Fast implicit difference schemes for time-space fractional diffusion equations with the integral fractional Laplacian

Xian-Ming Gu1,2,3 | Hai-Wei Sun2 | Yanzhi Zhang4 | Yong-Liang Zhao5

1School of Economic Mathematics/Institute of Mathematics, Southwestern University of Finance and Economics, Chengdu, 611130, Sichuan, China
2Department of Mathematics, University of Macau, Avenida da Universidade, Taipa, Macao, China
3Bernoulli Institute for Mathematics, Computer Science and Artificial Intelligence, University of Groningen, Nijenborgh 9, P.O. Box 407, Groningen, 9700 AK, The Netherlands
4Department of Mathematics and Statistics, Missouri University of Science and Technology, Rolla, 65409-0020, MO, USA
5School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, 611731, Sichuan, China

Correspondence
Yong-Liang Zhao, School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, Sichuan 611731, China.
Email: ylzhaofde@sina.com

Funding information
Applied Basic Research Project of Sichuan Province, Grant/Award Number: 2020YJ0007; Fundamental Research Funds for the Central Universities, Grant/Award Number: JBK1902028; National Natural Science Foundation of China, Grant/Award Number: 11801463; National Science Foundation, Grant/Award Number: DMS-1620465; Science and Technology Development Fund, Macau SAR, Grant/Award Number: 0118/2018/A3

1 | INTRODUCTION

In recent decades, fractional partial differential equations (FPDEs) have attracted growing attention in modeling phenomena with long-term memory and spatial heterogeneity arising in engineering, physics, chemistry, and other applied
In physics, fractional derivatives are used to model anomalous diffusion. Anomalous diffusion is the theory of
diffusing particles in environments that are not locally homogeneous. A physical–mathematical model to anomalous
diffusion may be based on FPDEs containing derivatives of fractional order in both space and time, where the subdif-
fusion appears in time and the superdiffusion occurs in space simultaneously. On the other hand, although most of
time–space fractional diffusion models are initially defined with the spatially integral fractional Laplacian (IFL), many
previous studies (cf., e.g., previous works) always substitute the space Riesz fractional derivative for the IFL. In
fact, such two kinds of definitions are not equivalent in high-dimensional cases. It means that the “direct” study of
time–space fractional diffusion models with the IFL should be worthily considered.

In this paper, we study an alternative time–space fractional diffusion equation (TSFDE) with variable coefficients in
one space dimension

\[
\begin{aligned}
\frac{\partial}{\partial t} D_t^\gamma u(x, t) &= -\kappa(x, t)(-\Delta)^{\gamma/2} u(x, t) + f(x, t), \\
u(x, t) &= 0, \\
u(x, 0) &= \phi(x),
\end{aligned}
\]

(1.1)

where \(\kappa(x, t) > 0\) denotes the diffusivity coefficients, \(\Omega = (-l, l), \Omega^c = \mathbb{R} \setminus \Omega\), and the initial condition \(\phi(x)\) and the source
term \(f(x, t)\) are known functions. Meanwhile, \(\frac{\partial}{\partial t} D_t^\gamma\) is the Caputo derivative of order \(\gamma\)—that is,

\[
\frac{\partial}{\partial t} D_t^\gamma u(x, t) = \left\{ \begin{array}{ll}
\frac{1}{\Gamma(\gamma-1)} \int_0^t \frac{1}{t-s} \frac{\partial u(x,s)}{\partial s} ds, & 0 < \gamma < 1, \\
\frac{\partial}{\partial t} u(x, t), & \gamma = 1,
\end{array} \right.
\]

(1.2)

and throughout the paper, we always assume that \(0 < \gamma < 1\). Here, the fractional Laplacian \((-\Delta)^{\gamma/2}\) is defined by

\[(\Delta)^{\gamma/2} u(x) = c_{1,\alpha} \text{P.V.} \int_{\mathbb{R}} \frac{u(x) - u(x')}{|x - x'|^{1+\alpha}} dx', \quad \alpha \in (0, 2),\]

(1.3)

where P.V. stands for the Cauchy principal value and \(|x - x'|\) denotes the Euclidean distance between points \(x\) and \(x'\). The
normalization constant \(c_{1,\alpha}\) is defined as

\[
c_{1,\alpha} = \frac{2^{\alpha-1} \alpha \Gamma(\frac{\alpha+1}{2})}{\sqrt{\pi} \Gamma(1 - \alpha/2)}
\]

(1.4)

with \(\Gamma(\cdot)\) denoting the gamma function. From a probabilistic point of view, the IFL represents the infinitesimal generator of a symmetric \(\alpha\)-stable Lévy process. Mathematically, the well-posedness/regularity of the Cauchy problem or
uniqueness of the solutions of the TSFDE (1.1) has been studied in previous works.

Due to the nonlocality, the analytical (or closed-form) solutions of TSFDEs (1.1) on a finite domain are rarely available.
Therefore, we have to rely on numerical treatments that produce approximations to the desired solutions; refer, for
example, to previous studies and references therein for a description of such approaches. In fact, utilizing the
suitable temporal discretization, most of the early established numerical methods including the finite difference (FD)
method, finite element (FE) method, and matrix (named it as all-at-once) method for the TSFDE (1.1) were
developed via the fact that the IFL is equivalent to the Riesz fractional derivative in one space dimension. However,
such a numerical framework cannot be directly extended to solve the two- and three-dimensional TSFDEs due to the IFL
\((-\Delta)^{\gamma/2} u(x, y) \neq -\frac{\partial^2 u(x, y)}{\partial x^2} - \frac{\partial^2 u(x, y)}{\partial y^2}\). Therefore, it will hinder the development of numerical solutions for TSFDEs
from the stated objective.

In order to remedy the above drawback, Duo et al. replaced the IFL in TSFDE (1.1) by the spectral fractional Lapla-
cian and presented a fast numerical approach that combines the matrix transfer method with inverse Laplace
transform for solving the one-dimensional and multidimensional TSFDEs (1.1) with constant coefficients. Although
the numerical results show that their proposed method converges with the second-order accuracy in both time and space variables, the spectral fractional Laplacian on a bounded domain is also not equivalent to the IFL at all. On the other hand, Nochetto et al. used the Caffarelli–Silvestre extension to rewrite the TSFDE (1.1) with \(\kappa(x, t) \equiv \kappa\) as a two-dimensional quasi-stationary elliptic problem with dynamic boundary condition. Then, they established an FE scheme for solving the converted elliptic problem and showed that the numerical scheme cannot reach the error estimates of order \(O(\tau^{2-\gamma})\)
claimed in the literature. Later, Hu et al.\textsuperscript{35,36} successively exploited the similar strategy with FD approximation for the converted elliptic problem of one-dimensional and multidimensional TSFDEs (1.1) with $\kappa(x,t) \equiv \kappa$. Nevertheless, the numerical results showed that such FD schemes often converge with the less than first- and second-order accuracy in time and space, respectively, even for TSFDEs with sufficient smooth solutions.

In fact, it is important to set up numerical schemes that utilize the "direct" discretizations of IFL for solving the TSFDEs (1.1). Moreover, the discretizations of (multidimensional) TSFDEs became a recently hot topic, with the main numerical challenge stemming from the approximation of the hypersingular integral (see, e.g., previous works\textsuperscript{10,37-45}). Indeed, there are some numerical schemes that utilize the temporal $L_1$ formula\textsuperscript{46} (or numerical Laplace inversion\textsuperscript{24}) and spatial FE discretization\textsuperscript{10,26,41-43} for solving the (multidimensional) constant-coefficient TSFDEs (1.1).\textsuperscript{7,25,26} Both the theoretical and numerical results are reported to show that such numerical schemes are efficient to solve the (multidimensional) TSFDEs (1.1) with $\kappa(x,t) \equiv \kappa$. In addition, there are some other kinds of time–space fractional diffusion models but related to TSFDEs (1.1), where the spatial (or temporal) nonlocal operator is a replacement for the IFL (or the Caputo fractional derivative). This is mainly because the nonlocal operators with suitable kernels can exactly embrace the IFL and the Caputo fractional derivative, respectively.\textsuperscript{9,10,47-49} For such novel model problems, Guan and Gunzburger\textsuperscript{47} established a class of numerical methods that utilized the $\theta$ schemes and piecewise-linear FE discretization. Their fully discrete scheme is analyzed for all to determine conditional and unconditional stability regimes for the scheme and also to obtain error estimates for the approximate solution. Later, Liu et al.\textsuperscript{48} improved the idea of Guan and Gunzburger by giving the proof of convergence behavior with $O(r^{2-\gamma} + h^2)$. Meanwhile, Liu et al. considered the piecewise-quadratic FE discretization to improve the spatial convergence rate. The efficient implementation based on fast Toeplitz-matrix multiplications\textsuperscript{50-52} of their proposed scheme is also reported. For space–time nonlocal diffusion equations, Chen et al.\textsuperscript{49} proposed a numerical scheme, by exploiting the quadrature-based FD method in time and the Fourier spectral method in space, and showed its stability. Moreover, it is shown that the convergence is uniform at a rate of $O(\delta + \sigma^2)$ (where $\delta$ and $\sigma$ are the time and space horizon parameters) under certain regularity assumptions on initial and source data. Even though these are several methods with linear solvers of quasilinear complexity, the implementation of the above methods is still complicated, especially the computation of entries of stiffness matrix in FE discretization or finding the modes in terms of expansion basis in spectral method (cf. previous works\textsuperscript{26,44,48,49}). In particular, it is pointed out that “more than 95% of CPU time is used to assembly routine” for their FE methods.\textsuperscript{43}

On the other hand, most of the abovementioned methods overlook that the presence of the kernel $(t-s)^{-\gamma}$ results in a weak initial singularity in the solution of Equation (1.1), so that approximation methods (e.g., $L_1$ formula) on uniform meshes have a poor convergent rate and high computational cost.\textsuperscript{34,49,53,54} In this work, we devote ourselves in developing fast implicit difference schemes (IDSs) for solving the TSFDE (1.1); the direct scheme utilizes the simple FD discretization\textsuperscript{58} and the graded $L_1$ formula\textsuperscript{54} for approximating the IFL and the Caputo fractional derivative, where the nonuniform temporal discretization can overcome the initial singularity. Due to the repeated summation of numerical solutions in the previous steps, the direct scheme always needs much CPU time and memory cost, especially for the larger number of time steps.\textsuperscript{55-57} In order to alleviate the computational cost, the sum-of-exponential (SOE) approximation\textsuperscript{28} of the kernel $(t-s)^{-\gamma}$ in the graded $L_1$ formula for Caputo fractional derivative can be efficiently evaluated via the recurrence method. Thus, we can derive the fast implicit difference scheme. In particular, we revisit the matrix properties of the discretized IFL and prove the discretized matrix is a strictly diagonally dominant and symmetric $M$-matrix with positive diagonal elements (i.e., the symmetric positive definite matrix), which is not studied in Duo et al.\textsuperscript{38} Based on such matrix properties, we strictly prove that the fast numerical schemes for the TSFDE (1.1) are unconditionally stable and present the corresponding error estimates of $O(M^{-\min(1/\gamma,2-\gamma)}+h^2)$ ($h$ is the spatial grid size) under certain regularity assumptions on the smooth solutions. To our best knowledge, there are few successful attempts to derive the efficient IDSs for solving the TSFDE (1.1) with rigorous theoretical analyses. This is one of the main attractive advantages of our proposed methods compared with the above mentioned methods.

In addition, the nonlocality of IFL results in dense discretized linear systems, which is the leading time-consuming part in practical implementations.\textsuperscript{26,43,44} Fortunately, the coefficient matrix of discretized linear systems enjoys the Toeplitz-like structure;\textsuperscript{38-40} it means that we solve the sequence of discretized linear systems in a matrix-free pattern;\textsuperscript{26,52,59} because the Toeplitz matrix–vector products can be computed via fast Fourier transforms (FFTs) in $O(N \log N)$ operations. More precisely, we will adapt the circulant preconditioners\textsuperscript{50-52} for accelerating the Krylov subspace solvers\textsuperscript{51} for the sequence of discretized linear systems. Moreover, the benefit of circulant preconditioners will be verified via both theoretical and numerical results. It notes that fast schemes greatly reduce the computational work of solving the discretized
linear systems from $\mathcal{O}(MN^3 + M^2N)$ by a direct solver to $\mathcal{O}(MN(\log N + N_{\exp}))$ per preconditioned Krylov subspace iteration and a memory requirement from $\mathcal{O}(MN^2)$ to $\mathcal{O}(NN_{\exp})$ (see Section 2.1 for defining $N_{\exp}$).

