipulated in Lemmas 2 and 3. Considering the main equation in problem (1) at (x_i,y_j,t_k), and averaging two adjacent time layers, we have

$$\begin{aligned}&\frac{1}{2}\left(\frac{\partial u(x_i,y_j,t_{n+1})}{\partial t}+\frac{\partial u(x_i,y_j,t_n)}{\partial t}\right)=\\&\frac{1}{2}K_{\beta}(W_x^{\beta}u(x_i,y_j,t_{n+1})+W_x^{\beta}u(x_i,y_j,t_n))+\\&\frac{1}{2}K_{\gamma}(W_y^{\gamma}u(x_i,y_j,t_{n+1})+W_y^{\gamma}u(x_i,y_j,t_n))+\\&\frac{1}{2}(f(u_{ij}^{n+1})+f(u_{ij}^{n}))+p_{ij}^{n+\frac{1}{2}},(i,j)\in\omega,\\&n\in[[0,N-1]]\end{aligned}\tag{2}$$

Applying $\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}$ to Eq.(2) and using Lemmas 2–

4, we have

$$\mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}} = K_\beta \mathcal{H}_y^\gamma \delta_x^\beta U_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma U_{i,j}^{n+\frac{1}{2}} + \frac{3}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^n) - \frac{1}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^{n-1}) + \mathcal{H}_x^\beta \mathcal{H}_y^\gamma p_{i,j}^{n+\frac{1}{2}} + (R_1)_{i,j}^{n+\frac{1}{2}}, (i,j) \in \omega, n \in [[0,N-1]] \tag{3}$$

where $U_{i,j}^{-1} = U_{i,j}^0$, and

$$\begin{cases} (R_1)_{i,j}^{n+\frac{1}{2}} = O(\Delta t^2 + \Delta x^4 + \Delta y^4), \ n \geqslant 1 \\ (R_1)_{i,j}^{\frac{1}{2}} = O(\Delta t + \Delta x^4 + \Delta y^4) \end{cases}$$

Adding $\frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}}$ to both sides of Eq. (3) gives

$$\mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}} = K_\beta \mathcal{H}_y^\gamma \delta_x^\beta U_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma U_{i,j}^{n+\frac{1}{2}} + \frac{3}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^n) - \frac{1}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^{n-1}) + \mathcal{H}_x^\beta \mathcal{H}_y^\gamma p_{i,j}^{n+\frac{1}{2}} + R_{i,j}^{n+\frac{1}{2}}, \ (i,j) \in \omega, \ n \in [[0,N-1]] \tag{4}$$

where

$$\begin{cases} R_{i,j}^{n+\frac{1}{2}} = O(\Delta t^2 + \Delta x^4 + \Delta y^4), \ n \geqslant 1 \\ R_{i,j}^{\frac{1}{2}} = O(\Delta t + \Delta x^4 + \Delta y^4) \end{cases} \tag{5}$$

By omitting the small terms in Eq. (4) and substituting the exact solution $U_{i,j}^n$ with an approximated solution $u_{i,j}^n$, the following linearized numerical scheme for solving problem (1) is obtained

$$\begin{cases} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t u_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \delta_t u_{i,j}^{n+\frac{1}{2}} = \\ \quad K_\beta \mathcal{H}_y^\gamma \delta_x^\beta u_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma u_{i,j}^{n+\frac{1}{2}} + \mathcal{H}_x^\beta \mathcal{H}_y^\gamma g_{i,j}^{n+\frac{1}{2}}, \\ \quad (i,j) \in \omega, \ n \in [[0,N-1]] \\ u_{i,j}^0 = \varphi(x_i, y_j), (i,j) \in \omega \\ u_{i,j}^n = 0, \ (i,j) \in \partial\omega, \ n \in [[0,N-1]] \end{cases} \tag{6}$$

where $g_{i,j}^{n+\frac{1}{2}} = \frac{3}{2} f(u_{i,j}^n) - \frac{1}{2} f(u_{i,j}^{n-1}) + p_{i,j}^{n+\frac{1}{2}}$.

Multiplying both sides of the main equation in Eq.(6) by Δt yields the resulting form

$$(\mathcal{H}_x^\beta - \frac{K_\beta \Delta t}{2} \delta_x^\beta)(\mathcal{H}_y^\gamma - \frac{K_\gamma \Delta t}{2} \delta_y^\gamma) u_{i,j}^{n+1} = (\mathcal{H}_x^\beta + \frac{K_\beta \Delta t}{2} \delta_x^\beta)(\mathcal{H}_y^\gamma + \frac{K_\gamma \Delta t}{2} \delta_y^\gamma) u_{i,j}^n + \Delta t \mathcal{H}_x^\beta \mathcal{H}_y^\gamma g_{i,j}^{n+\frac{1}{2}}$$

By introducing the intermediate variable $v_{i,j}^*$, we have the following linearized ADI scheme

$$\begin{cases} (\mathcal{H}_x^\beta - \frac{\Delta t}{2} \delta_x^\beta) v_{i,j}^* = (\mathcal{H}_x^\beta + \frac{\Delta t}{2} \delta_x^\beta) \cdot \\ \quad (\mathcal{H}_y^\gamma + \frac{\Delta t}{2} \delta_y^\gamma) u_{i,j}^n + \Delta t \mathcal{H}_x^\beta \mathcal{H}_y^\gamma g_{i,j}^{n+\frac{1}{2}} \\ (\mathcal{H}_y^\gamma - \frac{\Delta t}{2} \delta_y^\gamma) u_{i,j}^{n+1} = v_{i,j}^* \end{cases} \tag{7}$$

We now develop a linearized LOD scheme. Adding small term

$$\frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \left(\frac{3}{2} f(U_{i,j}^n) - \frac{1}{2} f(U_{i,j}^{n-1}) + p_{i,j}^{n+\frac{1}{2}} \right)$$

to both sides of Eq. (4) yields

$$\begin{aligned} &\mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \delta_t U_{i,j}^{n+\frac{1}{2}} = \\ &K_\beta \mathcal{H}_y^\gamma \delta_x^\beta U_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma U_{i,j}^{n+\frac{1}{2}} + \\ &\frac{3}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^n) - \frac{1}{2} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma f(U_{i,j}^{n-1}) + \\ &\mathcal{H}_x^\beta \mathcal{H}_y^\gamma p_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \cdot \\ &\left(\frac{3}{2} f(U_{i,j}^n) - \frac{1}{2} f(U_{i,j}^{n-1}) + p_{i,j}^{n+\frac{1}{2}} \right) + \bar{R}_{i,j}^{n+\frac{1}{2}}, \\ &(i,j) \in \omega, \ n \in [[0,N-1]] \end{aligned} \tag{8}$$

where

$$\begin{cases} \bar{R}_{ij}^{n+\frac{1}{2}} = O(\Delta t^2 + \Delta x^4 + \Delta y^4), \ n \geqslant 1 \\ \bar{R}_{ij}^{\frac{1}{2}} = O(\Delta t + \Delta x^4 + \Delta y^4) \end{cases} \tag{9}$$

By omitting the small terms in Eq. (8) and substituting the exact solution $U_{i,j}^n$ with an approximated solution $u_{i,j}^n$, the following linearized numerical scheme for solving problem (1) is obtained

$$\begin{cases} \mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t u_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma \delta_t u_{i,j}^{n+\frac{1}{2}} = \\ \quad K_\beta \mathcal{H}_y^\gamma \delta_x^\beta u_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma u_{i,j}^{n+\frac{1}{2}} + \\ \quad \mathcal{H}_x^\beta \mathcal{H}_y^\gamma g_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4} \delta_x^\beta \delta_y^\gamma g_{i,j}^{n+\frac{1}{2}}, \\ \quad (i,j) \in \omega, n \in [[0,N-1]] \\ u_{i,j}^0 = \varphi(x_i, y_j), (i,j) \in \omega \\ u_{i,j}^n = 0, (i,j) \in \partial\omega, n \in [[0,N-1]] \end{cases} \tag{10}$$

We multiply both sides of the main equation in Eq.(10) by Δt and reformulate it as

$$(\mathcal{H}_x^\beta - \frac{K_\beta \Delta t}{2} \delta_x^\beta)(\mathcal{H}_y^\gamma - \frac{K_\gamma \Delta t}{2} \delta_y^\gamma) u_{i,j}^{n+1} = (\mathcal{H}_x^\beta + \frac{K_\beta \Delta t}{2} \delta_x^\beta)(\mathcal{H}_y^\gamma + \frac{K_\gamma \Delta t}{2} \delta_y^\gamma) u_{i,j}^n +$$

$$\Delta t\,\mathcal{H}_x^\beta\,\mathcal{H}_y^\gamma g_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^3}{4}\delta_x^\beta \delta_y^\gamma g_{i,j}^{n+\frac{1}{2}}$$

By introducing the intermediate variable $u_{i,j}^*$, we have the following linearized LOD scheme

$$\begin{cases} (\mathcal{H}_x^\beta - \frac{K_\beta \Delta t}{2}\delta_x^\beta) u_{i,j}^* = (\mathcal{H}_x^\beta + \frac{K_\beta \Delta t}{2}\delta_x^\beta) u_{i,j}^n + \\ \quad \frac{\Delta t}{2}(\mathcal{H}_x^\beta + \frac{K_\beta \Delta t}{2}\delta_x^\beta) g_{i,j}^{n+\frac{1}{2}} \\ (\mathcal{H}_y^\gamma - \frac{K_\gamma \Delta t}{2}\delta_y^\gamma) u_{i,j}^{n+1} = (\mathcal{H}_y^\gamma + \frac{K_\gamma \Delta t}{2}\delta_y^\gamma) u_{i,j}^* + \\ \quad \frac{\Delta t}{2}(\mathcal{H}_y^\gamma - \frac{K_\gamma \Delta t}{2}\delta_y^\gamma) g_{i,j}^{n+\frac{1}{2}} \end{cases} \tag{11}$$

Let

$$\mu_\beta = \frac{K_\beta \Delta t}{4\Delta x^\beta \cos(\beta\pi/2)}, \mu_\gamma = \frac{K_\gamma \Delta t}{4\Delta y^\gamma \cos(\gamma\pi/2)}$$

and we assume that $|\mu_\beta|$ and $|\mu_\gamma|$ share a common upper bound, i.e.,

$$\max\{|\mu_\beta|, |\mu_\gamma|\} < \mu \tag{12}$$

where μ is a real number.