The contributions of the current work can be summarized as follows.

- We present two IDSs for solving the TSFDE (1.1) with nonsmooth initial data and such numerical schemes can be easily extended to solve the multidimensional cases.
- Both the stability and convergence of these IDSs are rigorously proved via the discretized matrix properties.
- We provide the efficient implementation of fast IDSs with theoretical guarantee for reducing the computation and memory cost deeply.

The rest of this paper is organized as follows. In Section 2, both direct and fast IDSs are derived for the TSFDE (1.1) in details, and their stability and convergence are proved by revisiting the properties of spatial discretized matrix. In Section 3, the efficient implementation based on fast preconditioned Krylov subspace solvers of the proposed IDSs are given and the details, and their stability and convergence are proved by revisiting the properties of spatial discretized matrix. In Section 3, the efficient implementation of fast IDSs with theoretical guarantee for reducing the computation and memory cost deeply.

2 | DIRECT AND FAST IMPLICIT DIFFERENCE SCHEMES

In this section, we will establish two implicit difference schemes for solving the problem (1.1). Meanwhile, the stability, convergence, and error analysis of such difference schemes are investigated and proved in details.

2.1 | Two implicit difference schemes

As mentioned above, we assume that the problem (1.1) has a solution $u(x,t)$ such that

$$
\left| \frac{\partial^k u(x,t)}{\partial t^k} \right| \leq c_0 t^{-k}, \quad 0 \leq k \leq 3.
$$

(2.1)

Here and in what follows, $\hat{C}$ and $c_j$ $(j = 0, 1, 2)$ are positive constants, which depend on the problem but not on the mesh parameters.\(^{54}\) Let $M, N, r \in \mathbb{N}^+$ (positive integers), $h = 2L/N \Delta x = -l + ih$, $t_m = (m/M)^r T$, and $\tau_m = t_m - t_{m-1}$. We also consider the sets

$$
\Omega_h = \{ x_i | 0 \leq i \leq N \}, \quad \Omega_r = \{ t_m | 0 \leq m \leq M \}, \quad \Omega_{hr} = \{ (x_i, t_m) | 0 \leq i \leq N, \ 0 \leq m \leq M \},
$$

and let $v = \{ v_i^m | 0 \leq i \leq N, \ 0 \leq m \leq M \}$ be a grid function on $\Omega_{hr}$. We define the set $\mathcal{V}_h = \{ v | v = (v_0, v_1, \ldots, v_{N-1}, v_N) \}$ of grid functions on $\Omega_h$ and provide it with the norm

$$
\| v \|_\infty = \max_{0 \leq i \leq N} | v_i |.
$$

For $m \geq 1$, we approximate the Caputo fractional derivative (1.2) by the $L1$ formula on the graded mesh, which can capture the weak initial singularity of (1.1):

$$
\mathcal{C}D^\gamma_0 D^0_1 u(t_m) \approx D^\gamma_1 D^0_1 u(t_m) \triangleq \frac{1}{\Gamma(1 - \gamma)} \left[ a^{m,\gamma}_m u(t_m) - \sum_{k=1}^{m-1} (a^{m,\gamma}_k - a^{m,\gamma}_{k+1}) u(t_k) - a^{m,\gamma}_1 u(t_0) \right],
$$

where

$$
a^{m,\gamma}_k = \frac{1}{\tau_k} \int_{t_{k-1}}^{t_k} ds \frac{ds}{(t_m - s)^\gamma}, \quad k \geq 1.
$$

(2.2)

The truncation error $\psi^m$ can be defined by $\psi^m \triangleq \mathcal{C}D^\gamma_0 D^0_1 u(t_m) - D^\gamma_1 D^0_1 u(t_m)$. From Shen et al.\(^{54}\), Lemma 2.1 we obtain the boundedness of the truncation error

$$
| \psi^m | \leq \hat{C} m^{-\min\{r(1+\gamma), 2-\gamma\}}, \quad m = 1, 2, \ldots, M.
$$

(2.3)

if $|u''(t)| \leq c_0 t^{-2}$, $0 \leq t \leq T$, and $\hat{C} > 0$. 


On the other hand, it notes that the above graded $L1$ scheme always needs much computational cost in practical applications due to the repeatedly weighted sum of the solutions of previous time steps. To reduce the cost, here, it is useful to develop the fast approximation of Caputo fractional derivative on a nonuniform temporal mesh.

**Lemma 2.1** (Jiang et al\textsuperscript{58}). Let $\epsilon, \delta$, and $T$ denote the tolerance error, cut-off time restriction, and final time, respectively. Then there exist $N_{\text{exp}} \in \mathbb{N}^+$ and $s_i, w_j > 0$, $j = 1, 2, \ldots, N_{\text{exp}}$ such that

$$| \epsilon^\gamma - \sum_{j=1}^{N_{\text{exp}}} w_j e^{-s_j t_j} | \leq \epsilon, \quad \text{for any } t \in [\delta, T],$$

where $N_{\text{exp}} = \mathcal{O}((\log \epsilon^{-1})(\log \log \epsilon^{-1} + \log(T\delta^{-1}))) + (\log \delta^{-1})(\log \log \epsilon^{-1} + \log \delta^{-1})).$

Based on Lemma 2.1, we set $\delta = (1/M)^T$, and then the fast approximation of Caputo fractional derivative on a graded temporal grid can be drawn as follows ($m \geq 1$).

$$\mathcal{D}_t^\gamma u(t_m) = \mathcal{F}\mathcal{D}_t^\gamma u(t_m) + \mathcal{O}(m^{-\min(r(1+\gamma),2-\gamma)} + \epsilon)$$

where the estimate of truncation error holds when $|u'(t)| \leq c_0 t^{-1}$ and $|u''(t)| \leq c_0 t^{-2}$ and

$$b_k^{(m, \gamma)} = \left\{ \begin{array}{ll}
\sum_{j=1}^{N_{\text{exp}}} w_j \int_{t_j}^{t_k} e^{-\gamma s} (s^m) ds, & k = 1, 2, \ldots, m - 1, \\
\alpha_k^{(m, \gamma)}, & k = m.
\end{array} \right. \quad (2.5)$$

Moreover, we provide the information for approximating the IFL with extended Dirichlet boundary conditions in (1.1). According to the idea in Duo et al\textsuperscript{38}, the approximation is given by

$$(-\Delta)^{\alpha/2} u(x_i) = c_{a, \mu} \left\{ \sum_{\ell=2}^{N} \left( (\ell + 1)^\mu - (\ell - 1)^\mu \right) \right/ \ell \mu + \frac{N^\mu - (N - 1)^\mu}{N^\mu} + \frac{2^\nu + \kappa - 1}{\mu (\alpha/a)} \right\} u(x_i)$$

$$- \frac{2^\nu + \kappa - 1}{2} \left( u(x_{i-1}) + u(x_{i+1}) \right) - \frac{1}{2} \sum_{j=1, j \neq i \pm 1}^{N-1} \left( |j - i| + 1 \right)^\mu - \left( |j - i| - 1 \right)^\mu \right/ |j - i|^\mu u(x_j), \quad (2.6)$$

where $i = 1, 2, \ldots, N - 1$, $C_{a, \mu} = c_{1, a}/(\nu h^\alpha) > 0$ and the constant $\kappa = 1$ for $\mu \in (a, 2)$, while $\kappa = 2$ if $\mu = 2$. Meanwhile, we denote $\nu = \mu - a$ for notational simplicity.

**Lemma 2.2** (Duo et al\textsuperscript{38}). Suppose that $u(x) \in C^{k+\frac{\nu}{2}}(\mathbb{R})$ has the finite support on an open set $\Omega \subset \mathbb{R}$, and Equation (2.6) is a finite difference approximation of the fractional Laplacian $(-\Delta)^{\alpha/2}$. Then, for any $\mu \in (a, 2)$, there is

$$\|(-\Delta)^{\alpha/2} u(x) - (-\Delta)^{\alpha/2} u(x)\|_{\|\cdot\|_p} \leq \tilde{C} h^p \quad (2.7)$$

with $\tilde{C} > 0$ depending on $a$ and $\mu$. Here, $p \in (0, 2]$ would be determined via the regularity (i.e., the index $s \in \mathbb{N}$) of $u(x)$.

Lemma 2.2 provides a direct discretization for the IFL that appeared in the TSFDE (1.1). At present, the spatial and temporal discretizations are ready for developing the numerical methods. Evaluating Equation (1.1) at points $(x_i, t_m)$, we have

$$\sum_{j=1}^{N} (\alpha_j - \lambda_j) D_t^\gamma u(x_j, t_m) = -\kappa(x_i, t_m)(-\Delta)^{\alpha/2} u(x_i, t_m) + f(x_i, t_m), \quad (2.8)$$

where $1 \leq i \leq N - 1, 1 \leq m \leq M$. Let $u = \{ u_{i,m}^j \}_{0 \leq i \leq N, 0 \leq m \leq M}$ be a grid function defined by

$$U_{i,m}^j := u(x_i, t_m), \quad f_m^i = f(x_i, t_m), \quad \kappa_m^i = \kappa(x_i, t_m), \quad 0 \leq i \leq N, 0 \leq m \leq M.$$
Using these notations and recalling Equation (2.4) along with Lemma 2.2, we can approximate Equation (1.1) at grid point \((x_i, t_m)\) as follows:

\[
\begin{align*}
\frac{\partial}{\partial t} U_i^m & = -\kappa_i^m (-\Delta)_{h,\mu}^{q/2} U_i^m + f_i^m + R_i^m, \quad 1 \leq i \leq N - 1, \quad 1 \leq m \leq M, \\
U_0^m & = U_N^m = 0, \\
U_i^0 & = \phi(x_i), \\
1 \leq i & \leq N - 1,
\end{align*}
\]  

(2.9)

where the terms \([R_i^m]\) are small and satisfy the inequality

\[
|R_i^m| \leq c_2 (m - \min(n(1+\gamma),2-\gamma) + h^p + \epsilon), \quad 1 \leq i \leq N - 1, \quad 1 \leq m \leq M.
\]

We omit the above small terms and arrive at the following implicit difference scheme

\[
\begin{align*}
\frac{1}{\Gamma(1-\gamma)} \left[ b_i^{(m,\gamma)} u_i^m - \sum_{k=1}^{m-1} (b_{k+1}^{(m,\gamma)} - b_k^{(m,\gamma)}) u_k^k - b_1^{(m,\gamma)} u_0^0 \right] & = -\kappa_i^m (-\Delta)_{h,\mu}^{q/2} u_i^m \\
+ f_i^m, & \quad 1 \leq i \leq N - 1, \quad 1 \leq m \leq M, \\
0 \leq m \leq M, \\
1 \leq i & \leq N - 1,
\end{align*}
\]  

(2.10)

which is named as the fast implicit difference scheme (FIDS). Similarly, we combine the graded L1 formula with Lemma 2.2 for deriving the following difference scheme

\[
\begin{align*}
\frac{1}{\Gamma(1-\gamma)} \left[ a_i^{(m,\gamma)} u_i^m - \sum_{k=1}^{m-1} (a_{k+1}^{(m,\gamma)} - a_k^{(m,\gamma)}) u_k^k - a_1^{(m,\gamma)} u_0^0 \right] & = -\kappa_i^m (-\Delta)_{h,\mu}^{q/2} u_i^m \\
+ f_i^m, & \quad 1 \leq i \leq N - 1, \quad 1 \leq m \leq M, \\
0 \leq m \leq M, \\
1 \leq i & \leq N - 1,
\end{align*}
\]  

(2.11)

which is labeled as direct implicit difference scheme (DIDS).

At each time level, both FIDS (2.10) and DIDS (2.11) are the resultant linear systems, which can be solved by the direct method (e.g., Gauss elimination) with total computational cost \(O(N^3M + NMN_{\exp})\) for FIDS and \(O(N^3M + NM^2)\) for DIDS. Note that, generally, \(N_{\exp} < 100^{54,58}\) and (if) \(M\) is very large, so that FIDS requires smaller computational cost than DIDS. Moreover, FIDS only requires \(O(N^2 + NN_{\exp})\) memory units rather than \(O(N^2 + NM)\) for DIDS. In Section 3, we will further reduce the computational cost of both FIDS and DIDS by means of matrix-free preconditioned iterative solvers.

### 2.2 The stability and convergence

In this subsection, we discuss the stability and convergence of the difference scheme for the problem (1.1). In order to analyze the stability and convergence, we rewrite the FIDS (2.10) into the matrix form

\[
\mathcal{M}^{(m)} u^m = \frac{1}{\Gamma(1-\gamma)} \left[ \sum_{k=1}^{m-1} (b_{k+1}^{(m,\gamma)} - b_k^{(m,\gamma)}) u_k^k - b_1^{(m,\gamma)} u_0^0 \right] + f^m,
\]  

(2.12)

where \(\mathcal{M}^{(m)} = \frac{1}{\Gamma(1-\gamma)} b_i^{(m,\gamma)} I + K^{(m)} A\) and refer to Equation (2.13) for the definition of \(A\), \(I\) is the identity matrix of order \(N - 1\), \(K^{(m)} = \text{diag}(\kappa_1^m, \kappa_2^m, \ldots, \kappa_{N-1}^m)\), \(u^m = [u_1^m, u_2^m, \ldots, u_{N-1}^m]^T\), \(f^m = [f_1^m, f_2^m, \ldots, f_{N-1}^m]^T\), and \(b_i^{(m,\gamma)} > 0^{54, \text{Lemma 2.4}}\). First of all, we revisit the properties of spatial discretization, which is not deeply studied in the original paper. In fact, the spatial discretization (2.6) of \((-\Delta)^{q/2} u(x,t)\) can be expressed in the matrix–vector product form \((-\Delta)^{q/2} u^m = A u^m\),
where $A = [a_{ij}]_{i,j=1, \ldots, N-1}$ is the matrix representation of the (discretized) fractional Laplacian, defined as

$$a_{ij} = C_{a,\mu}^h \left\{ \begin{array}{ll}
\sum_{\ell=2}^{N-1} \frac{(\ell+1)^\nu - (\ell-1)^\nu}{\ell^\mu} + \frac{N^\nu - (N-1)^\nu}{N^\mu} + (2^\nu + \kappa \mu - 1) + \frac{2\nu}{a N^\nu}, & j = i, \\
\frac{2}{(j-i+1)^\nu - (j-i-1)^\nu}, & j = i \pm 1, \\
\frac{2}{2|j-i|^\nu}, & j \neq i, i \pm 1,
\end{array} \right. \quad (2.13)$$

where $i,j = 1,2, \ldots, N-1$. It is easy to see that the matrix $A$ is a real symmetric Toeplitz matrix, which can be stored with only $(N-1)$ entries.\textsuperscript{38,50,51} Moreover, we can give the following conclusions.

**Proposition 2.1.** According to the definition of $A$, the following holds:

1) $A$ is a strictly diagonally dominant $M$-matrix;
2) $A$ is symmetric positive definite;
3) The absolute values of the entries $a_{ij}$ away from the diagonals decay gradually, that is, $a_{11} > |a_{12}| > \ldots > |a_{1,N-1}|$ and $\lim_{i \to \infty} |a_{1,i-1}| = 0$.

**Proof.** (1) Because $A$ is a symmetric Toeplitz matrix, then the diagonal entries are equal to $a_{11} > 0$. Moreover, it is not hard to see that $a_{ij} < 0$ ($i \neq j$). So we conclude that $A$ is an $M$-matrix\textsuperscript{60, p. 533} and obtain

$$a_{11} - \sum_{j \neq 1} |a_{1j}| = a_{11} - \sum_{j=2}^{N-1} |a_{1j}| > 0, \quad (2.14)$$

and

$$a_{N-1,N-1} - \sum_{j \neq N-1} |a_{N-1,j}| = a_{11} - \sum_{j=2}^{N-1} |a_{N-1,j}| > 0. \quad (2.15)$$

Similarly, it follows that

$$a_{ij} - \sum_{j \neq i} |a_{ij}| = a_{ij} - \sum_{j \neq i} a_{ij} > 0, \quad i = 2, 3, \ldots, N - 2. \quad (2.16)$$

A combination of the aforementioned three inequalities verifies that $A$ is a strictly diagonally dominant $M$-matrix.

(2) Because $A$ is a symmetric strictly diagonally dominant $M$-matrix and all its diagonal entries are positive, that is, $a_{ii} = a_{11} > 0$, then $A$ is indeed a symmetric positive definite matrix.\textsuperscript{60, Corollary 7.2.3}

(3) First of all, we rewrite the matrix $A = C_{a,\mu}^h A = C_{a,\mu}^h [\tilde{a}_{ij}]_{i,j=1,\ldots,N-1}$ (cf. Equation 2.13). Meanwhile, it is easy to note that $a_{11} > |a_{12}|$, then we find

$$|\tilde{a}_{12}| - |\tilde{a}_{13}| = \frac{2^\nu + \kappa \mu - 1}{2} - \frac{3^\nu - 1}{2^{\nu+1}} \geq \frac{4^{\nu+2} - 3^\nu + 1}{2^{\nu+1}} > 0. \quad (2.17)$$

For $j = 3, 4, \ldots$, we set $|j - i| = |j - 1| = k$; thus, $k \geq 2$ and

$$|\tilde{a}_{1j}| := f(k) = \frac{k^{a} - 1}{2} \left[ \left( \frac{1 + 1}{k} \right)^\nu - \left( \frac{1 - 1}{k} \right)^\nu \right] = \left[ \left( \frac{\nu}{3} \right) k^{1-a} + \left( \frac{\nu}{5} \right) k^{3-a} + \ldots \right]. \quad (2.18)$$

which should imply that $f(k) > f(k+1)$. Therefore, it follows that $a_{11} > |a_{12}| > \ldots > |a_{1,N-1}|$.\]^
According to Proposition 2.1, if we define

$$D(C) = \min_{1 \leq i \leq N-1} \left( |C_{ii}| - \sum_{1 \leq j \leq N-1, j \neq i} |C_{ij}| \right)$$

(2.19)

for any matrix \( C = [C_{ij}]_{i=1, \ldots, N-1} \), then it follows that \( D(A) \geq 0 \), which is helpful in the next context. The following properties of the operator \( D(\cdot) \) can be given as follows.

**Lemma 2.3** (Lin and Ng\(^{61, \text{Lemma 3}}\)). Let \( C_1, C_2 \in \mathbb{R}^{(N-1) \times (N-1)} \). Suppose both \( C_1 \) and \( C_2 \) have positive diagonal entries. Then it follows that \( D(C_1 + C_2) \geq D(C_1) + D(C_2) \).

**Lemma 2.4** (Lin and Ng\(^{61, \text{Lemma 4}}\)). Let \( C \in \mathbb{R}^{(N-1) \times (N-1)} \). Suppose \( D(C) \geq 0 \). Then for any nonnegative diagonal matrix \( K \in \mathbb{R}^{(N-1) \times (N-1)} \), it holds \( D(KC) \geq D(C) \min_{1 \leq j \leq N-1} K_{jj} \geq 0 \).

Next, we exploit the above two lemmas to give the following estimation about the coefficient matrices \( M^{(m)} \) of Equation (2.12).

**Theorem 2.1.** For any \( \frac{b_{m}^{(m,y)}}{\Gamma(1-\gamma)} > 0 \) and \( 1 \leq m \leq M \), it holds

$$\min_{1 \leq m \leq M} D\left( \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,y)} I + K^{(m)} A \right) \geq \frac{b_{m}^{(m,y)}}{\Gamma(1-\gamma)}.$$  

(2.20)

Proof. From Lemmas 2.3 and 2.4 and Proposition 2.1, we obtain

$$D\left( \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,y)} I + K^{(m)} A \right) \geq D\left( \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,y)} I \right) + D(K^{(m)} A) \geq \frac{b_{m}^{(m,y)}}{\Gamma(1-\gamma)}, \quad 1 \leq m \leq M,$$

from which the result follows.

Before proving the final result of this section on the unconditional stability and convergence property of the FIDS (2.10), we recall the following useful lemma.

**Lemma 2.5.** Suppose \( C \in \mathbb{R}^{(N-1) \times (N-1)} \) satisfies \( D(C) \geq \lambda > 0 \). Then, for any \( y \in \mathbb{R}^{N-1} \), it holds \( \lambda \|y\|_{\infty} \leq \|Cy\|_{\infty} \).

Proof. Because \( D(C) \geq \lambda > 0 \), then \( D\left( \frac{1}{\lambda} C \right) \geq 1 \) and \( \|y\|_{\infty} \geq \left\| \frac{C}{\lambda} y \right\|_{\infty} \), which proves the above result.

**Theorem 2.2.** The proposed FIDS (2.10) with \( c \leq c_{1} M^{\gamma} \) is uniquely solvable and unconditionally stable in the sense that

$$\|u^{k}\|_{\infty} \leq \|u^{0}\|_{\infty} + \Gamma(1-\gamma) \max_{1 \leq s \leq k} \frac{\|f^{s}\|_{\infty}}{b_{1}^{(s,y)}}, \quad k = 1, 2, \ldots, M,$$

(2.21)

where \( \|f\|_{\infty} \leq \max_{1 \leq \eta \leq N-1} |f^{\eta}| \).

Proof. It is easy to see that proving the unique solvability of FIDS (2.10) is equivalent to showing the invertibility of coefficient matrices \( M^{(m)} \) with each \( 1 \leq m \leq M \). By means of Theorem 2.1 and Lemma 2.5, it follows that

$$\|M^{(m)} y\|_{\infty} = \left\| \left[ \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,y)} I + K^{(m)} A \right] y \right\|_{\infty} \geq \frac{b_{m}^{(m,y)}}{\Gamma(1-\gamma)} \|y\|_{\infty}, \quad \forall y \in \mathbb{R}^{N-1},$$

(2.22)

where \( 1 \leq m \leq M \). Therefore, \( M^{(m)} : \mathbb{R}^{N-1} \rightarrow \mathbb{R}^{N-1} \) is clearly an injection for each \( 1 \leq m \leq M \), whose null space is simply \( \{0\} \). Hence, \( M^{(m)} \)'s are nonsingular, which proves the unique solvability.

On the other hand, we apply Equation (2.22) and the monotonicity of \( \{b_{k}^{(m,y)}\} \) \( 1 \leq m \leq M \)\(^{54, \text{Lemma 2.4}}\) to obtain

$$b_{m}^{(m,y)} \|u^{m}\|_{\infty} \leq \sum_{k=1}^{m-1} (c_{k}^{(m,y)} - c_{k}^{(m,y)}) \|u^{k}\|_{\infty} + b_{1}^{(m,y)} \left[ \|u^{0}\|_{\infty} + \frac{\Gamma(1-\gamma)}{b_{1}^{(m,y)}} \|f^{m}\|_{\infty} \right], \quad 1 \leq m \leq M.$$
Then the inequality (2.21) can be proved by the method of mathematical induction, which is similar to the proof of Shen et al.\(^5\), Theorem 4.1. We omit the details here.

On the other hand, we replace the coefficients \(\{b_k^{(m,\gamma)}\}\) with \(\{a_k^{(m,\gamma)}\}\) in Theorem 2.1 and Equations (2.12) and (2.22), and then we can obtain the following conclusion, which is helpful to analyze the stability and convergence of DIDS (2.11).

**Theorem 2.3.** The proposed DIDS (2.11) is uniquely solvable and unconditionally stable in the sense that

\[
\|u^k\|_\infty \leq \|u^0\|_\infty + \Gamma(1 - \gamma) \max_{1 \leq i \leq k} \frac{\|f_i\|_\infty}{a_1^{(s,\gamma)}}, \quad k = 1, 2, \ldots, M. \tag{2.23}
\]

**Proof.** The proof of this theorem is similar to Theorem 2.3; we omit the details here.

From Theorems 2.1–2.3, we can see that both FIDS (2.10) and DIDS (2.11) are stable to the initial value \(\phi\) and the right-hand term \(f\). Now, we consider the convergence of these two difference schemes.

**Theorem 2.4.** Let \(\{U_i^m\}^N_0 \leq i \leq N, 0 \leq m \leq M\) and \(\{U_i^m\}^N_0 \leq i \leq N, 0 \leq m \leq M\) be, respectively, the solutions of the problem (1.1) and the difference scheme (2.10). If \(\epsilon \leq \min\{c_1 M^p, T^{-\gamma/2}\}\), then

\[
\|e_i^m\|_\infty \leq 2\Gamma(1 - \gamma)c_2 T^p (M^{-\min(\gamma, 2-\gamma)} + h^p + \epsilon), \quad 1 \leq m \leq M, \tag{2.24}
\]

where \(e_i^m = U_i^m - u_i^m\), \(0 \leq i \leq N, 0 \leq m \leq M\), and \(p\) would be determined by the spatial regularity of \(u(x,t)\).