Let

$$s_k^{(\alpha,\pm)} = \begin{cases} -2d_2^\alpha \pm 2\mu_\alpha \hat{w}_1^{(\alpha)}, & k = 0 \\ d_2^\alpha \pm \mu_\alpha(\hat{w}_0^{(\alpha)} + \hat{w}_2^{(\alpha)}), & |k| = 1 \\ \pm\mu_\alpha \hat{w}_{k+1}^{(\alpha)}, & |k| \in [[2, M-1]] \end{cases} \tag{13}$$

where $(M,\alpha) = (M_1,\beta)$ or (M_2,γ).

Define

$$\boldsymbol{B}_\alpha^\pm = \boldsymbol{I} + \boldsymbol{S}_\pm^{(\alpha)}, \alpha = \beta, \gamma \tag{14}$$

where

$$\boldsymbol{S}_\pm^{(\alpha)} = \begin{bmatrix} s_0^{(\alpha,\pm)} & s_{-1}^{(\alpha,\pm)} & \cdots & s_{2-M}^{(\alpha,\pm)} & s_{1-M}^{(\alpha,\pm)} \\ s_1^{(\alpha,\pm)} & s_0^{(\alpha,\pm)} & \cdots & s_{3-M}^{(\alpha,\pm)} & s_{2-M}^{(\alpha,\pm)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{M-2}^{(\alpha,\pm)} & s_{M-3}^{(\alpha,\pm)} & \cdots & s_0^{(\alpha,\pm)} & s_{-1}^{(\alpha,\pm)} \\ s_{M-1}^{(\alpha,\pm)} & s_{M-2}^{(\alpha,\pm)} & \cdots & s_1^{(\alpha,\pm)} & s_0^{(\alpha,\pm)} \end{bmatrix} \tag{15}$$

where $(M,\alpha) = (M_1,\beta)$ or (M_2,γ).

Clearly, $\boldsymbol{B}_\alpha^\pm$ and $\boldsymbol{S}_\pm^{(\alpha)}$ are the symmetric Toeplitz (ST) matrices.

We define the following M-order symmetric tridiagonal matrices:

$$\boldsymbol{C}^{(\alpha)} = \mathrm{tridiag}(d_2^\alpha, 1 - 2d_2^\alpha, d_2^\alpha)$$

where $(M,\alpha) = (M_1,\beta)$ or (M_2,γ).

Define

$$\boldsymbol{u}^n = (u_{i,j}^n)_{M_1\times M_2}, \boldsymbol{g}^{n+\frac{1}{2}} = (g_{i,j}^{n+\frac{1}{2}})_{M_1\times M_2}$$

$$\bar{\boldsymbol{U}}^n = \boldsymbol{B}_\beta^- ((\boldsymbol{B}_\gamma^-)^{\mathrm{T}} (\boldsymbol{u}^n)^{\mathrm{T}})^{\mathrm{T}} = (\bar{U}_{i,j}^n)_{M_1\times M_2}$$

$$\boldsymbol{G}^{n+\frac{1}{2}} = \boldsymbol{C}^{(\beta)} ((\boldsymbol{C}^{(\gamma)})^{\mathrm{T}} (\boldsymbol{g}^{n+\frac{1}{2}})^{\mathrm{T}})^{\mathrm{T}} = (G_{i,j}^{n+\frac{1}{2}})_{M_1\times M_2}$$

$$\bar{\boldsymbol{U}}_{\Delta x,j}^n = (\bar{U}_{1,j}^n, \bar{U}_{2,j}^n, \cdots, \bar{U}_{M_1,j}^n)^{\mathrm{T}}$$

$$\boldsymbol{G}_{\Delta x,j}^{n+\frac{1}{2}} = (G_{1,j}^{n+\frac{1}{2}}, G_{2,j}^{n+\frac{1}{2}}, \cdots, G_{M_1,j}^{n+\frac{1}{2}})^{\mathrm{T}}$$

$$\boldsymbol{v}_{\Delta x,j}^* = (u_{1,j}^*, u_{2,j}^*, \cdots, u_{M_1,j}^*)^{\mathrm{T}}$$

$$\boldsymbol{v}_{i,\Delta y}^* = (u_{i,1}^*, u_{i,2}^*, \cdots, u_{i,M_2}^*)^{\mathrm{T}}$$

$$\boldsymbol{u}_{\Delta x,j}^n = (u_{1,j}^n, u_{2,j}^n, \cdots, u_{M_1,j}^n)^{\mathrm{T}}$$

$$\boldsymbol{u}_{i,\Delta y}^{n+1} = (u_{i,1}^{n+1}, u_{i,2}^{n+1}, \cdots, u_{i,M_2}^{n+1})^{\mathrm{T}}$$

$$\boldsymbol{u}_{\Delta x,j}^* = (u_{1,j}^*, u_{2,j}^*, \cdots, u_{M_1,j}^*)^{\mathrm{T}}$$

$$\boldsymbol{u}_{i,\Delta y}^* = (u_{i,1}^*, u_{i,2}^*, \cdots, u_{i,M_2}^*)^{\mathrm{T}}$$

$$\boldsymbol{g}_{\Delta x,j}^{n+\frac{1}{2}} = (g_{1,j}^{n+\frac{1}{2}}, g_{2,j}^{n+\frac{1}{2}} \cdots, g_{M_1,j}^{n+\frac{1}{2}})^{\mathrm{T}}$$

$$\boldsymbol{g}_{i,\Delta y}^{n+\frac{1}{2}} = (g_{i,1}^{n+\frac{1}{2}}, g_{i,2}^{n+\frac{1}{2}} \cdots, g_{i,M_2}^{n+\frac{1}{2}})^{\mathrm{T}}$$

We then derive the matrix form of the linearized ADI scheme as follows:

$$\begin{cases} \boldsymbol{B}_\beta^+ \boldsymbol{v}_{\Delta x,j}^* = \bar{\boldsymbol{U}}_{\Delta x,j}^n + \Delta t \boldsymbol{G}_{\Delta x,j}^{n+\frac{1}{2}}, j = 1,2,\cdots,M_2 \\ \boldsymbol{B}_\gamma^+ \boldsymbol{u}_{i,\Delta y}^{n+1} = \boldsymbol{v}_{i,\Delta y}^*, i = 1,2,\cdots,M_1 \end{cases} \tag{16}$$

Let

$$\boldsymbol{\sigma}_{\Delta x,j} = \begin{bmatrix} \hat{w}_2^{(\beta)}\rho_{0,j}^n \\ \hat{w}_3^{(\beta)}\rho_{0,j}^n \\ \vdots \\ \hat{w}_{M_1}^{(\beta)}\rho_{0,j}^n \\ \hat{w}_{M_1+1}^{(\beta)}\rho_{0,j}^n \end{bmatrix} + \begin{bmatrix} \hat{w}_{M_1+1}^{(\beta)}\rho_{M_1+1,j}^n \\ \hat{w}_{M_1}^{(\beta)}\rho_{M_1+1,j}^n \\ \vdots \\ \hat{w}_3^{(\beta)}\rho_{M_1+1,j}^n \\ \hat{w}_2^{(\beta)}\rho_{M_1+1,j}^n \end{bmatrix} + \begin{bmatrix} \hat{w}_0^{(\beta)}\rho_{0,j}^n \\ 0 \\ \vdots \\ 0 \\ \hat{w}_0^{(\beta)}\rho_{M_1+1,j}^n \end{bmatrix} + \begin{bmatrix} d_2^\beta \varrho_{0,j}^n \\ 0 \\ \vdots \\ 0 \\ d_2^\beta \varrho_{M_1+1,j}^n \end{bmatrix}$$

$$\boldsymbol{\sigma}_{i,\Delta y} = \begin{bmatrix} \hat{w}_2^{(\gamma)}\xi_{i,0}^n \\ \hat{w}_3^{(\gamma)}\xi_{i,0}^n \\ \vdots \\ \hat{w}_{M_2}^{(\gamma)}\xi_{i,0}^n \\ \hat{w}_{M_2+1}^{(\gamma)}\xi_{i,0}^n \end{bmatrix} + \begin{bmatrix} \hat{w}_{M_2+1}^{(\gamma)}\xi_{i,M_2+1}^n \\ \hat{w}_{M_2}^{(\gamma)}\xi_{i,M_2+1}^n \\ \vdots \\ \hat{w}_3^{(\gamma)}\xi_{i,M_2+1}^n \\ \hat{w}_2^{(\gamma)}\xi_{i,M_2+1}^n \end{bmatrix} + \begin{bmatrix} \hat{w}_0^{(\gamma)}\xi_{i,0}^n \\ 0 \\ \vdots \\ 0 \\ \hat{w}_0^{(\gamma)}\xi_{i,M_2+1}^n \end{bmatrix} + \begin{bmatrix} d_2^\gamma \zeta_{i,0}^n \\ 0 \\ \vdots \\ 0 \\ d_2^\gamma \zeta_{i,M_2+1}^n \end{bmatrix}$$

where

$$\rho_{k,j}^n = -\mu_\beta(\frac{\Delta t}{2} g_{k,j}^{n+\frac{1}{2}} + u_{k,j}^n + u_{k,j}^*)$$

$$\varrho_{k,j}^n = \frac{\Delta t}{2} g_{k,j}^{n+\frac{1}{2}} + u_{k,j}^n - u_{k,j}^*, k = 0, M_1 + 1$$

$$\xi_{i,k}^n = \mu_\gamma(\frac{\Delta t}{2} g_{i,k}^{n+\frac{1}{2}} - u_{i,k}^{n+1} - u_{i,k}^*)$$

$$\zeta_{i,k}^{n}=\frac{\Delta t}{2}g_{i,k}^{n+\frac{1}{2}}-u_{i,k}^{n+1}+u_{i,k}^{*},k=0,M_2+1$$

Then, we arrive at the following matrix form of the linearized LOD scheme:

$$\begin{cases}\boldsymbol{B}_{\beta}^{+}\boldsymbol{u}_{\Delta x,j}^{*}=\boldsymbol{B}_{\beta}^{-}\boldsymbol{u}_{\Delta x,j}^{n}+\dfrac{\Delta t}{2}\boldsymbol{B}_{\beta}^{-}\boldsymbol{g}_{\Delta x,j}^{n+\frac{1}{2}}+\boldsymbol{\sigma}_{\Delta x,j},\\ j=0,1,\cdots,M_2+1\\ \boldsymbol{B}_{\gamma}^{+}\boldsymbol{u}_{i,\Delta y}^{n+1}=\boldsymbol{B}_{\gamma}^{-}\boldsymbol{u}_{i,\Delta y}^{*}+\dfrac{\Delta t}{2}\boldsymbol{B}_{\gamma}^{+}\boldsymbol{g}_{i,\Delta y}^{n+\frac{1}{2}}+\boldsymbol{\sigma}_{i,\Delta y},\\ i=1,2,\cdots,M_1\end{cases}\tag{17}$$

1.2 Error Estimate

Let $\mathring{V}_h=\{u \mid u\in V_h;\ \text{if } (i,j)\in\partial\omega, u_{i,j}=0\}$. Given any discrete functions $u,\ v\in\mathring{V}_h$, we define

$$(u,\ v)=\Delta x\Delta y\sum_{i=1}^{M_1}\sum_{j=1}^{M_2}u_{i,j}v_{i,j},\quad \|u\|=\sqrt{(u,\ u)}$$

Since $\mathcal{H}_x^{\beta}$ and $\mathcal{H}_y^{\gamma}$ are positive definite and self-adjoint, we have $\mathcal{H}_x^{\beta}=(\mathcal{Q}_x)^2$ and $\mathcal{H}_y^{\gamma}=(\mathcal{Q}_y)^2$ [18].