**Proof.** Writing the system (2.9) as

\[
\begin{aligned}
&\frac{1}{\Gamma(1 - \gamma)} \left[ b_m^{(m,\gamma)} U_i^m - \sum_{k=1}^{m-1} (b_k^{(m,\gamma)} - b_k^{(m,\gamma)}) U_i^k - b_1^{(m,\gamma)} U_i^0 \right] + f_i^m + R_i^m, & 1 \leq i \leq N - 1, \quad 1 \leq m \leq M, \\
&U_0^m = U_N^m = 0, & 0 \leq m \leq M, \\
&U_0^0 = \phi(x), & 1 \leq i \leq N - 1,
\end{aligned}
\]

and subtracting Equation (2.10) from the above corresponding system,

\[
\begin{aligned}
&\frac{1}{\Gamma(1 - \gamma)} \left[ b_m^{(m,\gamma)} e_i^m - \sum_{k=1}^{m-1} (b_k^{(m,\gamma)} - b_k^{(m,\gamma)}) e_i^k - b_1^{(m,\gamma)} e_i^0 \right] + R_i^m, & 1 \leq i \leq N - 1, \quad 1 \leq m \leq M, \\
&e_0^m = e_N^m = 0, & 0 \leq m \leq M, \\
&e_0^0 = 0, & 1 \leq i \leq N - 1.
\end{aligned}
\tag{2.25}
\]

By means of Theorem 2.1 and the matrix analysis described above, it follows that

\[
\|e_i^m\|_\infty \leq \Gamma(1 - \gamma) \max_{1 \leq i \leq m} \frac{\|R_i^m\|_\infty}{b_1^{(s,\gamma)}}, \quad m = 2, \ldots, M.
\]

The rest of this proof is also similar to the work of Shen et al.\(^5\), Theorem 4.2.

Again, we employ the similar strategy to give the error analysis of DIDS (2.11) as follows.

**Theorem 2.5.** Let \(\{U_i^m\}^N_0 \leq i \leq N, 0 \leq m \leq M\) and \(\{U_i^m\}^N_0 \leq i \leq N, 0 \leq m \leq M\) be, respectively, the solutions of the problem (1.1) and the difference scheme (2.11), then

\[
\|e_i^m\|_\infty \leq O(M^{-\min(\gamma, 2-\gamma)} + h^p), \quad 1 \leq m \leq M, \tag{2.26}
\]
where \( e_i^m = U_i^m - u_i^m, \ 0 \leq i \leq N, \ 0 \leq m \leq M. \)

In practice, the value of \( \epsilon \) is sufficiently small such that the tolerance error in (2.24) can be negligible compared with the space and time errors. Then it also finds that the numerical errors for DIDS and FIDS are almost identical but the later is often faster (cf. Section 4). With the help of arguments in proving, 38, Theorems 3.1–3.2 it is not hard to make the convergence results described in Theorems 2.4 and 2.5 more specific.

**Remark 2.1.** For determining the value of \( p \), it reads

- suppose that the solution of the problem (1.1) satisfies the condition (2.1) and \( u(x, \cdot) \in C^{\alpha/2}(\mathbb{R}) \) with \( (x, t) \in \mathbb{R} \times [0, T] \) and \( s \geq 1 \), then solutions of FIDS (2.10) and DIDS (2.11) with \( \mu \in (a, 2] \) converge to the exact solutions of Equation (1.1), respectively.
- for \( s = 1, \ a \in (0, 2) \) and \( \mu \in (a, 2] \), the convergence rates of FIDS (2.10) and DIDS (2.11) are (at least)

\[
\mathcal{O}(M^{-\min\{r,2-\gamma\}} + h^{1-\frac{s}{2}} + \epsilon) \quad \text{and} \quad \mathcal{O}(M^{-\min\{r,2-\gamma\}} + h^{1-\frac{s}{2}}),
\]

respectively; see Duo et al 38 and Duo and Zhang 39 for details.
- for \( s \geq 3 \) and \( a \in (0, 2) \), the convergence rates of FIDS (2.10) and DIDS (2.11) with \( \mu = 2 \) or \( \mu = 1 + \alpha/2 \) are \( \mathcal{O}(M^{-\min\{r,2-\gamma\}} + h^2 + \epsilon) \) and \( \mathcal{O}(M^{-\min\{r,2-\gamma\}} + h^2) \), respectively.

Moreover, we will work out some numerical results for supporting the above theoretical convergence behaviors described in Section 4.

### 3 Efficient Implementation Based on Preconditioning of the Difference Schemes

In the section, we analyze both the implementation and computational complexity of FIDS (2.10) and DIDS (2.11) and propose an efficient implementation utilized preconditioned Krylov subspace solvers. Noting that \( a_m^{(m,r)} = b_m^{(m,r)} > 0 \), we start the efficient implementation from the following matrix form of these two implicit difference schemes at the time level \( 1 \leq m \leq M \), which are given by Equation (2.12) and

\[
\mathcal{M}^{(m)}u^m = \frac{1}{\Gamma(1-\gamma)} \left[ \sum_{k=1}^{m-1} (a_k^{(m,r)} - a_{k+1}^{(m,r)}) u^k + a_1^{(m,r)} u^0 \right] + f^m,
\]

(3.1)

respectively. From Theorems 2.2 and 2.3, it knows that both Equations (2.12) and (3.1) have the unique solutions. In addition, it is meaningful to remark that Equations (2.12) and (3.1) corresponding to FIDS (2.10) and DIDS (2.11) are inherently sequential; thus, both of them are difficult to parallelize over time.

#### 3.1 The circulant preconditioner

On the other hand, it is useful to note that the matrix–vector product \( \mathcal{M}^{(m)}v \) can be efficiently calculated by

\[
\mathcal{M}^{(m)}v = \frac{1}{\Gamma(1-\gamma)} K^{(m,r)}v + K^{(m)}(Av),
\]

(3.2)

where \( v \in \mathbb{R}^{N-1} \) is any vector and the Toeplitz matrix–vector \( Av \) can be implicitly evaluated via the FFTs in \( \mathcal{O}(N \log N) \) operations. In other words, we can use a matrix-free method to compute \( \mathcal{M}^{(m)}v \) quickly. Based on such observations, the Krylov subspace method should be the most suitable solver for Equation (2.12) or Equation (3.1) one by one. However, when the coefficients and the order of integral fractional Laplacian are not small, then the coefficient matrices \( \mathcal{M}^{(m)} \) will be increasingly ill-conditioned (cf. Section 4). This fact deeply slows down the convergence of the Krylov subspace method, while the preconditioning techniques are often used to overcome this difficulty. 29,50-52 In the literature on Toeplitz systems, circulant preconditioners always played important roles. 50,51 In fact, circulant preconditioners have been theoretically and numerically studied with applications to fractional partial differential equations for recent years; see, for instance, previous studies. 29,40,52
In this work, we design a family of the Strang preconditioners for accelerating the convergence of Krylov subspace solvers. More precisely, the circulant preconditioners are given for Equation (2.12) and/or Equation (3.1) as follows:

\[ P^{(m)} = \frac{1}{\Gamma(1-\gamma)} b^{(m,\gamma)} I + \kappa^{(m)} A = F^* \left[ \frac{1}{\Gamma(1-\gamma)} b^{(m,\gamma)} + \kappa^{(m)} \Lambda \right] F, \tag{3.3} \]

where \( F \) and \( F^* \) are the Fourier matrix and its conjugate transpose, respectively, and the scalar \( \kappa^{(m)} = \frac{1}{N-1} \sum_{i=1}^{N-1} \kappa_i^{(m)} \). Meanwhile, \( s(A) = F^* A F \) is the Strang circulant approximation of the Toeplitz matrix \( A \) and the diagonal matrix \( \Lambda \) contains all the eigenvalues of \( s(A) \) with the first column: \( \mathbf{e}_S = [a_{11}, \ldots, a_{11}, a_{11,1}, \ldots, a_{11}]^T \in \mathbb{R}^{N-1} \). Therefore, the matrix \( \Lambda = \text{diag}(F\mathbf{e}_S) \) can be computed in advance and only one time during each time level. Besides, as \( P^{(m)} \) are the circulant matrices, we observe from Equation (3.3) that the inverse matrix–vector product \( z = [P^{(m)}]^{-1} v \) can be carried out in \( \mathcal{O}(N \log N) \) operations via the (inverse) FFTs. In a word, we exploit a fast preconditioned Krylov subspace method with only \( \mathcal{O}(N) \) memory requirement and \( \mathcal{O}(N \log N) \) computational cost per iteration, while the number of iterations and the computational cost are greatly reduced.

To investigate the properties of the proposed preconditioners, the following lemma is the key to prove the invertibility of \( P^{(m)} \) in Equation (3.3).

**Lemma 3.1.** All eigenvalues of \( s(A) \) fall inside the open disk

\[ \{ z \in \mathbb{C} : |z - a_{11}| < a_{11} \}, \tag{3.4} \]

and all the eigenvalues of \( s(A) \) are strictly positive for all \( N \).

**Proof.** First of all, because the matrix \( A \) is symmetric, then \( s(A) \) is also symmetric and its eigenvalues should be real. All the Gershgorin disk of the circulant matrix \( s(A) \) are centered at \( a_{11} \) with radius

\[ r_N = 2 \sum_{i=2}^{[\frac{N}{2}]} |a_{1,i}| < a_{11}. \tag{3.5} \]

The above inequality holds due to the expression of Equation (2.13), where the expression of \( a_{11} \) contains exactly the sum of \( 2a_{1,i} \) (\( \ell = 2, 3, \ldots, N-1 \)). In conclusion, all the eigenvalues of \( s(A) \) are strictly positive for all \( N \). \( \square \)

According to Lemma 3.1, it means that \( s(A) \) is a real symmetric positive definite matrix. Moreover, the invertibility of circulant preconditioners \( P^{(m)} \) (3.3) can be given for all \( m = 1, 2, \ldots, M \) as follows.

**Lemma 3.2.** Let \( a \in (0, 2) \). The preconditioner \( P^{(m)} \) is invertible and

\[ \left\| (P^{(m)})^{-1} \right\|_2 \leq \frac{1}{\min_{1 \leq k \leq N-1} [\Lambda]_{k,k}} \cdot \frac{1}{\int b_{m}^{(m,\gamma)} + \kappa^{(m)} \Lambda} \tag{3.6} \]

**Proof.** According to Lemma 3.1, we have \( [\Lambda]_{k,k} > 0 \). Noting that \( a_{m}^{(m,\gamma)} > 0 \) (or \( b_{m}^{(m,\gamma)} > 0 \)) and \( \kappa^{(m)} > 0 \), we have

\[ \left[ \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,\gamma)} + \kappa^{(m)} \Lambda \right]_{k,k} > 0 \tag{3.7} \]

for \( k = 1, 2, \ldots, N-1 \). Therefore, \( P^{(m)} \) is invertible. Furthermore, we have

\[ \left\| (P^{(m)})^{-1} \right\|_2 \leq \min_{1 \leq k \leq N-1} \left[ \frac{1}{\Gamma(1-\gamma)} b_{m}^{(m,\gamma)} + \kappa^{(m)} \Lambda \right]_{k,k} \cdot \frac{1}{\int b_{m}^{(m,\gamma)} + \kappa^{(m)} \Lambda} \cdot \min_{1 \leq k \leq N-1} [\Lambda]_{k,k}. \]

The other inequality can be similarly obtained. \( \square \)
3.2 Spectrum of the preconditioned matrix

In this subsection, we study the spectrum of the preconditioned matrix, which can help us to understand the convergence of preconditioned Krylov subspace solvers. For convenience of our investigation, we first assume that the diffusion coefficient function \( \kappa(x,t) \equiv \kappa(t) \), then Equation (2.12) or Equation (3.1) will be a sequence of real symmetric positive definite linear systems, where the coefficient matrices reduce to \( \mathcal{M}^{(m)} = \frac{1}{1-\gamma} b^{(m)} \mathcal{A}^m + \kappa^{(m)} \mathcal{A} \) corresponding to Equations (2.12) and (3.1), respectively. The preconditioned CG (PCG) method\(^{51}\) should be a suitable candidate for solving such linear systems one by one. Moreover, the spectrum of the preconditioned matrix \((\mathcal{P}^{(m)})^{-1} \mathcal{M}^{(m)}\) is available for both Equations (2.12) and (3.1) at each time level \( m \), so we take Equation (2.12) as the research object in the next context.