Lemma 5[18] For any grid function $u\in\mathring{V}_h$, we have

$$\frac{1}{9}\|u\|^2\leqslant(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}u,u)=(\mathcal{Q}_x\mathcal{Q}_yu,\mathcal{Q}_x\mathcal{Q}_yu)\leqslant \|u\|^2,(\delta_x^{\beta}u,u)\leqslant 0,(\delta_y^{\gamma}u,u)\leqslant 0$$

According to the definitions of δ_x^{β} and δ_y^{γ}, it is straightforward to show that both $-\delta_x^{\beta}$ and $-\delta_y^{\gamma}$ are symmetric positive definite operators. Therefore, there exist fractional symmetric positive definite difference operators Λ_x and Λ_y such that the following conclusions hold[18]:

$$(\delta_x^{\beta}u,u)=-(\Lambda_xu,\Lambda_xu)\leqslant 0$$
$$(\delta_y^{\gamma}u,u)=-(\Lambda_yu,\Lambda_yu)\leqslant 0$$
$$(\delta_x^{\beta}\delta_y^{\gamma}u,u)=(\Lambda_x\Lambda_yu,\Lambda_x\Lambda_yu)$$

Lemma 6 For any grid function $u\in\mathring{V}_h$, we have

$$((\mathcal{H}_y^{\gamma}\delta_x^{\beta}+\mathcal{H}_x^{\beta}\delta_y^{\gamma})u,u)\leqslant 0$$

Proof: By Lemma 5, we get

$$(\mathcal{H}_y^{\gamma}\delta_x^{\beta}u,u)=(\delta_x^{\beta}\mathcal{Q}_yu,\mathcal{Q}_yu)\leqslant 0$$
$$(\mathcal{H}_x^{\beta}\delta_y^{\gamma}u,u)=(\delta_y^{\gamma}\mathcal{Q}_xu,\mathcal{Q}_xu)\leqslant 0$$

The proof is completed.

Lemma 7[30] Assume that $\{\varsigma_n\}$ and $\{\eta_n\}$ are nonnegative sequences, and that the sequence $\{\theta_n\}$ satisfies

$$\phi_0\geqslant 0,\theta_0\leqslant\phi_0$$
$$\theta_n\leqslant\phi_0+\sum_{l=0}^{n-1}\eta_l+\sum_{l=0}^{n-1}\varsigma_l\theta_l,n\geqslant 1$$

Then it holds that

$$\theta_n\leqslant\left(\phi_0+\sum_{l=0}^{n-1}\eta_l\right)\exp\left(\sum_{l=0}^{n-1}\varsigma_l\right),n\geqslant 1$$

We begin with an analysis of the stability and convergence of the linearized ADI scheme in Eq. (7). Let $e_{i,j}^n=U_{i,j}^n-u_{i,j}^n$, by subtracting the master equation in Eq. (6) from Eq. (4), we get

$$\begin{aligned}&\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}\delta_te_{i,j}^{n+\frac{1}{2}}+\frac{K_{\beta}K_{\gamma}\Delta t^2}{4}\delta_x^{\beta}\delta_y^{\gamma}\delta_te_{i,j}^{n+\frac{1}{2}}=\\&K_{\beta}\mathcal{H}_y^{\gamma}\delta_x^{\beta}e_{i,j}^{n+\frac{1}{2}}+K_{\gamma}\mathcal{H}_x^{\beta}\delta_y^{\gamma}e_{i,j}^{n+\frac{1}{2}}+\\&\frac{3}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}(f(U_{i,j}^n)-f(u_{i,j}^n))-\\&\frac{1}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}(f(U_{i,j}^{n-1})-f(u_{i,j}^{n-1}))+R_{i,j}^{n+\frac{1}{2}},\\&(i,j)\in\omega,n\in[[0,N-1]]\end{aligned}\tag{18}$$

Theorem 1 Let $e_{i,j}^n$ be the solution of Eq. (18) and $\Delta t<1/(18(L+\frac{1}{2}))$, we have

$$\|e^n\|\leqslant c(\Delta t^2+\Delta x^4+\Delta y^4),n\in[[1,N]]$$

where c is a constant independent of Δx, Δy and Δt, and L is the Lipschitz constant of the function $f(u)$.

Proof: Forming the inner product of Eq. (18) with $e^{n+\frac{1}{2}}$ gives

$$\begin{aligned}&(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}\delta_te^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})+\frac{K_{\beta}K_{\gamma}\Delta t^2}{4}\cdot\\&(\delta_x^{\beta}\delta_y^{\gamma}\delta_te^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})=(K_{\beta}\mathcal{H}_y^{\gamma}\delta_x^{\beta}e^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})+\\&(K_{\gamma}\mathcal{H}_x^{\beta}\delta_y^{\gamma}e^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})+(\frac{3}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}(f(U^n)-\\&f(u^n)),e^{n+\frac{1}{2}})-(\frac{1}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}(f(U^{n-1})-\\&f(u^{n-1})),e^{n+\frac{1}{2}})+(R^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}}),\\&(i,j)\in\omega,n\in[[0,N-1]]\end{aligned}\tag{19}$$

Let $E^n=(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}e^n,\ e^n)=\|\mathcal{Q}_x\mathcal{Q}_ye^n\|^2,F^n=(\delta_x^{\beta}\delta_y^{\gamma}e^n,\ e^n)=\|\Lambda_x\Lambda_ye^n\|^2$, we have

$$(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}\delta_te^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})=\frac{1}{2\Delta t}(E^{n+1}-E^n)\tag{20a}$$

$$(\delta_x^{\beta}\delta_y^{\gamma}\delta_te^{n+\frac{1}{2}},\ e^{n+\frac{1}{2}})=\frac{1}{2\Delta t}(F^{n+1}-F^n)\tag{20b}$$

When $n\geqslant 1$, an application of Lemma 5, coupled with the Lipschitz condition and the Schwarz inequality, leads to

$$\begin{aligned}&\left(\frac{3}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}(f(U^n)-f(u^n)),e^{n+\frac{1}{2}}\right)\leqslant\\&\left(\frac{3}{2}\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}|f(U^n)-f(u^n)|,|e^{n+\frac{1}{2}}|\right)\leqslant\\&\frac{3L}{2}(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}|e^n|,|e^{n+\frac{1}{2}}|)\leqslant\frac{3L}{4}\Big(\Big(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}|e^n|,\\&|e^{n+1}|\Big)+\Big(\mathcal{H}_x^{\beta}\mathcal{H}_y^{\gamma}|e^n|,|e^n|\Big)\Big)\leqslant\end{aligned}$$

$$\frac{9L}{8}\|e^n\|^2+\frac{3L}{8}\|e^{n+1}\|^2 \tag{21}$$

By the same reasoning, we obtain

$$-\left(\frac{1}{2}\mathcal{H}_x^\beta\mathcal{H}_y^\gamma(f(U^{n-1})-f(u^{n-1})),e^{n+\frac{1}{2}}\right)\leqslant \frac{L}{2}(\mathcal{H}_x^\beta\mathcal{H}_y^\gamma|e^{n-1}|,|e^{n+\frac{1}{2}}|)\leqslant\frac{L}{4}\left(\left(\mathcal{H}_x^\beta\mathcal{H}_y^\gamma|e^{n-1}|,|e^{n+1}|\right)+\left(\mathcal{H}_x^\beta\mathcal{H}_y^\gamma|e^{n-1}|,|e^n|\right)\right)\leqslant \frac{L}{8}(2\|e^{n-1}\|^2+\|e^n\|^2+\|e^{n+1}\|^2) \tag{22}$$

Inserting Eq.(20) and inequalities (21)-(22) into Eq.(19), and using Lemma 6, yields

$$E^{n+1}+\frac{K_\beta K_\gamma\Delta t^2}{4}F^{n+1}\leqslant E^n+\frac{K_\beta K_\gamma\Delta t^2}{4}F^n+\frac{L\Delta t}{2}\|e^{n-1}\|^2+\frac{5L\Delta t}{2}\|e^n\|^2+L\Delta t\|e^{n+1}\|^2+2\Delta t(R^{n+\frac{1}{2}},e^{n+\frac{1}{2}})$$