Throughout this subsection, we rewrite Equation (2.12) into the following equivalent form:

\[
\tilde{\mathcal{M}}^{(m)} \mathbf{u}^m = \frac{1}{1-\gamma} C_{\alpha,\mu}^\gamma \left[ \sum_{k=1}^{m-1} \left( b_{k+1}^{(m)} - b_k^{(m)} \right) \mathbf{u}^k + b_1^{(m)} \mathbf{u}^0 \right] + \frac{1}{C_{\alpha,\mu}} \mathbf{f}^m, \tag{3.8}
\]

where \( \tilde{\mathcal{M}}^{(m)} = \frac{b_{m}}{1-\gamma} \mathcal{A}^m + \kappa^{(m)} \mathcal{A} \) and its corresponding circulant preconditioner \( \mathcal{P}^{(m)} \) reduces to \( \hat{\mathcal{P}}^{(m)} = \frac{b_{m}}{1-\gamma} \mathcal{I} + \kappa^{(m)} \mathcal{A} \), which is still invertible (cf. Lemma 3.2). Moreover, we assume that \( M \) and \( r \) are properly chosen, depending on \( N \), such that \( \tilde{\eta}_{N,M,r}^{(m)} = \frac{b_{m}}{1-\gamma} \) in (3.8) is bounded away from 0; that is, there exist two real numbers “\( \tilde{\eta} \)” and “\( \tilde{\eta} \)” such that

\[
0 < \tilde{\eta} \leq \tilde{\eta}_{N,M,r}^{(m)} \leq \hat{\tilde{\eta}}, \quad \forall N \text{ and } m = 1, 2, \ldots, M - 1. \tag{3.9}
\]

We add a subscript \( N \) to each matrix to denote the matrix size. Under the above assumption in (3.9), the matrix \( \tilde{\mathcal{M}}^{(m)} \), \( K^{(m)}, \tilde{\eta}_{N,M,r}^{(m)} \), and \( \hat{\mathcal{P}}^{(m)} \) are independent of \( m \), and we therefore simply denote them as \( \tilde{\mathcal{M}}_{N-1}, \kappa \) (constant), \( \tilde{\eta}_{N,M,r} \) (constant), and \( \hat{\mathcal{P}}_{N-1} \), respectively. Now the coefficient matrix \( \tilde{\mathcal{M}}^{(m)} \) in (2.12) becomes

\[
\tilde{\mathcal{M}}_{N-1} = \tilde{\eta}_{N,M,r} I + \kappa \tilde{\mathcal{A}}_{N-1} \nonumber
\]

\[
= \left[ \begin{array}{cccc}
\phi & \kappa \tilde{a}_{12} & \cdots & \kappa \tilde{a}_{1,N-2} & \kappa \tilde{a}_{1,N-1} \\
\kappa \tilde{a}_{12} & \phi & \kappa \tilde{a}_{12} & \cdots & \kappa \tilde{a}_{1,N-2} \\
\vdots & \kappa \tilde{a}_{12} & \phi & \ddots & \vdots \\
\kappa \tilde{a}_{1,N-2} & \cdots & \ddots & \ddots & \kappa \tilde{a}_{12} \\
\kappa \tilde{a}_{1,N-1} & \kappa \tilde{a}_{1,N-2} & \cdots & \kappa \tilde{a}_{12} & \phi
\end{array} \right] \tag{3.10}
\]

where we set \( |\tilde{a}_{12}| < \phi < \tilde{a}_{11} \) (without loss of generality), \( G_{N-1} = [\tilde{g}_{ij}]_{|N-1| \times |N-1|} = [\tilde{g}_{ij}]_{|N-1| \times |N-1|} \), and \( \tilde{g}_{ij} = \kappa \tilde{a}_{ij} \).

To study the spectrum of the preconditioned matrix \((\mathcal{P}^{(m)})^{-1} \tilde{\mathcal{M}}^{(m)}\), we first introduce the generating function of the sequence of Toeplitz matrices \( \{G_N\}_{N=1}^\infty \) :

\[
p(\theta) = \sum_{k=-\infty}^{\infty} \tilde{g}_k e^{ik\theta}, \tag{3.11}
\]

where \( \tilde{g}_k \) is the \( k \)th diagonal of \( G_N = [\tilde{g}_{i-j}]_{N \times N} \) and \( i = \sqrt{-1} \). The generating function \( p(\theta) \) is in the Wiener class\(^{50,51}\) if and only if

\[
\sum_{k=-\infty}^{\infty} |\tilde{g}_k| < \infty. \tag{3.12}
\]

For \( G_{N-1} \) defined in (3.8), we have the following conclusion.

**Lemma 3.3.** Under the above assumptions, it finds that \( p(\theta) \) is real valued and in the Wiener class.

**Proof.** For convenience of our investigation, we can rewrite \( p(\theta) \) for the matrix \( G_N \) defined in (3.8) as

\[
p(\theta) = \tilde{g}_0 + 2 \sum_{k=1}^{\infty} \tilde{g}_k \cos(k\theta) = \rho - 2 \sum_{k=1}^{\infty} (-\tilde{g}_k) \cos(k\theta). \tag{3.13}
\]
With the help of Lemma 3.4, we note that
\[
\begin{align*}
\sum_{k=1}^{\infty} (-\tilde{g}_k) \cos(k\theta) &= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right) \\
&= \frac{\sin(\theta/2)}{\sin(\theta/2)} \left(1 - \frac{1}{\pi} \theta \right)
\end{align*}
\]

with the series \(\sum_{k=1}^{\infty} (-\tilde{g}_k) \cos(k\theta)\) being convergent, the series \(\sum_{k=-\infty}^{\infty} (-\tilde{g}_k) \cos(k\theta)\) converges to a real-valued function for \(\forall \theta \in [-\pi, \pi]\), which also implies that \(p(\theta)\) is real valued. According to Proposition 2.1 and its proof, it is not hard to note that \(\lim_{k \to \infty} |\tilde{g}_k| = 0\). Therefore, it follows that \(\sum_{k=-\infty}^{\infty} |\tilde{g}_k| < \infty\), which completes the proof.

In fact, Lemma 3.3 ensures the following property that the given Toeplitz matrix \(G_N\) can be approximated via a circulant matrix well.

**Lemma 3.4.** If \(p(\theta)\), the generating function of \(G_N\), is in the Wiener class, then for any \(\epsilon > 0\), there exist \(N'\) and \(M' > 0\) such that for all \(N > N'\),
\[
G_N - s(G_N) = U_N + V_N, \quad (3.14)
\]
where \(\text{rank}(U_N) \leq M'\) and \(\|V_N\|_2 < \epsilon\).

Now we consider the spectrum of \((\hat{P}_{N-1})^{-1} \hat{M}_{N-1} - I\) is clustered around 1.

**Theorem 3.1.** If \(\eta_{N,M,r}\) satisfies the assumption (3.9), for any \(0 < \epsilon < 1\), there exists \(N'\) and \(M' > 0\) such that, for all \(N > N'\), at most \(2M'\) eigenvalues of the matrix \(\hat{M}_{N-1} - \hat{P}_{N-1}\) have absolute values exceeding \(\epsilon\).

**Proof.** With the help of Lemma 3.4, we note that
\[
\hat{M}_{N-1} - \hat{P}_{N-1} = \kappa \hat{A} - \kappa s(\hat{A})
= G_N - s(G_N) = U_N + V_N, \quad (3.15)
\]

Because both \(V_{N-1}\) and \(U_{N-1}\) are real symmetric with \(\|V_{N-1}\|_2 < \epsilon\) and \(\text{rank}(U_{N-1}) < M'\), hence the spectrum of \(V_{N-1}\) lies in \((- \epsilon, \epsilon)\). By the celebrated Weyl's theorem, we see that at most \(2M'\) eigenvalues of \(\hat{M}_{N-1} - \hat{P}_{N-1}\) have absolute values exceeding \(\epsilon\).

At this stage, we can see from Lemma 3.2 that
\[
\|((\hat{P}_{N-1})^{-1})_2\| = \min_{1 \leq k \leq N-1} \left| \frac{1}{\eta_{N,M,r} + \hat{A}} \right|_{k,k} < \frac{1}{\eta_{N,M,r}} < \frac{1}{\eta}, \quad (3.16)
\]
where $s(\kappa \tilde{A}) = F^* \tilde{A} F$ and the diagonal matrix $\tilde{A}$ contains all the eigenvalues of $s(\kappa \tilde{A})$. Meanwhile, we employ the fact that

$$
(\tilde{P}_{N-1})^{-1} \tilde{M}_{N-1} - I = (\tilde{P}_{N-1})^{-1} U_{N-1} - (\tilde{P}_{N-1})^{-1} V_{N-1}.
$$

(3.17)

then we have the following corollary.

**Corollary 3.1.** If $\eta_{NM'}$ satisfies the assumption (3.9), for any $0 < \epsilon < 1$, there exists $N'$ and $M' > 0$ such that, for all $N > N'$, at most $2M'$ eigenvalues of the matrix $(\tilde{P}_{N-1})^{-1} \tilde{M}_{N-1} - I$ have absolute values exceeding $\epsilon$.

Thus, the spectrum of $(\tilde{P}_{N-1})^{-1} \tilde{M}_{N-1}$ is clustered around 1 for $N$ that is large enough. It follows that the convergence rate of the PCG method is superlinear; refer to Chan and Ng\(^{50}\) and Ng\(^{51}\) for details. Based on such observations, the preconditioner $P^{(m)}$ is fairly predictable to accelerating the convergence of PCG for solving both Equations (2.12) and (3.1) at each time level $m = 1, 2, \ldots, M$ well, respectively; refer to numerical results in the next section.

Besides, although the theoretical analysis in Section 3.2 is only available for handling the model problem (1.1) with time-varying diffusion coefficients, that is, $\kappa(\chi, t) \equiv \kappa(t)$, the preconditioner $P^{(m)}$ is still efficient to accelerate the convergence of nonsymmetric Krylov subspace solvers for Equation (2.12) and/or Equation (3.1) corresponding to the problem (1.1). The variable diffusion coefficients and nonsymmetric discretized linear systems make it greatly challenging to theoretically study the eigenvalue distributions of preconditioned matrices $(P^{(m)})^{-1} M^{(m)}$, but we provide numerical results.
to show the clustering eigenvalue distributions of some specified preconditioned matrices in Section 4. In summary, we can analyze the computational complexity and memory requirement for both FIDS and DIDS as follows.

**Proposition 3.1.** The FIDS (or DIDS) has $O(NN_{\text{exp}})$ (or $O(NM)$) memory requirement and $O(MN(\log N + N_{\text{exp}}))$ (or $O(MN(\log N + M))$) computational complexity.

### 4 Numerical Experiments

In this section, numerical experiments are presented to achieve our twofold objective. We show that the proposed FIDS and DIDS can indeed converge with the theoretical accuracy in both space and time. Meanwhile, we assess the computational efficiency and theoretical results on circulant preconditioners described in Section 3. For the Krylov subspace method and direct solver, we exploit built-in functions for the preconditioned BiCGSTAB (PBiCGSTAB) method in MATLAB's backslash in Examples 1 and 2, respectively. For the BiCGSTAB method with circulant preconditioners, the stopping criterion of those methods is $||r^{(k)}||_2/||r^{(0)}||_2 < tol = 10^{-10}$, where $r^{(k)}$ is the residual vector of the linear system after $k$ iterations, and the initial guess is chosen as the zero vector. All experiments were performed

---

1 Here, we note that the MATLAB codes of all the numerical tests are available from the authors' emails and will make them public in the GitHub repository: https://github.com/Hsien-Ming-Ku/Group-of-FDEs.
TABLE 6 The $L_\infty$- and $L_2$-norm of errors, temporal convergence orders for solving Example 1 with $\alpha = 0.4$, $N(M) = \lfloor 2M^{\min\{\gamma_r - 2, 2-\gamma_r\}/2} \rfloor$, $\mu = 2$, $\kappa_\mu = 2$, and $s = 3$

| $(r, \gamma)$ | $M$ | $\text{DIDS (2.11)}$ | $\text{FIDS (2.10)}$ |
|---------------|-----|----------------|----------------|
|               | $\text{Err}_{\infty}$ | Rate | $\text{Err}_2$ | Rate | $\text{CPU(s)}$ | $\text{Err}_{\infty}$ | Rate | $\text{Err}_2$ | Rate | $\text{CPU(s)}$ |
| $(1, 0.8)$    | $2^8$ | 1.829e-3 | - | 1.468e-3 | - | 0.061 | 1.829e-3 | - | 1.468e-3 | - | 0.046 |
|               | $2^9$ | 1.070e-3 | 0.774 | 8.633e-4 | 0.766 | 0.179 | 1.070e-3 | 0.774 | 8.633e-4 | 0.766 | 0.094 |
|               | $2^{10}$ | 6.206e-4 | 0.785 | 5.027e-4 | 0.780 | 0.592 | 6.206e-4 | 0.785 | 5.027e-4 | 0.780 | 0.199 |
|               | $2^{11}$ | 3.589e-4 | 0.790 | 2.911e-4 | 0.788 | 2.667 | 3.589e-4 | 0.790 | 2.911e-4 | 0.788 | 0.553 |
| $(2, 0.5)$    | $2^7$ | 1.657e-3 | - | 1.343e-3 | - | 0.022 | 1.657e-3 | - | 1.343e-3 | - | 0.033 |
|               | $2^8$ | 8.409e-4 | 0.978 | 6.829e-4 | 0.976 | 0.045 | 8.409e-4 | 0.978 | 6.829e-4 | 0.976 | 0.062 |
|               | $2^9$ | 4.233e-4 | 0.990 | 3.446e-4 | 0.987 | 0.183 | 4.233e-4 | 0.990 | 3.446e-4 | 0.987 | 0.151 |
|               | $2^{10}$ | 2.128e-4 | 0.992 | 1.731e-4 | 0.993 | 0.744 | 2.128e-4 | 0.992 | 1.731e-4 | 0.993 | 0.448 |
| $(3, 0.8)$    | $2^7$ | 4.758e-4 | - | 3.854e-4 | - | 0.029 | 4.758e-4 | - | 3.854e-4 | - | 0.042 |
|               | $2^8$ | 2.071e-4 | 1.200 | 1.687e-4 | 1.193 | 0.070 | 2.071e-4 | 1.200 | 1.687e-4 | 1.193 | 0.101 |
|               | $2^9$ | 8.963e-5 | 1.208 | 7.315e-5 | 1.205 | 0.183 | 8.963e-5 | 1.208 | 7.315e-5 | 1.205 | 0.268 |
|               | $2^{10}$ | 3.930e-5 | 1.190 | 3.212e-5 | 1.187 | 1.069 | 3.930e-5 | 1.190 | 3.212e-5 | 1.187 | 0.821 |

Abbreviations: DIDS, direct implicit difference scheme; FIDS, fast implicit difference scheme.

FIGURE 1 The spatial convergence order of two difference schemes for Example 1 with $\alpha = 1.6$, $M(N) = \lfloor (N/2)^{\min\{\gamma_r - 2, 2-\gamma_r\}/2} \rfloor$, $\mu = 2$, $\kappa_\mu = 2$, and $s = 3$. Left: direct implicit difference scheme (DIDS) (2.11); right: fast implicit difference scheme (FIDS) (2.10) [Colour figure can be viewed at wileyonlinelibrary.com]

FIGURE 2 The spatial convergence order of two difference schemes for Example 1 with $\alpha = 0.4$, $M(N) = \lfloor (N/2)^{\min\{\gamma_r - 2, 2-\gamma_r\}/2} \rfloor$, $\mu = 2$, $\kappa_\mu = 2$, and $s = 3$. Left: direct implicit difference scheme (DIDS) (2.11); right: fast implicit difference scheme (FIDS) (2.10) [Colour figure can be viewed at wileyonlinelibrary.com]

on a Windows 10 (64 bit) PC-Intel(R) Core(TM) i5-8265U CPU (1.6–3.9 GHz), 8 GB of RAM, using MATLAB 2017b with machine epsilon $10^{-16}$ in double precision floating point arithmetic. The computing time reported is an average over 20 runs of our algorithms. We also choose the tolerance error $\varepsilon = 10^{-10}$, $10^{-9}$ for FIDS in Examples 1 and 2, respectively. The use of bold emphasis in the following tables means that the elapsed CPU time is the least. Moreover, some notations on numerical errors are introduced as follows:

$$\text{Error}_{\infty}(N, M) = \max_{0 \leq j \leq M} \| E_j \|_{\infty} \quad \text{and} \quad \text{Error}_2(N, M) = \max_{0 \leq j \leq M} \| E_j \|_2.$$
\[ \text{Example 1 (Accuracy test). In this example, we consider Equation (1.1) with the spatial domain } \Omega = (-1, 1) \text{ and the time interval } [0, T] = [0, 1]. \text{ The diffusion coefficients } \kappa(x, t) = (1 + t) e^{0.8x + 1} \text{ and the source term is given} \]

\begin{align*}
  f(x, t) &= \Gamma(1 + \gamma)(1 - x^2)^{s + \alpha/2} + \kappa(x, t) \frac{2^\gamma \Gamma\left(\frac{s + 1}{2}\right) \Gamma(s + 1 + \alpha/2)}{\sqrt{\pi} \Gamma(s + 1)} \times \\
  &\quad \mathbf{2F1}\left(\frac{\alpha + 1}{2}, -s; \frac{1}{2}; x^2\right)(t^{\gamma/2} + 1), \quad s \in \mathbb{N}^+,
\end{align*}

then

\[
\text{Rate}_\infty = \begin{cases} 
\log_2 \left( \frac{\text{Error}_{\infty}(N(M/2), M/2)}{\text{Error}_{\infty}(N(M), M)} \right), & \text{(temporal convergence order),} \\
\log_2 \left( \frac{\text{Error}_{\infty}(N/2, M/2)}{\text{Error}_{\infty}(N, M)} \right), & \text{(spatial convergence order)}.
\end{cases}
\]

and

\[
\text{Rate}_2 = \begin{cases} 
\log_2 \left( \frac{\text{Error}_{\infty}(N(M/2), M/2)}{\text{Error}_{\infty}(N(M), M)} \right), & \text{(temporal convergence order),} \\
\log_2 \left( \frac{\text{Error}_{\infty}(N/2, M/2)}{\text{Error}_{\infty}(N, M)} \right), & \text{(spatial convergence order)}.
\end{cases}
\]
TABLE 9  The $L_{\infty}$- and $L_{1}$-norm of errors, spatial convergence orders for solving Example 2 with $a = 1.9$, $M(N) = [(N/2)^{\mu / \min (r,2-r)}]$, $\mu = 1 + a/2$, and $\kappa = 1$

| $(r,s)$ | $N$ | DIDS (2.11) | FIDS (2.10) |
|---------|-----|-------------|-------------|
|         | $N$ | $\text{Err}_{\infty}$ | Rate | $\text{Err}_{1}$ | Rate | $\text{CPU(s)}$ | $\text{Err}_{\infty}$ | Rate | $\text{Err}_{1}$ | Rate | $\text{CPU(s)}$ |
| $(1, 0.5)$ | 9  | 7.770e-2 | - | 8.304e-2 | - | 0.006 | 7.770e-2 | - | 8.304e-2 | - | 0.007 |
| $18$ | - | 2.010 | 2.052e-2 | 2.017 | 0.035 | 1.929e-2 | 2.010 | 2.052e-2 | 2.017 | 0.033 |
| $36$ | - | 4.719e-3 | 5.045e-3 | 2.024 | 0.834 | 4.719e-3 | 2.031 | 5.045e-3 | 2.024 | 0.254 |
| $72$ | - | 1.153e-3 | 1.238e-3 | 2.026 | 2.551 | 1.153e-3 | 2.033 | 1.238e-3 | 2.026 | 1.927 |
| $(2, 0.5)$ | 9  | 7.663e-2 | - | 8.193e-2 | - | 0.002 | 7.663e-2 | - | 8.193e-2 | - | 0.005 |
| $18$ | - | 1.903e-2 | 2.010 | 2.026e-2 | 2.016 | 0.008 | 1.903e-2 | 2.010 | 2.026e-2 | 2.016 | 0.014 |
| $36$ | - | 4.657e-3 | 4.983e-3 | 2.023 | 0.053 | 4.657e-3 | 2.031 | 4.983e-3 | 2.023 | 0.077 |
| $72$ | - | 1.138e-3 | 1.223e-3 | 2.026 | 0.879 | 1.138e-3 | 2.033 | 1.223e-3 | 2.026 | 0.426 |
| $(3, 0.5)$ | 9  | 7.663e-2 | - | 8.195e-2 | - | 0.002 | 7.663e-2 | - | 8.195e-2 | - | 0.004 |
| $18$ | - | 1.902e-2 | 2.011 | 2.025e-2 | 2.017 | 0.005 | 1.902e-2 | 2.011 | 2.025e-2 | 2.017 | 0.008 |
| $36$ | - | 4.650e-3 | 4.977e-3 | 2.025 | 0.016 | 4.650e-3 | 2.032 | 4.977e-3 | 2.025 | 0.033 |
| $72$ | - | 1.135e-3 | 1.221e-3 | 2.028 | 0.114 | 1.135e-3 | 2.034 | 1.221e-3 | 2.028 | 0.162 |

Abbreviations: DIDS, direct implicit difference scheme; FIDS, fast implicit difference scheme.

TABLE 10  The $L_{\infty}$- and $L_{1}$-norm of errors, spatial convergence orders for solving Example 2 with $a = 0.5$, $M(N) = [(N/2)^{\mu / \min (r,2-r)}]$, $\mu = 1 + a/2$, and $\kappa = 1$

| $(r,s)$ | $N$ | DIDS (2.11) | FIDS (2.10) |
|---------|-----|-------------|-------------|
|         | $N$ | $\text{Err}_{\infty}$ | Rate | $\text{Err}_{1}$ | Rate | $\text{CPU(s)}$ | $\text{Err}_{\infty}$ | Rate | $\text{Err}_{1}$ | Rate | $\text{CPU(s)}$ |
| $(1, 0.8)$ | 2 | 1.871e-3 | - | 1.871e-3 | - | 0.065 | 1.871e-3 | - | 1.871e-3 | - | 0.063 |
| $2$ | - | 8.355e-4 | 1.163 | 8.323e-4 | 1.169 | 0.539 | 8.355e-4 | 1.163 | 8.323e-4 | 1.169 | 0.364 |
| $4$ | - | 3.642e-4 | 1.198 | 3.616e-4 | 1.203 | 0.535 | 3.642e-4 | 1.198 | 3.616e-4 | 1.203 | 2.438 |
| $8$ | - | 1.556e-4 | 1.227 | 1.543e-4 | 1.229 | 0.604 | 1.556e-4 | 1.227 | 1.543e-4 | 1.229 | 51.194 |
| $(2, 0.5)$ | 2 | 2.666e-3 | - | 2.654e-3 | - | 0.013 | 2.666e-3 | - | 2.654e-3 | - | 0.023 |
| $2$ | - | 1.157e-3 | 1.204 | 1.149e-3 | 1.208 | 0.737 | 1.157e-3 | 1.204 | 1.149e-3 | 1.208 | 0.091 |
| $4$ | - | 4.970e-4 | 1.219 | 4.933e-4 | 1.219 | 0.584 | 4.970e-4 | 1.219 | 4.933e-4 | 1.219 | 0.571 |
| $8$ | - | 2.115e-4 | 1.232 | 2.098e-4 | 1.234 | 0.585 | 2.115e-4 | 1.232 | 2.098e-4 | 1.234 | 9.277 |
| $(3, 0.5)$ | 2 | 2.581e-3 | - | 2.739e-3 | - | 0.003 | 2.581e-3 | - | 2.739e-3 | - | 0.006 |
| $2$ | - | 1.114e-3 | 1.211 | 1.169e-3 | 1.228 | 0.007 | 1.114e-3 | 1.211 | 1.169e-3 | 1.228 | 0.014 |
| $4$ | - | 4.677e-4 | 1.253 | 4.874e-4 | 1.262 | 0.025 | 4.677e-4 | 1.253 | 4.874e-4 | 1.262 | 0.046 |
| $8$ | - | 2.000e-4 | 1.226 | 2.075e-4 | 1.232 | 0.161 | 2.000e-4 | 1.226 | 2.075e-4 | 1.232 | 0.232 |

Abbreviations: DIDS, direct implicit difference scheme; FIDS, fast implicit difference scheme.

where $\text{F}_1(\cdot)$ denotes the Gauss hypergeometric function, which can be computed via the MATLAB built-in function “hypergeom.m” and the initial-boundary value conditions are

$$u(x, t) = 0, \; x \notin \Omega, \quad \text{and} \quad u(x, 0) = (1 - x^2)^{s+a/2}.$$  

Thus, the exact solution of this problem is $u(x, t) = (1 - x^2)^{s+a/2}(e^t + 1)$. The numerical results involving both spatial and temporal convergence orders of FIDS (2.10) and DIDS (2.11) are shown in Tables 1–6 and Figures 1 and 2. Here, it should mentioned that we only use the direct method for solving the resultant linear systems of FIDS (2.10) and DIDS (2.11), respectively, because the maximal size of such resultant linear systems is still smaller than 128 and the superiority of Krylov subspace solvers with circulant preconditioners are slightly less remarkable compared with the direct solvers in terms of the elapsed CPU time; see, for example, Gu et al$^{29}$ and Pang and Sun$^{59}$ and the context in the next example for a discussion.