Therefore for all $n\geqslant 2$,

$$\begin{aligned}E^n+\frac{K_\beta K_\gamma\Delta t^2}{4}F^n&\leqslant E^1+\frac{K_\beta K_\gamma\Delta t^2F^1}{4}+\frac{L\Delta t}{2}\sum_{m=1}^{n-1}\|e^{m-1}\|^2+\frac{5L\Delta t}{2}\sum_{m=1}^{n-1}\|e^m\|^2+L\Delta t\sum_{m=1}^{n-1}\|e^{m+1}\|^2+2\Delta t\sum_{m=1}^{n-1}(R^{m+\frac{1}{2}},e^{m+\frac{1}{2}})\leqslant\\&E^1+\frac{K_\beta K_\gamma\Delta t^2F^1}{4}+4L\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+L\Delta t\|e^n\|^2+\Delta t\sum_{m=1}^{n-1}(\|R^{m+\frac{1}{2}}\|^2+\|e^{m+\frac{1}{2}}\|^2)\leqslant\\&E^1+\frac{K_\beta K_\gamma\Delta t^2F^1}{4}+4L\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+L\Delta t\|e^n\|^2+\frac{\Delta t}{2}\sum_{m=1}^{n-1}(\|e^{m+1}\|^2+\|e^m\|^2)+\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2\leqslant E^1+\frac{K_\beta K_\gamma\Delta t^2F^1}{4}+\\&(4L+1)\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+(L+\frac{1}{2})\Delta t\|e^n\|^2+\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2\end{aligned} \tag{23}$$

Upon taking the inner product of Eq. (18) with $e^{\frac{1}{2}}$ for the case $n=0$, and noting the identity $f(U^0)-f(u^0)=0$, we establish that

$$(\mathcal{H}_x^\beta\mathcal{H}_y^\gamma\delta_te^{\frac{1}{2}},e^{\frac{1}{2}})+\left(\frac{K_\beta K_\gamma\Delta t^2}{4}\delta_x^\beta\delta_y^\gamma\delta_te^{\frac{1}{2}},e^{\frac{1}{2}}\right)=(K_\beta\mathcal{H}_y^\gamma\delta_x^\beta e^{\frac{1}{2}},e^{\frac{1}{2}})+(K_\gamma\mathcal{H}_x^\beta\delta_y^\gamma e^{\frac{1}{2}},e^{\frac{1}{2}})+(R^{\frac{1}{2}},e^{\frac{1}{2}}),(i,j)\in\omega \tag{24}$$

Using Eq.(20) and Lemma 6, we obtain

$$E^1+\frac{K_\beta K_\gamma\Delta t^2}{4}F^1\leqslant 2\Delta t(R^{\frac{1}{2}},e^1)\leqslant 2\Delta t\|R^{\frac{1}{2}}\|\|e^1\| \tag{25}$$

It follows from Lemma 5, the definition of E^1, and inequality (25) that

$$\|e^1\|\leqslant 18\Delta t\|R^{\frac{1}{2}}\| \tag{26}$$

Substituting inequality (26) into (25), we obtain

$$E^1+\frac{K_\beta K_\gamma\Delta t^2}{4}F^1\leqslant 36\Delta t^2\|R^{\frac{1}{2}}\|^2 \tag{27}$$

Inserting inequality (27) into (23) yields

$$E^n+\frac{K_\beta K_\gamma\Delta t^2}{4}F^n\leqslant 36\Delta t^2\|R^{\frac{1}{2}}\|^2+(4L+1)\cdot\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+(L+\frac{1}{2})\Delta t\|e^n\|^2+\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2$$

Using Lemma 5 and the definition of E^n, we get

$$(1-9(L+\frac{1}{2})\Delta t)\|e^n\|^2\leqslant 324\Delta t^2\|R^{\frac{1}{2}}\|^2+9(4L+1)\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+9\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2$$

When $\Delta t<1/(18(L+\frac{1}{2}))$,

$$\|e^n\|^2\leqslant 648\Delta t^2\|R^{\frac{1}{2}}\|^2+18(4L+1)\cdot\Delta t\sum_{m=1}^{n-1}\|e^m\|^2+18\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2$$

By Lemma 7, we get

$$\|e^n\|^2\leqslant\exp(18(4L+1)T)\left(648\Delta t^2\|R^{\frac{1}{2}}\|^2+18\Delta t\sum_{m=1}^{n-1}\|R^{m+\frac{1}{2}}\|^2\right)$$

From Eqs.(5) and (26), we obtain the required result immediately.

Remark 1 The stability of the linearized ADI scheme (7) can be readily established in a similar manner.

In what follows, we focus exclusively on the error estimates of the LOD scheme for linear equations. We consider the linear version of problem (1), which is formulated by discarding the nonlinear term. For this linear equation, the compact LOD scheme is given by

$$\begin{cases}\mathcal{H}_x^\beta\mathcal{H}_y^\gamma\delta_tu_{i,j}^{n+\frac{1}{2}}+\dfrac{K_\beta K_\gamma\Delta t^2}{4}\delta_x^\beta\delta_y^\gamma\delta_tu_{i,j}^{n+\frac{1}{2}}=\\\quad K_\beta\mathcal{H}_y^\gamma\delta_x^\beta u_{i,j}^{n+\frac{1}{2}}+K_\gamma\mathcal{H}_x^\beta\delta_y^\gamma u_{i,j}^{n+\frac{1}{2}}+\mathcal{H}_x^\beta\mathcal{H}_y^\gamma p_{i,j}^{n+\frac{1}{2}}+\\\quad\dfrac{K_\beta K_\gamma\Delta t^2}{4}\delta_x^\beta\delta_y^\gamma p_{i,j}^{n+\frac{1}{2}},(i,j)\in\omega,n\in[[0,N-1]]\\u_{i,j}^0=\varphi(x_i,y_j),(i,j)\in\omega\\u_{i,j}^n=0,(i,j)\in\partial\omega,n\in[[0,N-1]]\end{cases} \tag{28}$$

Let $e_{i,j}^n=U_{i,j}^n-u_{i,j}^n$, by subtracting the master equation in Eq.(28) from Eq. (8) and neglecting the nonlinear terms, we get

$$\mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t e_{i,j}^{n+\frac{1}{2}} + \frac{K_\beta K_\gamma \Delta t^2}{4}\delta_x^\beta \delta_y^\gamma \delta_t e_{i,j}^{n+\frac{1}{2}} = K_\beta \mathcal{H}_y^\gamma \delta_x^\beta e_{i,j}^{n+\frac{1}{2}} + K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma e_{i,j}^{n+\frac{1}{2}} + \bar{\bar{R}}_{i,j}^{n+\frac{1}{2}},\quad (i,j)\in\omega, n\in[[0,N-1]] \tag{29}$$

where $\bar{\bar{R}}_{i,j}^{n+\frac{1}{2}} = O(\Delta t^2 + \Delta x^4 + \Delta y^4), n \geqslant 0.$

Theorem 2 Let $e_{i,j}^n$ be the solution of Eq.(29) and $\Delta t < \frac{1}{9}$, we have

$$\| e^n \| \leqslant c(\Delta t^2 + \Delta x^4 + \Delta y^4),\quad n\in[[1,N]]$$

where c is a constant independent of Δx , Δy and Δt .

Proof: By forming the inner product of Eq. (29) with $e^{n+\frac{1}{2}}$, we get

$$(\mathcal{H}_x^\beta \mathcal{H}_y^\gamma \delta_t e^{n+\frac{1}{2}}, e^{n+\frac{1}{2}}) + \frac{K_\beta K_\gamma \Delta t^2}{4}\cdot (\delta_x^\beta \delta_y^\gamma \delta_t e^{n+\frac{1}{2}}, e^{n+\frac{1}{2}}) = (K_\beta \mathcal{H}_y^\gamma \delta_x^\beta e^{n+\frac{1}{2}}, e^{n+\frac{1}{2}}) + (K_\gamma \mathcal{H}_x^\beta \delta_y^\gamma e^{n+\frac{1}{2}}, e^{n+\frac{1}{2}}) + (\bar{\bar{R}}^{n+\frac{1}{2}}, e^{n+\frac{1}{2}}),\quad (i,j)\in\omega, n\in[[0,N-1]]$$

Let $E^n = (\mathcal{H}_x^\beta \mathcal{H}_y^\gamma e^n, e^n) = \|\mathcal{Q}_x \mathcal{Q}_y e^n\|^2$, $F^n = (\delta_x^\beta \delta_y^\gamma e^n, e^n) = \|\Lambda_x \Lambda_y e^n\|^2$. By Lemma 6, we obtain

$$E^{n+1} + \frac{K_\beta K_\gamma \Delta t^2}{4}F^{n+1} \leqslant E^n + \frac{K_\beta K_\gamma \Delta t^2}{4}F^n + 2\Delta t(\bar{\bar{R}}^{n+\frac{1}{2}}, e^{n+\frac{1}{2}})$$

Therefore for all $n \geqslant 1$,

$$E^n + \frac{K_\beta K_\gamma \Delta t^2}{4}F^n \leqslant E^0 + \frac{K_\beta K_\gamma \Delta t^2 F^0}{4} + 2\Delta t\sum_{m=0}^{n-1}(\bar{\bar{R}}^{m+\frac{1}{2}}, e^{m+\frac{1}{2}}) \leqslant E^0 + \frac{K_\beta K_\gamma \Delta t^2 F^0}{4} + \Delta t\sum_{m=0}^{n-1}(\|\bar{\bar{R}}^{m+\frac{1}{2}}\|^2 + \|e^{m+\frac{1}{2}}\|^2) \leqslant E^0 + \frac{K_\beta K_\gamma \Delta t^2 F^0}{4} + \Delta t\sum_{m=0}^{n-1}\|\bar{\bar{R}}^{m+\frac{1}{2}}\| + \frac{\Delta t}{2}\sum_{m=0}^{n-1}(\|e^{m+1}\|^2 + \|e^m\|^2)$$

By Lemma 5 and the definitions of E^n and F^n , we get

$$(1 - \frac{9}{2}\Delta t)\ \|e^n\|^2 \leqslant 9\Delta t\sum_{m=0}^{n-1}\|\bar{\bar{R}}^{m+\frac{1}{2}}\| + 9\Delta t\sum_{m=0}^{n-1}\|e^m\|^2$$

When $\Delta t < 1/9$,

$$\|e^n\|^2 \leqslant 18\Delta t\sum_{m=0}^{n-1}\|\bar{\bar{R}}^{m+\frac{1}{2}}\| + 18\Delta t\sum_{m=0}^{n-1}\|e^m\|^2$$

By Lemma 7, we get

$$\|e^n\|^2 \leqslant \exp(18T)\left(18\Delta t\sum_{m=0}^{n-1}\|\bar{\bar{R}}^{m+\frac{1}{2}}\|^2\right)$$

From the definition of $\bar{\bar{R}}_{i,j}^{n+\frac{1}{2}}$, the needed result follows immediately.

Remark 2 The stability of the LOD scheme (28) can be easily proved through an analogous approach.