Tables 1–4 present the numerical errors, CPU time (in seconds), and spatial/temporal convergence rates of both FIDS and DIDS for solving the problem (1.1), which satisfy the smooth condition mentioned in Remark 2.1. When we refine the discretized grid size, it is easily seen that for the temporal direction, the numerical convergence order is consistent with the theoretical estimate $O(M^{-\min (r,2-r)})$ for different $a$. Meanwhile, it can find that the numerically spatial convergence...
order is exactly consistent with the theoretical estimation $O(h^2)$ for different orders of the IFL. In addition, the results of CPU time demonstrate that the FIDS outperforms the DIDS, especially when the integer $M$ is increasingly large.

On the other hand, there is another splitting parameter $\mu = 2$ for discretizing the IFL, and it makes the spatial discretization of IFL enjoy the second-order accuracy. According to Tables 5 and 6 and Figures 1 and 2, it is not hard to find that both FIDS and DIDS under such a spatial discretization for solving Example 1 can reach still the spatial convergence order $O(h^2)$ and $O(M^{-\min\{r,2-\gamma\}})$ with different settings. Thus, such results are also consistent with the theoretical estimate described in Section 2.2. Again, the results of the elapsed CPU time show that the FIDS consumes much less time than the DIDS, especially when the integer $M$ is very large.

**Example 2.** The second example is similar to the setting in Example 1, while we choose $s = 1$ and $\kappa(\chi,t) = 7[\ln(5 + 2\chi + t) + \cos(\chi t)]/4$. Then the exact solution $u(x,t)$ and source term $f(x,t)$ can be computed via the form described in Example 1, and the corresponding initial-boundary conditions are also similarly obtained.

In this example, it notes that the exact solution $u(x, \cdot) \in C^{1+\frac{\gamma}{2}}(\mathbb{R})$ satisfies the less smoother condition than that in Example 1. Moreover, it is seen from Tables 7–10 that for the temporal direction, the numerical convergence rate of both FIDS and DIDS is consistent with the theoretical estimate $O(M^{-\min\{r,2-\gamma\}})$ for different settings. However, it remarks that the spatial convergence rate of both FIDS and DIDS can at least approach to $1 + \frac{\alpha}{2}$, especially when $\alpha$ increasingly goes to 2, and the spatial convergence orders of both FIDS and DIDS are almost 2. These results on spatial convergence rate of both FIDS and DIDS are fairly better than the theoretical estimate in Remark 2.1. It implies that the error analysis and smooth condition of the numerical discretization of IFL used to establish the DIDS and FIDS can be further sharpened and weakened, respectively. Analogously, the average CPU time of FIDS is smaller than that of DIDS for problem (1.1), when the number of time levels is increasingly large.

On the other hand, Table 11 and Figures 3 and 4 are carried out to show the effectiveness of the proposed circulant preconditioners, which is especially useful for the order $\alpha (\rightarrow 2)$ of IFL; refer to Figures 3 and 4 as well. For Figure 3A, it is implied that if we increase $N$, then the number of time level will be too huge to make a concise comparison of FIDS and DIDS (with no/circulant preconditioners). Moreover, due to the large number $M$, the family of FIDS should be more efficient than the counterparts of DIDS for solving the problem (1.1). As seen from Table 11, the proposed circulant preconditioner is efficient to accelerate the implementation of both FIDS and DIDS in terms of the reduction of “Its” and “CPU,” especially for large integers $M$ and $N$. This observation can be also supported by the clustering eigenvalue distributions shown in Figures 3 and 4. Moreover, the number of iterations of “DIDS + $P$” and “FIDS + $P$” is roughly independent of decreasing spatial grid size. The above results of circulant preconditioners are exactly consistent with the theoretical investigations given in Section 3. In a word, the “FIDS + $P$” is the most promising numerical method for solving the problem (1.1), especially with large integers $M, N, \text{and } M > N$.

### Abbreviations

- DIDS: Direct implicit difference scheme
- FIDS: Fast implicit difference scheme

### Table 11

| $(r, \gamma, \alpha)$ | $N$ | $F_{\text{err}}$ | $F_{\text{err}a}$ | DIDS | DIDS + $I$ | DIDS + $P$ | FIDS | FIDS + $I$ | FIDS + $P$ | CPU(s) | Its | CPU(s) | Its | CPU(s) | Its | CPU(s) | Its |
|----------------------|-----|----------------|------------------|------|-----------|-----------|------|-----------|-----------|--------|-----|--------|-----|--------|-----|--------|-----|
| $(2, 0.5, 1)$        | $2^6$ | 3.632e-4 | 7.727e-4 | 0.162 | 0.156 | 44.1 | 0.658 | 7.6 | 0.332 | 44.1 | 0.655 | 7.7 | 0.347 |
|                      | $2^7$ | 1.542e-4 | 1.889e-4 | 1.793 | 1.014 | 66.8 | 5.401 | 7.9 | 2.380 | 66.7 | 4.693 | 7.9 | 1.689 |
|                      | $2^8$ | 4.469e-5 | 4.643e-5 | 22.217 | 92.7 | 45.111 | 8.0 | 20.777 | 92.7 | 30.132 | 8.0 | $6.989$ |
|                      | $2^9$ | 1.333e-5 | 1.236e-5 | 413.96 | 125.7 | 570.75 | 8.2 | 285.50 | 125.7 | 306.26 | 8.2 | $43.338$ |

| $(3, 0.8, 1)$        | $2^7$ | 1.381e-4 | 1.734e-4 | 0.286 | 0.327 | 41.5 | 0.935 | 6.6 | 0.446 | 41.6 | 0.900 | 6.6 | 0.439 |
|                      | $2^8$ | 3.418e-5 | 4.146e-5 | 2.142 | 48.6 | 4.479 | 6.9 | 21.04 | 48.6 | 4.091 | 6.9 | $1.545$ |
|                      | $2^9$ | 8.979e-6 | 9.699e-6 | 0.293 | 5.401 | 0.935 | 7.0 | 16.488 | 5.39 | 26.938 | 7.1 | $1.813$ |
|                      | $2^{10}$ | 2.319e-6 | 2.396e-6 | 624.02 | 511.16 | 206.13 | 7.0 | 138.18 | 57.8 | 97.137 | 7.0 | $3.545$ |

| $(2, 0.5, 1.9)$      | $2^7$ | 1.446e-3 | 1.553e-3 | 0.507 | 0.386 | 65.6 | 2.123 | 7.9 | 0.853 | 65.6 | 1.879 | 7.9 | 0.748 |
|                      | $2^8$ | 3.353e-4 | 3.811e-4 | 7.831 | 2.347 | 116.4 | 22.072 | 7.9 | 8.574 | 116.4 | 17.458 | 7.9 | 3.838 |
|                      | $2^9$ | 8.636e-5 | 9.354e-5 | 84.271 | 213.65 | 187.3 | 8.1 | 135.58 | 198.3 | 203.71 | 8.1 | $1.807$ |
|                      | $2^{10}$ | 2.112e-5 | 2.297e-5 | 48.229 | 1029.78 | 316.1 | 7.0 | 2474.12 | 316.3 | 2620.56 | 8.1 | $1.320$ |

| $(3, 0.8, 1.9)$      | $2^7$ | 3.521e-4 | 3.800e-4 | 0.935 | 0.689 | 73.8 | 3.317 | 7.3 | 1.201 | 73.9 | 3.222 | 7.3 | 1.049 |
|                      | $2^8$ | 8.596e-5 | 9.318e-5 | 10.468 | 5.134 | 100.3 | 20.579 | 7.3 | 7.464 | 100.2 | 17.025 | 7.3 | $1.049$ |
|                      | $2^9$ | 2.100e-5 | 2.286e-5 | 57.942 | 173.492 | 84.047 | 7.4 | 81.104 | 127.5 | 203.71 | 7.4 | $21.436$ |
|                      | $2^{10}$ | 5.136e-6 | 5.614e-6 | 243.12 | 1505.84 | 151.0 | 7.3 | 1003.78 | 151.0 | 74.72 | 7.3 | $86.540$ |

**Abbreviations:** DIDS, direct implicit difference scheme; FIDS, fast implicit difference scheme.
FIGURE 3 The brief performance analysis of fast implicit difference scheme (2.10) and direct implicit difference scheme (2.11) with no/circulant preconditioners for solving Example 2 with $N = 2^7$ in panels (b)-(d) [Colour figure can be viewed at wileyonlinelibrary.com]

(A) The number of time levels $M$ versus the number $N$ of spatial grid nodes for $(r, \gamma) = (1, 0.8)$.

(B) The eigenvalue distributions of original and preconditioned matrices when $\alpha = 1.5$.

(C) The eigenvalue distributions of original and preconditioned matrices when $\alpha = 1.1$.

(D) The eigenvalue distributions of original and preconditioned matrices when $\alpha = 1.0$.

FIGURE 4 Eigenvalue distributions of the original and preconditioned matrices from both fast implicit difference scheme (2.10) and direct implicit difference scheme (2.11) for Example 2 with different setting and $N = 2^7$ [Colour figure can be viewed at wileyonlinelibrary.com]
In this work, we proposed two fast and easy-to-implement IDSs (i.e., FIDS and DIDS) for solving the TSFDE (1.1) with nonsmooth initial data, which was not well studied in the previous work. Meanwhile, the solvability, stability, and convergence rate of the proposed IDSs with nonuniform temporal steps are rigorously proved via the matrix properties, which are meticulously derived from the direct discretization of IFL. Numerical results in Section 4 are reported to support our theoretical findings. In addition, although the focus is on the one-dimensional spatial domain in this work, we note that the proposed methods utilizing spatial discretizations can be directly adapted and corresponding results remain valid for two- and three-dimensional cases, which will be precisely presented in our another coming manuscript.

On the other hand, due to the nonlocality of Caputo fractional derivative in the TSFDE (1.1), the numerical scheme needs to repeat the weighted sum of solutions at previous time levels. In order to reduce the computational cost, we exploit the fast SOE approximation of graded L1 formula to result in the FIDS, which is cheaper than the DIDS, especially for large integer $M$. However, they both need to solve the dense discretized systems, which is still time consuming. It implies that the efficient implementation of FIDS and DIDS should be further considered. With the help of Toeplitz-like matrix, we construct the BiCGSTAB with circulant preconditioners for solving the series of discretized linear systems (cf. Equations 2.12 and 3.1) without storing any matrices. It makes the FIDS (or DIDS) only require $O(NN_{exp})$ (or $O(NM)$) memory requirement and $O(MN(\log N + N_{exp}))$ (or $O(MN(\log N + M))$) computational complexity. It finds that the circulant preconditioners are efficient, because we can prove that the eigenvalues of preconditioned matrices cluster around 1, except for few outliers. The vast majority of these eigenvalues are well separated away from 0. It means that the BiCGSTAB with circulant preconditioners for solving the discretized linear systems can converge very fast. Numerical experiments are reported to show the effectiveness of the FIDS and DIDS with PBiCGSTAB solvers in terms of the elapsed CPU time and number of iterations, especially the former one.

Finally, it is meaningful to note that the numerically spatial convergence order of FIDS or DIDS is better than the theoretical estimate of FIDS or DIDS; see Example 2. It means that the error analysis of numerical schemes for solving time-dependent problems is different from the numerical IFL described in Duo et al. because their result is the error estimate for the IFL, but in the current work, it is the estimate for solutions of problem (1.1). The more refined error/convergence analysis of FIDS or DIDS is worth exploring in our future work; refer, for example, to previous works for a discussion. Our current work includes applying the FIDS and DIDS for solving the (nonlinear) multidimensional TSFDEs (on unbounded domains) with nonhomogeneous boundary conditions and designing the more efficient preconditioning techniques, such as $r$-algebra, multigrid, and banded preconditioners, for the corresponding two- and three-level Toeplitz discretized linear systems.

**ACKNOWLEDGEMENTS**

We are very grateful to the anonymous referees for their invaluable comments and insightful suggestions that have greatly improved the presentation of this paper. This research is supported by National Natural Science Foundation of China (11801463), the Applied Basic Research Program of Sichuan Province (2020YJ0007), the Fundamental Research Funds for the Central Universities (JBK1902028), the Science and Technology Development Fund, Macau SAR (file no. 0118/2018/A3), MYRG2018-00015-FST from University of Macau, and the U.S. National Science Foundation (DMS-1620465).

**CONFLICT OF INTEREST**

The authors declare that this work does not have any conflicts of interest.

**ORCID**

Xian-Ming Gu https://orcid.org/0000-0001-7895-2050

**REFERENCES**

1. Podlubny I. *Fractional Differential Equations*. New York, NY: Academic Press; 1998.
2. Sun HG, Zhang Y, Baleanu D, Chen W, Chen YQ. A new collection of real world applications of fractional calculus in science and engineering. *Commun Nonlinear Sci Numer Simul.* 2018;64:213–231.
3. Vázquez JL. The Mathematical Theories of Diffusion: Nonlinear and Fractional Diffusion. In: Bonforte M, ed. Nonlocal and Nonlinear Diffusions and Interactions: New Methods and Directions, Lecture Notes in Mathematics, vol. 2186. Cham, Switzerland: Springer; 2017:205–278. https://doi.org/10.1007/978-3-319-61494-6_5

4. Li C, Yi Q. Modeling and computing of convection diffusion equation. Commun Appl Math Comput. 2019;1(4):565–595.

5. Chen W, Sun H, Zhang X, Korosak D. Anomalous diffusion modeling by fractal and fractional derivatives. Comput Math Appl. 2010;59(5):1754–1758.

6. Shlesinger MF, West BJ, Klaver J. Lévy dynamics of enhanced diffusion: application to turbulence. Phys Rev Lett. 1987;58(11):1100–1103.

7. Estrada-Rodriguez G, Gimperlein H, Painter KJ, Stoeck J. Space-time fractional diffusion in cell movement models with delay. Math Models Meth Appl Sci. 2019;29(1):65–88.

8. Chi G, Li G, Sun C, Jia X. Numerical solution to the space-time fractional diffusion equation and inversion for the space-dependent diffusion coefficient. J Comput Theor Phys. 2017;46(2):122–146.

9. Du Q, Gunzburger M, Lehouc Q, Zhou K. Analysis and approximation of nonlocal diffusion problems with volume constraints. SIAM Rev. 2012;54(4):667–696.

10. Tian X, Du Q, Gunzburger M. Asymptotically compatible schemes for the approximation of fractional Laplacian and related nonlocal diffusion problems on bounded domains. Adv Comput Math. 2016;42(6):1363–1380.

11. Duo S, Wang H, Zhang Y. A comparative study on nonlocal diffusion operators related to the fractional Laplacian. Discrete Contin Dyn Syst Ser B. 2019;24(1):231–256.

12. del Teso F, Endal J, Jakobsen ER. Robust numerical methods for nonlocal (and local) equations of porous medium type Part II: Schemes and experiments. SIAM J Numer Anal. 2018;56(6):3611–3647.

13. Podlubny I, Chechkin A, Skovranek T, Chen YQ, Jarad BMV. Matrix approach to discrete fractional calculus II: partial fractional differential equations. J Comput Phys. 2009;228(8):3137–3153.

14. Yang Q, Liu F, Turner I. Numerical methods for fractional partial differential equations with Riesz space fractional derivatives. Appl Math Model. 2010;34(1):200–218.

15. Yang Q, Turner I, Liu F, Ilić M. Novel numerical methods for solving the time–space fractional diffusion equation in two dimensions. SIAM J Sci Comput. 2011;33(3):1159–1180.

16. Garbaczewski P, Stephanovich V. Fractional Laplacians in bounded domains: killed, reflected, censored, and taboo Lévy flights. Phys Rev E. 2019;99(4):42126. https://doi.org/10.1103/PhysRevE.99.042126

17. Guan Q-Y, Ma Z-M. Boundary problems for fractional Laplacians. Stoch Dynam. 2005;5(3):385–424.

18. Guan Q-Y, Ma Z-M. Reflected symmetric α-stable processes and regional fractional Laplacian. Probab Theory Rel. 2006;134(4):649–694.

19. Kolokoltsov VN, Veretennikova MA. Well-posedness and regularity of the Cauchy problem for nonlinear fractional in time and space equations. Fractional Differ Calc. 2014;4(1):1–30.

20. Hanyga A. Multi-dimensional solutions of space-time-fractional diffusion equations. Proc R Soc Lond A. 2002;458(2018):429–450. https://doi.org/10.1098/rspa.2001.0893

21. Padgett JL. The quenching of solutions to time–space fractional Kawarada problems. Comput Math Appl. 2018;76(7):1583–1592.

22. Li L, Liu J-G, Wang L. Cauchy problems for Keller–Segel type time–space fractional diffusion equation. J Differ Equ. 2018;265(3):1044–1096.

23. Jia J, Li K. Maximum principles for a time–space fractional convection-diffusion equation. Fract Calc Appl Anal. 2016;19(5-6):122–146.

24. Bonito A, Lei W, Pasciak JE. Numerical approximation of space-time fractional parabolic equations. Comput Methods Appl Math. 2017;17(4):679–705.

25. Acosta G, Beretsetche FM, Borthagaray JP. Finite element approximations for fractional evolution problems. Fract Calc Appl Anal. 2019;22(3):767–794.

26. Bonito A, Borthagaray JP, Nochetto RH, Otárola E, Salgado AJ. Numerical methods for fractional diffusion. Comput Vis Sci. 2018;19(5-6):19–46.

27. Arshad S, Huang J, Khaliq AQM, Tang Y. Trapezoidal scheme for time–space fractional diffusion equation with Riesz derivative. J Comput Phys. 2017;350:1–15.

28. Arshad S, Bu W, Huang J, Tang Y, Zhao Y. Finite difference method for time–space linear and nonlinear fractional diffusion equations. Int J Comput Math. 2018;95(1):202–217.

29. Gu Y-M, Huang T-Z, Ji C-C, Carpentieri B, Alikhanov AA. Fast iterative method with a second-order implicit difference scheme for time–space fractional convection-diffusion equation. J Sci Comput. 2017;72(3):957–985.

30. Feng LB, Zhuang P, Liu F, Turner I, Gu YT. Finite element method for space-time fractional diffusion equation. Numer Algorithms. 2016;72(3):749–767.

31. Yue XQ, Bu WP, Shu S, Liu MH, Wang S. Fully finite element adaptive AMG method for time–space Caputo–Riesz fractional diffusion equations. Adv Appl Math Mech. 2018;10(5):1103–1125.

32. Biala TA, Khaliq AQM. Parallel algorithms for nonlinear time–space fractional parabolic PDEs. J Comput Phys. 2018;375:135–154.

33. Duo S, Ju L, Zhang Y. A fast algorithm for solving the space–time fractional diffusion equation. Comput Math Appl. 2018;75(6):1929–1941.

34. Nochetto RH, Otárola E, Salgado AJ. A PDE approach to space–time fractional parabolic problems. SIAM J Numer Anal. 2016;54(2):848–873.

35. Hu Y, Li C, Li H. The finite difference method for Caputo-type parabolic equation with fractional Laplacian: one-dimension case. Chaos Soliton Fract. 2017;102:319–326.
36. Hu Y, Li C, Li H. The finite difference method for Caputo-type parabolic equation with fractional Laplacian: more than one space dimension. *Int J Comput Math.* 2018;95(6-7):1114–1130.
37. Huang Y, Oberman A. Numerical methods for the fractional Laplacian: a finite difference-quadrature approach. *SIAM J Numer Anal.* 2014;52(6):3056–3084.
38. Duo S, van Wyk HW, Zhang Y. A novel and accurate finite difference method for the fractional Laplacian and the fractional Poisson problem. *J Comput Phys.* 2018;355:233–252.
39. Duo S, Zhang Y. Accurate numerical methods for two and three dimensional integral fractional Laplacian with applications. *Comput Meth Appl Mech Eng.* 2019;355:639–662.
40. Hao Z, Zhang Z, Du R. Fractional centered difference scheme for high-dimensional integral fractional Laplacian. ResearchGate preprint. 11 Mar. 2020; 31 pages. Available online at: https://www.researchgate.net/publication/335888811
41. Bonito A, Lei W, Pasciak JE. Numerical approximation of the integral fractional Laplacian. *Numer Math.* 2019;142(2):235–278.
42. Acosta G, Borthagaray JP, Bruno O, Maas M. Regularity theory and high order numerical methods for the (1D)-fractional Laplacian. *Math Comp.* 2018;87(312):1821–1857.
43. Acosta G, Bersetche FM, Borthagaray JP. A short FE Implementation for a 2D homogeneous Dirichlet problem of a fractional Laplacian. *Comput Math Appl.* 2017;74(4):784–816.
44. Minden V, Ying L. A simple solver for the fractional Laplacian in multiple dimensions. *SIAM J Sci Comput.* 2020;42(2):A878–A900.
45. Ainsworth M, Glusa C. Hybrid finite element-spectral method for the fractional Laplacian: approximation theory and efficient solver. *SIAM J Sci Comput.* 2018;40(4):A2383–A2405.
46. Lin Y, Xu C. Finite difference/spectral approximations for the time-fractional diffusion equation. *J Comput Phys.* 2007;225(2):1533–1552.
47. Guan Q, Gunzburger M. \( \theta \) schemes for finite element discretization of the space-time fractional diffusion equations. *J Comput Appl Math.* 2015;288:264–273.
48. Liu Z, Cheng A, Li X, Wang H. A fast solution technique for finite element discretization of the space-time fractional diffusion equation. *Appl Numer Math.* 2017;119:146–163.
49. Chen A, Du Q, Li C, Zhou Z. Asymptotically compatible schemes for space-time nonlocal diffusion equations. *Chaos Soliton Fract.* 2017;102:361–371.
50. Chan RH, Ng MK. Conjugate gradient methods for Toeplitz systems. *SIAM Rev.* 1996;38(3):427–482.
51. Ng MK. *Iterative Methods for Toeplitz Systems.* New York, NY: Oxford University Press; 2004.
52. Lei S-L, Sun H-W. A circulant preconditioner for fractional diffusion equations. *J Comput Phys.* 2013;242:715–725.
53. Sakamoto K, Yamamoto M. Initial value/boundary value problems for fractional diffusion-wave equations and applications to some inverse problems. *J Math Anal Appl.* 2011;382(1):426–447.
54. Shen J-Y, Sun Z-Z, Du R. Fast finite difference schemes for time-fractional diffusion equations with a weak singularity at initial time. *East Asian J Appl Math.* 2018;8(4):834–858.
55. Fu H, Wang H. A preconditioned fast finite difference method for space-time fractional partial differential equations. *Fract Calc Appl Anal.* 2017;20(1):88–116.
56. Fu H, Ng MK, Wang H. A divide-and-conquer fast finite difference method for space-time fractional partial differential equation. *Comput Math Appl.* 2017;73(6):1233–1242.
57. Gu X-M, Wu S-L. A parallel-in-time iterative algorithm for Volterra partial integro-differential problems with weakly singular kernel. *J Comput Phys.* 2020;417:109576. https://doi.org/10.1016/j.jcp.2020.109576
58. Jiang S, Zhang J, Zhang Q, Zhang Z. Fast evaluation of the Caputo fractional derivative and its applications to fractional diffusion equations. *Commun Comput Phys.* 2017;21(3):650–678.
59. Pang H-K, Sun H-W. Multigrid method for fractional diffusion equations. *J Comput Phys.* 2018;231(2):693–703.
60. Horn RA, Johnson CR. *Matrix Analysis.* 2nd ed. New York, NY: Cambridge University Press; 2012.
61. Lin X-L, Ng MK. A fast solver for multidimensional time–space fractional diffusion equation with variable coefficients. *Comput Math Appl.* 2019;78(5):1477–1489.
62. van der Vorst HA. Bi-CGSTAB: A fast and smoothly converging variant of bi-CG for the solution of nonsymmetric linear systems. *SIAM J Sci Stat Comput.* 1992;13(2):631–644.

**How to cite this article:** Gu X-M, Sun H-W, Zhang Y, Zhao Y-L. Fast implicit difference schemes for time–space fractional diffusion equations with the integral fractional Laplacian. *Math Meth Appl Sci.* 2021;44:441-463. https://doi.org/10.1002/mma.6746