2 PCG Method

The CG method is a well-known and efficient technique for solving linear systems (16) and (17). However, experiments indicate that the coefficient matrix has a widely distributed eigenvalue spectrum, which significantly degrades the convergence rate of the standard CG method. To address this issue, this section introduces a PCG approach that utilizes the τ-preconditioner to accelerate convergence for these systems. Next, we introduce a τ-preconditioner derived from the sine transform.

For any Toeplitz matrix

$$\boldsymbol{S}_M = \begin{bmatrix} s_0 & s_{-1} & \cdots & s_{2-M} & s_{1-M} \\ s_1 & s_0 & \cdots & s_{3-M} & s_{2-M} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ s_{M-2} & s_{M-3} & \cdots & s_0 & s_{-1} \\ s_{M-1} & s_{M-2} & \cdots & s_1 & s_0 \end{bmatrix}$$

the generating function of $\{\boldsymbol{S}_M\}_{M=1}^{\infty}$ can be defined as

$$p(\theta) = \sum_{k=-\infty}^{\infty} s_k \exp(ik\theta)$$

where $p(\theta)$ is in the Wiener class if and only if $\sum_{k=-\infty}^{\infty}|s_k| < \infty$[25]. In particular, if $s_k = s_{-k}$, then $\boldsymbol{S}_M$ is symmetric, and we denote it as $\tilde{\boldsymbol{S}}_M$. Let $\boldsymbol{H}_M$ be a Hankel matrix with antidiagonal entries

$$[s_2, s_3, \cdots, s_{M-1}, 0, 0, 0, s_{M-1}, \cdots, s_3, s_2]$$

Then, the τ -matrix $\tau(\tilde{\boldsymbol{S}}_M)$ can be described by $\tau(\tilde{\boldsymbol{S}}_M) = \tilde{\boldsymbol{S}}_M - \boldsymbol{H}_M$ [25]. Hence, we propose the τ-preconditioners of the linear systems (16) and (17) as

$$\tau(\boldsymbol{B}_\alpha^+) = \boldsymbol{B}_\alpha^+ - \boldsymbol{H}_M \tag{30}$$

The dominant computational overhead of PCG when applied to linear systems (16) and (17) arises from evaluating $\boldsymbol{B}_\alpha^- \boldsymbol{v}$ and $(\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{v}$. Leveraging the Toeplitz property of $\boldsymbol{B}_\alpha^-$ and the diagonalizability of $(\tau(\boldsymbol{B}_\alpha^+))^{-1}$, these matrix-vector multiplications achieve an optimal complexity of $O(M\log(M))$ (M_1 =

$M_2 = M$) [25].

Lemma 8 [35] If $\boldsymbol{T}_M$ is a ST matrix with the first column being $[t_0, t_1, \cdots, t_{M-1}]^{\mathrm{T}}$. Then, the eigenvalues of $\tau(\boldsymbol{T}_M)$ can be given by

$$\lambda_m = t_0 + 2\sum_{k=1}^{M-1}\left(t_k\cos\left(\frac{\pi mk}{M+1}\right)\right), m \in [[1, M]]$$

Lemma 9 $\tau(\boldsymbol{B}_\alpha^+)$ in Eq. (30) is SPD, and $\|(\tau(\boldsymbol{B}_\alpha^+))^{-1}\|_2 < 3$, $(M, \alpha) = (M_1, \beta)$ or (M_2, γ).

Proof: Since $\boldsymbol{B}_\alpha^+$ is an ST matrix and $\tau(\boldsymbol{B}_\alpha^+)$ is defined by Eq. (30), it follows that $\tau(\boldsymbol{B}_\alpha^+)$ is a symmetric matrix. We apply Lemma 8 and Lemma 1 while taking into account $\mu_\alpha < 0$ and $d_2^\alpha \in \left(\frac{1}{12}, \frac{1}{6}\right)$ to obtain

$$\lambda_m(\tau(\boldsymbol{B}_\alpha^+)) = 1 - 2d_2^\alpha + 2d_2^\alpha\cos\left(\frac{\pi m}{M+1}\right) + 2\mu_\alpha(\hat{w}_1^{(\alpha)} + (\hat{w}_0^{(\alpha)} + \hat{w}_2^{(\alpha)})\cos\left(\frac{\pi m}{M+1}\right) + \sum_{k=3}^{M}\hat{w}_k^{(\alpha)}\cos\left(\frac{\pi m(k-1)}{M+1}\right)) \geqslant 1 - 2d_2^\alpha + 2d_2^\alpha\cos\left(\frac{\pi m}{M+1}\right) + 2\mu_\alpha(\hat{w}_1^{(\alpha)} + (\hat{w}_0^{(\alpha)} + \hat{w}_2^{(\alpha)}) + \sum_{k=3}^{M}\hat{w}_k^{(\alpha)}) = 1 - 2d_2^\alpha + 2d_2^\alpha\cos\left(\frac{\pi m}{M+1}\right) + 2\mu_\alpha\sum_{k=0}^{M}\hat{w}_k^{(\alpha)} > 1 - 4d_2^\alpha > \frac{1}{3}$$

Therefore, the matrix $\tau(\boldsymbol{B}_\alpha^+)$ is SPD, and we can further obtain that

$$\|(\tau(\boldsymbol{B}_\alpha^+))^{-1}\|_2 = \max_{1\leqslant m\leqslant M}|\lambda_m((\tau(\boldsymbol{B}_\alpha^+))^{-1})| = \frac{1}{\min\limits_{1\leqslant m\leqslant M}\lambda_m(\tau(\boldsymbol{B}_\alpha^+))} < 3$$

This completes the proof.

Lemma 10 Let $\boldsymbol{S}_+^{(\alpha)}$ be defined by Eq. (15). Under the assumption (12), the generating functions of $\{\boldsymbol{S}_+^{(\alpha)}\}_{M=1}^{\infty}$ ($(M, \alpha) = (M_1, \beta)$ or (M_2, γ)) are in the Wiener class.

Proof: Applying Lemma 1, and noting that $d_2^\alpha \in \left(\frac{1}{12}, \frac{1}{6}\right)$, we deduce that

$$\sum_{k=-\infty}^{\infty}|s_k^{(\alpha,+)}| = |2\mu_\alpha\hat{w}_1^{(\alpha)} - 2d_2^\alpha| + 2|\mu_\alpha(\hat{w}_0^{(\alpha)} + \hat{w}_2^{(\alpha)}) + d_2^\alpha| + 2\sum_{k=3}^{\infty}|\mu_\alpha\hat{w}_k^{(\alpha)}| \leqslant |2\mu_\alpha\hat{w}_1^{(\alpha)}| + |2d_2^\alpha| + 2|\mu_\alpha(\hat{w}_0^{(\alpha)} + \hat{w}_2^{(\alpha)})| + 2|d_2^\alpha| + 2|\mu_\alpha|\sum_{k=3}^{\infty}|\hat{w}_k^{(\alpha)}| = 4d_2^\alpha + 4\mu_\alpha\hat{w}_1^{(\alpha)} < \infty$$

This completes the proof.

Lemma 11 [35] If the generating function of $\{\boldsymbol{S}_+^{(\alpha)}\}_{M=1}^{\infty}$ is in the Wiener class, then for any $\varepsilon > 0$, there exists $M_\alpha > 0$ such that for all $M > M_\alpha$,

$$\boldsymbol{S}_+^{(\alpha)} - \tau(\boldsymbol{S}_+^{(\alpha)}) = \boldsymbol{U}_M' + \boldsymbol{V}_M'$$

where $\|\boldsymbol{U}_M'\|_2 < \frac{\varepsilon}{3}$ and $\mathrm{rank}(\boldsymbol{V}_M') \leqslant M_\alpha$.

Theorem 3 Let $\boldsymbol{B}_\alpha^+$ and $\tau(\boldsymbol{B}_\alpha^+)$ be defined by Eqs. (14) and (30), respectively. Under the assumption (12), and then for any $\varepsilon > 0$, there exists $M_\alpha > 0$ such that for all $M > M_\alpha$,

$$(\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{B}_\alpha^+ - \boldsymbol{I} = \boldsymbol{U}_M + \boldsymbol{V}_M$$

where $\|\boldsymbol{U}_M\|_2 < \varepsilon$ and $\mathrm{rank}(\boldsymbol{V}_M) \leqslant M_\alpha$, $(M, \alpha) = (M_1, \beta)$ or (M_2, γ).

Proof: From the definition of $\boldsymbol{B}_\alpha^+$ and Lemmas 10-11, we get

$$\boldsymbol{B}_\alpha^+ - \tau(\boldsymbol{B}_\alpha^+) = \boldsymbol{I} + \boldsymbol{S}_+^{(\alpha)} - \tau(\boldsymbol{I} + \boldsymbol{S}_+^{(\alpha)}) = \boldsymbol{S}_+^{(\alpha)} - \tau(\boldsymbol{S}_+^{(\alpha)}) = \boldsymbol{U}_M' + \boldsymbol{V}_M' \tag{31}$$

where

$$\|\boldsymbol{U}_M'\|_2 < \frac{\varepsilon}{3}, \ \mathrm{rank}(\boldsymbol{V}_M') \leqslant M_\alpha \tag{32}$$

Multiplying both sides of Eq. (31) by $(\tau(\boldsymbol{B}_\alpha^+))^{-1}$, we have

$$(\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{B}_\alpha^+ - \boldsymbol{I} = (\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{U}_M' + (\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{V}_M' = \boldsymbol{U}_M + \boldsymbol{V}_M$$

where $\boldsymbol{U}_M = (\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{U}_M'$ and $\boldsymbol{V}_M = (\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{V}_M'$. Applying Lemma 9 and inequality (32), we have

$$\|\boldsymbol{U}_M\|_2 \leqslant \|(\tau(\boldsymbol{B}_\alpha^+))^{-1}\|_2\|\boldsymbol{U}_M'\|_2 < \varepsilon$$

$$\mathrm{rank}(\boldsymbol{V}_M) = \mathrm{rank}(\boldsymbol{V}_M') \leqslant M_\alpha$$

This completes the proof.

According to the Weyl's Theorem, Theorem 3 indicates the spectrum of $(\tau(\boldsymbol{B}_\alpha^+))^{-1}\boldsymbol{B}_\alpha^+$ is clustered around 1 [36]. This typically leads to a fast convergence rate for the PCG method in solving linear systems (16) and (17) [24-25, 36].

3 Numerical Results

A series of numerical tests are performed to evaluate the method's effectiveness. We set the stopping criterion to be $\|r^{(k)}\|_2 / \|r^{(0)}\|_2 < 10^{-7}$. As shown in Tables 1-8, the acronyms "CG-LOD" and "CG-ADI" denote the base CG method for the linearized LOD and ADI schemes, respectively.

"PCG-LOD (τ)" and "PCG-ADI (τ)" represent our proposed PCG method with the τ-preconditioner. "PCG-LOD (S/T)" and "PCG-ADI (S/T)" indicate the PCG methods with Strang's[37] and T.Chan's[20] circulant preconditioners, respectively. "Time(s)" denotes the total CPU time in seconds. "Iter" represents the average number of iterations across all time steps for solving linear systems in both directions. Let $\Delta x = \Delta y = h$. Define

$$\mathrm{er} = \sqrt{\Delta x \Delta y \sum_{i=1}^{M_1} \sum_{j=1}^{M_2} (U_{i,j} - u_{i,j})^2}$$

$$\mathrm{or}_{\Delta t} = \log_2 \left[\frac{\mathrm{er}(h,\Delta t)}{\mathrm{er}(h,\Delta t/2)} \right]$$

$$\mathrm{or}_h = \log_2 \left[\frac{\mathrm{er}(h,\Delta t)}{\mathrm{er}(h/2,\Delta t)} \right]$$

Table 1 Temporal errors and convergence rates of the LOD scheme (11) with $h = 1/200$

β	Δt	$\gamma = 1.2$		$\gamma = 1.5$		$\gamma = 1.9$	
		er	$\mathrm{or}_{\Delta t}$	er	$\mathrm{or}_{\Delta t}$	er	$\mathrm{or}_{\Delta t}$
1.2	1/8	3.02065×10^{-7}	–	5.36413×10^{-7}	–	1.21187×10^{-6}	–
	1/16	7.31325×10^{-8}	2.046	1.28297×10^{-7}	2.064	2.85930×10^{-7}	2.084
	1/32	1.81416×10^{-8}	2.011	3.17351×10^{-8}	2.015	7.05046×10^{-8}	2.020
	1/64	4.52672×10^{-9}	2.003	7.91297×10^{-9}	2.004	1.75663×10^{-8}	2.005

Table 2 Spatial errors and convergence rates of the LOD scheme (11) with $\Delta t = 1/10000$

β	h	$\gamma = 1.2$		$\gamma = 1.5$		$\gamma = 1.9$	
		er	or_h	er	or_h	er	or_h
1.2	1/8	9.55062×10^{-9}	–	1.29432×10^{-8}	–	1.68654×10^{-8}	–
	1/16	7.30492×10^{-10}	3.709	9.66241×10^{-10}	3.744	1.18835×10^{-9}	3.827
	1/32	5.10642×10^{-11}	3.838	6.58259×10^{-11}	3.876	7.81579×10^{-11}	3.926
	1/64	3.47125×10^{-12}	3.879	4.40680×10^{-12}	3.901	5.24298×10^{-12}	3.898

Table 3 Temporal errors and convergence rates of the ADI scheme (7) with $h = 1/200$

β	Δt	$\gamma = 1.2$		$\gamma = 1.5$		$\gamma = 1.9$	
		er	$\mathrm{or}_{\Delta t}$	er	$\mathrm{or}_{\Delta t}$	er	$\mathrm{or}_{\Delta t}$
1.2	1/8	1.27311×10^{-8}	–	1.63531×10^{-8}	–	2.12941×10^{-8}	–
	1/16	3.15482×10^{-9}	2.013	4.03711×10^{-9}	2.018	5.23545×10^{-9}	2.024
	1/32	7.87512×10^{-10}	2.002	1.00670×10^{-9}	2.004	1.30391×10^{-9}	2.005
	1/64	1.96882×10^{-10}	2.000	2.51596×10^{-10}	2.000	3.25719×10^{-10}	2.001

Table 4 Spatial errors and convergence rates of the ADI scheme (7) with $\Delta t = 1/10000$

β	h	$\gamma = 1.2$		$\gamma = 1.5$		$\gamma = 1.9$	
		er	or_h	er	or_h	er	or_h
1.2	1/8	9.55053×10^{-9}	–	1.2943×10^{-8}	–	1.68651×10^{-8}	–
	1/16	7.30391×10^{-10}	3.709	9.66079×10^{-10}	3.744	1.18809×10^{-9}	3.827
	1/32	5.09608×10^{-11}	3.841	6.56595×10^{-11}	3.879	7.78913×10^{-11}	3.931
	1/64	3.36353×10^{-12}	3.921	4.22841×10^{-12}	3.957	4.93415×10^{-12}	3.981

Example 1 Consider the problem (1) with $K_\beta = K_\gamma = 1$, $L_1 = L_2 = T = 1$, $f(u) = \sin(u)$,

$$p(x,y,t) = \exp(-t)\Big[\frac{1}{2\cos(\beta\pi/2)} y^4(1-y)^4 \cdot \Phi(x,\beta) + \frac{1}{2\cos(\gamma\pi/2)} x^4(1-x)^4 \cdot \Phi(y,\gamma)\Big] - \exp(-t)\,(x(1-x)y(1-y))^4 - \sin(u)$$

and

$$\Phi(z,\alpha) = \frac{\Gamma(5)}{\Gamma(5-\alpha)}[z^{4-\alpha} + (1-z)^{4-\alpha}] - \frac{4\Gamma(6)}{\Gamma(6-\alpha)}[z^{5-\alpha} + (1-z)^{5-\alpha}] + \frac{6\Gamma(7)}{\Gamma(7-\alpha)}[z^{6-\alpha} + (1-z)^{6-\alpha}] - \frac{4\Gamma(8)}{\Gamma(8-\alpha)}[z^{7-\alpha} + (1-z)^{7-\alpha}] +$$

$$\frac{\Gamma(9)}{\Gamma(9-\alpha)}[z^{8-\alpha}+(1-z)^{8-\alpha}],$$

$(z,\alpha)=(x,\beta)$ or (y,γ)

The exact solution is

$$u(x,y,t)=\exp(-t)\,(x(1-x)y(1-y))^4$$

Tables 1−4 describe the numerical outcomes for solving Example 1, showing the errors and convergence rates for the linearized ADI scheme (7) and LOD scheme (11). It can be seen that both schemes achieve second-order temporal convergence and fourth-order spatial convergence.

Table 5 Comparisons of PCG-LOD(τ) method and the CG-LOD method with $\Delta t = 2h$

β	γ	τ	CG-LOD			PCG-LOD(τ)		
			Iter	Time(s)	er	Iter	Time(s)	er
1.2	1.2	1/200	10.65	10.6	1.85346×10^{-9}	4.00	11.0	1.85345×10^{-9}
		1/600	11.11	280.9	2.06294×10^{-10}	4.00	261.0	2.06261×10^{-10}
		1/1000	11.11	1517.2	7.43265×10^{-11}	4.00	1126.2	7.42801×10^{-11}
		1/1400	11.04	4019.4	3.79477×10^{-11}	4.00	3215.5	3.78936×10^{-11}
1.5	1.5	1/200	28.85	21.2	5.73690×10^{-9}	4.62	11.9	5.73689×10^{-9}
		1/600	36.54	673.2	6.38679×10^{-10}	5.00	287.8	6.38679×10^{-10}
		1/1000	39.89	3443.3	2.30143×10^{-10}	5.00	1257.6	2.30142×10^{-10}
		1/1400	42.39	11475.2	1.17490×10^{-10}	5.00	3643.5	1.17489×10^{-10}
1.9	1.9	1/200	89.74	63.4	2.92085×10^{-8}	4.00	11.0	2.92085×10^{-8}
		1/600	193.35	3118.6	3.24210×10^{-9}	4.00	258.5	3.24210×10^{-9}
		1/1000	246.25	18705.9	1.16721×10^{-9}	4.00	1145.0	1.16721×10^{-9}
		1/1400	–	–	–	4.00	3286.2	5.95558×10^{-10}

Table 6 Comparisons of PCG-ADI(τ) method and the CG-ADI method with $\Delta t = 2h$

β	γ	τ	CG-ADI			PCG-ADI(τ)		
			Iter	Time(s)	er	Iter	Time(s)	er
1.2	1.2	1/200	11.00	10.7	8.06613×10^{-11}	4.00	10.5	8.06955×10^{-11}
		1/600	11.01	280.8	8.97520×10^{-12}	4.00	252.9	9.06143×10^{-12}
		1/1000	11.01	1328.1	3.29484×10^{-12}	4.00	1133.8	3.38233×10^{-12}
		1/1400	11.01	3916.7	1.78150×10^{-12}	4.00	3259.6	1.86766×10^{-12}
1.5	1.5	1/200	28.00	20.7	1.37864×10^{-10}	4.03	11.0	1.37916×10^{-10}
		1/600	35.03	659.9	1.53155×10^{-11}	5.00	293.4	1.53158×10^{-11}
		1/1000	39.02	3438.3	5.51279×10^{-12}	5.00	1282.4	5.51335×10^{-12}
		1/1400	41.03	10811.9	2.81728×10^{-12}	5.00	3666.4	2.81250×10^{-12}
1.9	1.9	1/200	83.97	54.8	2.78381×10^{-10}	4.00	11.0	2.78381×10^{-10}
		1/600	184.69	3022.5	3.09189×10^{-11}	4.00	259.2	3.09205×10^{-11}
		1/1000	236.40	18191.1	1.11304×10^{-11}	4.00	1140.1	1.11328×10^{-11}
		1/1400	–	–	–	4.00	3263.8	5.68160×10^{-12}

In Tables 5−6, we compare the performance of our PCG-LOD(τ) and PCG-ADI(τ) methods with those of the CG-LOD and CG-ADI methods. It is evident that PCG-LOD (τ) and PCG-ADI (τ) achieve both a lower iteration count and reduced computational time. For instance, when $\beta=\gamma=1.9$, $h=1/1000$, the CG-LOD method achieves an accuracy of 1.16721×10^{-9} at the cost of an average of

246.25 iterations and 18705.9 s. In contrast, PCG-LOD(τ) achieves the same precision with only 4.00 iterations on average and a computation time of 1145.0 s. The results demonstrate that the proposed methods markedly enhance computational efficiency.

Existing circulant preconditioners have been developed for solving the linear systems (16) and (17) with the ST coefficient matrices[20,37]. As indicated in Tables 7-8, our PCG-LOD(τ) and PCG-ADI(τ) methods outperform the existing methods in terms of average iteration count, including PCG-LOD(S/T) and PCG-ADI(S/T).

The spectra of several preconditioned matrices are presented in Figs. 1 and 2. As clearly shown, matrices $(\tau(\boldsymbol{B}_\beta^+))^{-1}\boldsymbol{B}_\beta^+$ and $(\tau(\boldsymbol{B}_\gamma^+))^{-1}\boldsymbol{B}_\gamma^+$ exhibit significantly greater spectral concentration compared with matrices $(S(\boldsymbol{B}_\beta^+))^{-1}\boldsymbol{B}_\beta^+$, $(T(\boldsymbol{B}_\beta^+))^{-1}\boldsymbol{B}_\beta^+$, $(S(\boldsymbol{B}_\gamma^+))^{-1}\boldsymbol{B}_\gamma^+$ and $(T(\boldsymbol{B}_\gamma^+))^{-1}\boldsymbol{B}_\gamma^+$. The average iteration numbers for various preconditioning methods, presented in Tables 7-8, lend further support to this conclusion.

Table 7 Performance of different PCG methods with $\Delta t = 2h$

β	γ	h	PCG-LOD(τ)		PCG-LOD(S)		PCG-LOD(T)	
			Iter	er	Iter	er	Iter	er
1.2	1.2	1/200	4.00	1.85345×10^{-9}	5.00	1.85346×10^{-9}	4.05	1.85346×10^{-9}
		1/600	4.00	2.06261×10^{-10}	4.02	2.06299×10^{-10}	4.02	2.06287×10^{-10}
		1/1000	4.00	7.42801×10^{-11}	4.02	7.43102×10^{-11}	4.01	7.43173×10^{-11}
		1/1400	4.00	3.78936×10^{-11}	4.01	3.79176×10^{-11}	4.01	3.79352×10^{-11}
1.5	1.5	1/200	4.62	5.73689×10^{-9}	4.10	5.73692×10^{-9}	5.06	5.73690×10^{-9}
		1/600	5.00	6.38679×10^{-10}	5.00	6.38679×10^{-10}	5.03	6.38678×10^{-10}
		1/1000	5.00	2.30142×10^{-10}	5.00	2.30142×10^{-10}	4.61	2.30128×10^{-10}
		1/1400	5.00	1.17489×10^{-10}	5.00	1.17489×10^{-10}	4.60	1.17485×10^{-10}
1.9	1.9	1/200	4.00	2.92085×10^{-8}	4.00	2.92085×10^{-8}	8.41	2.92085×10^{-8}
		1/600	4.00	3.24210×10^{-9}	4.07	3.24209×10^{-9}	8.78	3.24210×10^{-9}
		1/1000	4.00	1.16721×10^{-9}	4.60	1.16721×10^{-9}	8.16	1.16721×10^{-9}
		1/1400	4.00	5.95558×10^{-10}	5.00	5.95559×10^{-10}	8.09	5.95559×10^{-10}

Table 8 Performance of different PCG methods with $\Delta t = 2h$

β	γ	h	PCG-ADI(τ)		PCG-ADI(S)		PCG-ADI(T)	
			Iter	er	Iter	er	Iter	er
1.2	1.2	1/200	4.00	8.06955×10^{-11}	5.00	8.06618×10^{-11}	4.01	8.06517×10^{-11}
		1/600	4.00	9.06143×10^{-12}	4.01	8.95070×10^{-12}	4.01	8.99115×10^{-12}
		1/1000	4.00	3.38233×10^{-12}	4.01	3.25564×10^{-12}	4.00	3.23887×10^{-12}
		1/1400	4.00	1.86766×10^{-12}	4.00	1.71162×10^{-12}	4.00	1.64816×10^{-12}
1.5	1.5	1/200	4.03	1.37916×10^{-10}	4.03	1.37826×10^{-10}	5.02	1.37865×10^{-10}
		1/600	5.00	1.53158×10^{-11}	5.00	1.53155×10^{-11}	4.98	1.53152×10^{-11}
		1/1000	5.00	5.51335×10^{-12}	5.00	5.51346×10^{-12}	4.04	5.64731×10^{-12}
		1/1400	5.00	2.81250×10^{-12}	5.00	2.81314×10^{-12}	4.03	2.85920×10^{-12}
1.9	1.9	1/200	4.00	2.78381×10^{-10}	4.00	2.78379×10^{-10}	8.02	2.78381×10^{-10}
		1/600	4.00	3.09205×10^{-11}	4.01	3.09311×10^{-11}	8.02	3.09189×10^{-11}
		1/1000	4.00	1.11328×10^{-11}	4.02	1.11556×10^{-11}	8.01	1.11302×10^{-11}
		1/1400	4.00	5.68160×10^{-12}	5.00	5.67876×10^{-12}	8.01	5.67836×10^{-12}

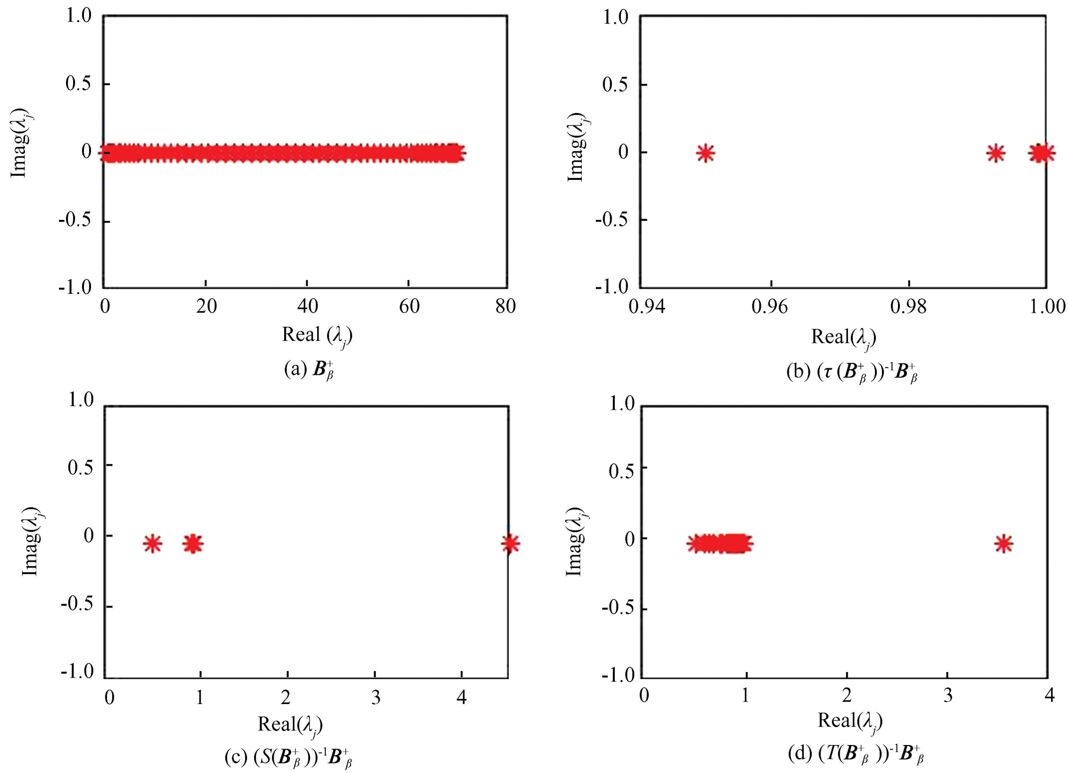

Fig.1 The spectrum of B_β^+, $(\tau(B_\beta^+))^{-1}B_\beta^+$, $(S(B_\beta^+))^{-1}B_\beta^+$ and $(T(B_\beta^+))^{-1}B_\beta^+$ with $h = \Delta t = 1/2^6$ and $\beta = \gamma = 1.9$ for Example 1

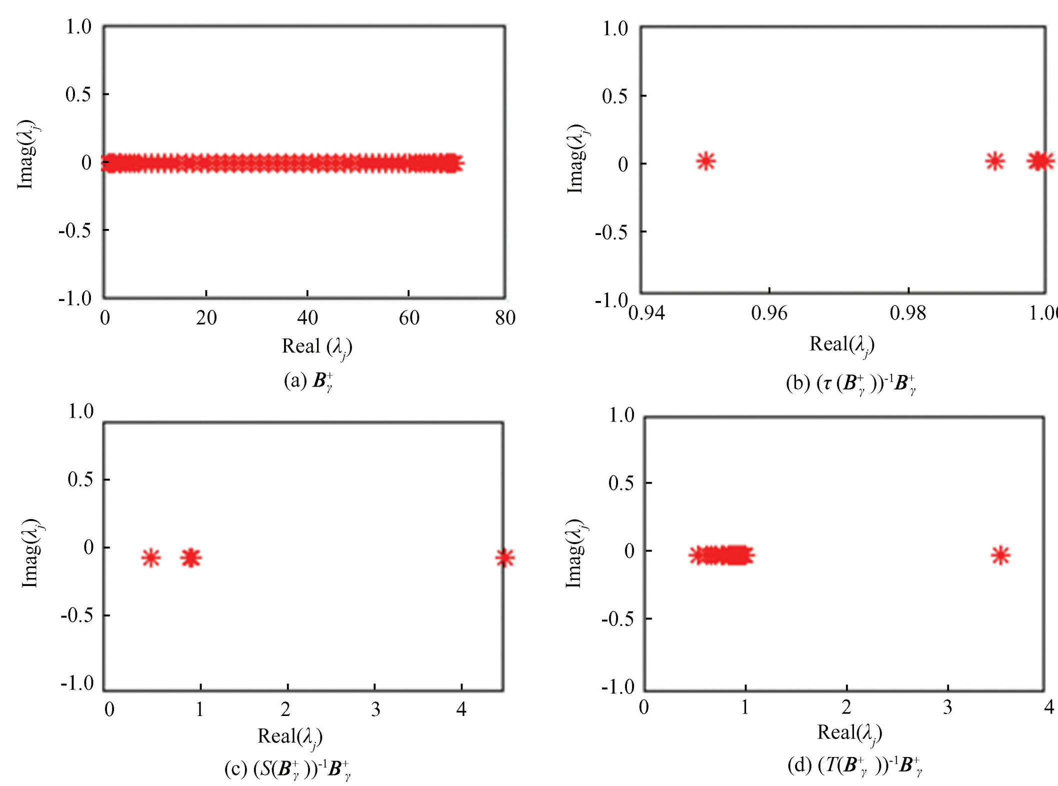

Fig. 2 The spectrum of B_γ^+, $(\tau(B_\gamma^+))^{-1}B_\gamma^+$, $(S(B_\gamma^+))^{-1}B_\gamma^+$ and $(T(B_\gamma^+))^{-1}B_\gamma^+$ with $h = \Delta t = 1/2^6$ and $\beta = \gamma = 1.9$ for Example 1

4 Conclusions

Fourth-order linearized ADI and LOD methods are derived herein to solve 2D nonlinear RSFDEs numerically. Discrete energy analysis provides explicit error estimates that establish the second-order temporal and fourth-order spatial accuracy of both schemes. Moreover, the PCG method, which incorporates a τ-preconditioner, is employed to solve these discrete linear systems. It has been theoretically proved that the preconditioned matrix can be decomposed into the sum of an identity matrix, a small-norm matrix, and a low-rank matrix. As demonstrated by numerical tests, both methods with τ-preconditioner outperform existing methods in both iteration efficiency and computational cost. Furthermore, preliminary evidence suggests that the proposed τ-preconditioner may also be applicable to the ADI scheme developed by Hu and Cao [33], which constitutes a pertinent subject for our further investigation.

References

[1] Metzler R, Nonnenmacher T F. Fractional relaxation processes and fractional rheological models for the description of a class of viscoelastic materials. International Journal of Plasticity, 2003, 19(7): 941-959. DOI: 10.1016/S0749-6419(02)00087-6.

[2] Benson D A, Wheatcraft S W, Meerschaert M M. Application of a fractional advection-dispersion equation. Water Resources Research, 2000, 36(6): 1403-1412. DOI: 10.1029/2000WR900031.

[3] Raberto M, Scalas E, Mainardi F. Waiting-times and returns in high-frequency financial data: An empirical study. Physica A: Statistical Mechanics & Its Applications, 2002, 314(1-4): 749-755. DOI: 10.1016/S0378-4371(02)01048-8.

[4] Sabatelli L, Keating S, Dudley J, et al. Waiting time distributions in financial markets. The European Physical Journal B-Condensed Matter and Complex Systems, 2002, 27: 273-275. DOI: 10.1140/epjb/e20020151.

[5] Meerschaert M M, Tadjeran C. Finite difference approximation for fractional advection-dispersion flow equations. Journal of Computational and Applied Mathematics, 2004, 172(1): 65-77. DOI: 10.1016/j.cam.2004.01.033.

[6] Langlands T A M, Henry B I. The accuracy and stability of an implicit solution method for the fractional diffusion equation. Journal of Computational Physics, 2005, 205(2): 719-736. DOI: 10.1016/j.jcp.2004.11.025.

[7] Meerschaert M M, Scheffler H P, Tadjeran C. Finite difference methods for two-dimensional fractional dispersion equation. Journal of Computational Physics, 2006, 211(1): 249-261. DOI: 10.1016/j.jcp.2005.05.017.

[8] Liu F, Anh V, Turner I. Numerical solution of the space fractional Fokker-Planck equation. Journal of Computational and Applied Mathematics, 2004, 166(1): 209-219. DOI: 10.1016/j.cam.2003.09.028.

[9] Ervin V J, Heuer N, Roop J P. Numerical approximation of a time dependent nonlinear space-fractional diffusion equation. SIAM Journal on Numerical Analysis, 2007, 45(2): 572-591. DOI: 10.1137/050642757.

[10] Deng W. Finite element method for the space and time fractional Fokker-Planck equation. SIAM Journal on Numerical Analysis, 2008, 47(1): 204-226. DOI: 10.1137/080714130.

[11] Fu H, Sun Y, Wang H, et al. Stability and convergence of a Crank-Nicolson finite volume method for space fractional diffusion equations. Applied Numerical Mathematics, 2019, 139: 38-51. DOI: 10.1016/j.apnum.2019.01.005.

[12] Hejazi H, Moroney T, Liu F. Stability and convergence of a finite volume method for the space fractional advection-dispersion equation. Journal of Computational & Applied Mathematics, 2014, 255: 684-697. DOI: 10.1016/j.cam.2013.06.039.

[13] Simmons A, Yang Q, Moroney T. A finite volume method for two-sided fractional diffusion equations on non-uniform meshes. Journal of Computational Physics, 2017, 335: 747-759. DOI: 10.1016/j.jcp.2017.01.061.

[14] Ortigueira M D. Riesz potential operators and inverses via fractional centred derivatives. International Journal of Mathematics and Mathematical Sciences, 2006, 2006(1): 048391. DOI: 10.1155/IJMMS/2006/48391.

[15] Çelik C, Duman M. Crank-Nicolson method for the fractional diffusion equation with the Riesz fractional derivative. Journal of Computational Physics, 2012, 231(4): 1743-1750. DOI: 10.1016/j.jcp.2011.11.008.

[16] Tian W, Zhou H, Deng W. A class of second order difference approximations for solving space fractional diffusion equations. Mathematics of Computation, 2015, 84(294): 1703-1727. DOI: 10.1090/s0025-5718-2015-02917-2.

[17] Zhou H, Tian W Y, Deng W H. Quasi-compact finite difference schemes for space fractional diffusion equations. Journal of Scientific Computing, 2013, 56: 45-66. DOI: 10.1007/s10915-012-9661-0.

[18] Hao Z P, Sun Z Z, Cao W R. A fourth-order approximation of fractional derivatives with its applications. Journal of Computational Physics, 2015, 281: 787-805. DOI: 10.1016/j.jcp.2014.10.053.

[19] Yu J W, Zhang C H, Huang X, et al. A class of preconditioner for solving the Riesz distributed-order nonlinear space-fractional diffusion equations. Japan Journal of Industrial and Applied Mathematics, 2023, 40: 537-562. DOI: 10.1007/s13160-022-00551-w.

[20] Chan T F. An optimal circulant preconditioner for Toeplitz

systems. SIAM Journal on Scientific and Statistical Computing, 1988, 9(4): 766-771. DOI: 10.1137/0909051.

[21] Chan R H. Circulant preconditioners for Hermitian Toeplitz systems. SIAM Journal on Matrix Analysis and Applications, 1989, 10: 876-896. DOI: 10.1137/0610039.

[22] Tang S P, Yang A L, Zhou J L, et al. R. Chan's circulant-based approximate inverse preconditioning iterative method for solving second-order space fractional advection-dispersion equations with variable coefficients. Computational and Applied Mathematics, 2024, 43(2): article number 97. DOI: 10.1007/s40314-024-02592-y.

[23] Zhang C H, Yu J W, Wang X. A fast second-order scheme for nonlinear Riesz space fractional diffusion equations. Numerical Algorithms, 2023, 92: 1813-1836. DOI: 10.1007/s11075-022-01367-y.

[24] Lin X L, Huang X, Ng M K, et al. A τ-preconditioner for a non-symmetric linear system arising from multidimensional Riemann-Liouville fractional diffusion equation. Numerical Algorithms, 2023, 92: 795-813. DOI: 10.1007/s11075-022-01342-7.

[25] Huang Y Y, Qu W, Lel S L. On τ-preconditioner for a novel fourth-order difference scheme of two-dimensional Riesz space-fractional diffusion equations. Computers & Mathematics with Applications, 2023, 145: 124-140. DOI: 10.1016/j.camwa.2023.06.015.

[26] Tadjeran C, Meerschaert M M. A second-order accurate numerical method for the two-dimensional fractional diffusion equation. Journal of Computational Physics, 2007, 220(2): 813-823. DOI: 10.1016/j.jcp.2006.05.030.

[27] Wang H, Wang K. An O ($N \log^2 N$) alternating-direction finite difference method for two-dimensional fractional diffusion equations. Journal of Computational Physics, 2011, 230(21): 7830-7839. DOI: 10.1016/j.jcp.2011.07.003.

[28] Wang H, Du N. Fast alternating-direction finite difference methods for three-dimensional space-fractional diffusion equations. Journal of Computational Physics, 2014, 258: 305-318. DOI: 10.1016/j.jcp.2013.10.040.

[29] Qin J, Wang T. A compact locally one-dimensional finite difference method for nonhomogeneous parabolic differential equations. International Journal for Numerical Methods in Biomedical Engineering, 2011, 27(1): 128-142. DOI: 10.1002/cnm.1299.

[30] Wang T, Wang Y M. A higher-order compact LOD method and its extrapolations for nonhomogeneous parabolic differential equations. Applied Mathematics & Computation, 2014, 237: 512-530. DOI: 10.1016/j.amc.2014.03.139.

[31] Tang S P, Huang Y M. A fast ADI based matrix splitting preconditioning method for the high dimensional space fractional diffusion equations in conservative form. Computers and Mathematics with Applications, 2023, 144: 210-220. DOI: 10.1016/j.camwa.2023.05.028.

[32] Chen M, Wang Y, Cheng X, et al. Second-order LOD multigrid method for multidimensional Riesz fractional diffusion equation. Bit Numer Math, 2014, 54: 623-647. DOI: 10.1007/s10543-014-0477-1.

[33] Hu D, Cao X. A fourth-order compact ADI scheme for two-dimensional Riesz space fractional nonlinear reaction-diffusion equation. International Journal of Computer Mathematics, 2014, 97(9): 1928-1948. DOI: 10.1080/00207160.2019.1671587.

[34] Sun Z Z. The Method of Order Reduction and Its Application to the Numerical Solutions of Partial Differential Equations. Beijing: Science Press, 2009.

[35] Bini D, Benedetto F. A new preconditioner for the parallel solution of positive definite Toeplitz systems. Proceedings of the Second Annual ACM Symposium on Parallel Algorithms and Architectures, New York: ACM, 1990: 220-223. DOI: 10.1145/97444.97688.

[36] Jin X. Preconditioning Techniques for Toeplitz Systems. Beijing: Higher Education Press, 2010.

[37] Strang G. A proposal for Toeplitz matrix calculations. Studies in Applied Mathematics, 1986, 74(2): 171-176.