J Sci Comput manuscript No.
(will be inserted by the editor)

Fast numerical contour integral method for fractional diffusion equations

Hong-Kui Pang · Hai-Wei Sun

Received: date / Accepted: date

Abstract The numerical contour integral method with hyperbolic contour is exploited to solve space-fractional diffusion equations. By making use of the Toeplitz-like structure of spatial discretized matrices and the relevant properties, the regions that the spectra of resulting matrices lie in are derived. The resolvent norms of the resulting matrices are also shown to be bounded outside of the regions. Suitable parameters in the hyperbolic contour are selected based on these regions to solve the fractional diffusion equations. Numerical experiments are provided to demonstrate the efficiency of our contour integral methods.

Keywords Fractional diffusion equation · Numerical contour integral · Laplace transform · Hyperbolic contour · Toeplitz matrix

Mathematics Subject Classification (2000) 26A33 · 35K15 · 65F10 · 65M06 · 47B35

1 Introduction

Fractional diffusion equations have been studied extensively in recent years, since they can provide an adequate and accurate description of transport processes that exhibit anomalous diffusion, which cannot be modeled properly by second-order diffusion equations [26,38]. Applications of such equations have been found in a variety of fields, for instance, in hydrology [1], biology [16], physics [4], chemistry [7], finance [28] and so on. As most of fractional diffusion equations can not be solved analytically, numerical methods are major tools to investigate them and have been developed intensively [10,11,22,23,25,34,35,38–40,43].

H. Pang
School of Mathematics and Statistics, Jiangsu Normal University, Xuzhou 221116, Jiangsu, People's Republic of China.
E-mail: panghongkui@163.com

H. Sun
Corresponding author. Department of Mathematics, University of Macau, Macao, People's Republic of China.
E-mail: HSun@umac.mo

This paper concerns the space-fractional diffusion equations in which the spatial derivatives are fractional order. There many works have been done on this topic. Liu, Anh, and Turner [13] transformed the space-fractional Fokker-Planck equation into a system of ordinary differential equations and then solved them using backward differentiation formulas. Shen and Liu [30] proposed an explicit finite difference approximation for the space-fractional diffusion equation and gave an error analysis. Meerschaert, Tadjeran, and Scheffler, in their series of articles [22,23,34,35], approximated the space fractional derivatives by shifted Grünwald-Letnikov formulas and successfully got stable finite difference schemes. As the fractional derivatives are nonlocal operators, the resulting coefficient matrices generally tend to be full [23]. Therefore, using the Gaussian elimination method to solve the resulting linear systems often results in $\mathcal{O}(n^3)$ of computational work per time step and $\mathcal{O}(n^2)$ of memory requirement, where n is the number of spatial grid points. Wang, Wang, and Sircar [40] observed the Toeplitz-like structure of resulting coefficient matrices and designed a fast finite difference scheme which only requires $\mathcal{O}(n \log^2 n)$ operations per time step and $\mathcal{O}(n)$ storage requirement. After that, many fast iterative methods for solving the resulting linear systems have been proposed and analyzed; see [10,12,24,25,27,39] and the references therein.

Nevertheless, the numerical schemes used in literatures are commonly time-stepping schemes. The popular time-stepping schemes usually require many time steps in order to balance the errors arising from the spatial discretization. Moreover, the consistency and stability issues also need to be discussed carefully, especially for fractional diffusion equations since the nonlocal property of fractional derivatives may cause the stability issue of the scheme very sensitive [22,23]. In the last decade or two, an alternative technique based on Laplace transformation and numerical contour integration has become popular [6,8,14,15,31,32,36,41, 42]. This method represents the solution of the differential equation as an inverse Laplace transform and then approximates it by quadrature along a suitably chosen contour. Sheen, Sloan, and Thomée [31,32] firstly introduced this idea for parabolic equations and later McLean and Thomée [19–21] developed it to fractional-order evolution equations. The attractivity of such numerical integral method is that it can offer high order convergence rate in time direction and during the whole numerical process, it only requires to solve several shifted linear systems which can be further implemented in parallel [6,31,32,41]. More discussions and applications can be found in Gavrilyuk and Makarov [3], in't Hout and Weideman [6], Lee, Lee, and Sheen [8], Mclean, Sloan, Thomée [18] and the references therein.

Based on efficient numerical schemes for the Laplace inversion, in the present paper we employ the numerical contour integral method instead of traditional time-stepping method for solving space-fractional diffusion equations. Note that the key point for the success of the contour integral method is to choose a suitable type of contour and parameterize it optimally. Up to now, three types of contours, namely parabolic, hyperbolic, as well as Talbot contour, have been proposed and widely investigated [3,6,14,15,36,41,42]. Here we choose the hyperbolic contour due to the sectorial nature of the spatial discretized matrices. By making use of the Toeplitz-like structure of coefficient matrices and the relevant properties, we give the concrete regions in which the spectra of coefficient matrices are contained. Moreover, we prove that the resolvent norms of the resulting coefficient matrices are bounded outside of these regions. Therefore, suitable parameters in the hyperbolic contour can be determined according to these regions. To reduce the

computational burden, the preconditioned GMRES method with circulant preconditioners [2,10] is used to solve the resulting shifted linear systems. Numerical examples demonstrate the efficiency of our methods.

The paper is arranged as follows. In Section 2, we start with a review of the numerical contour integral method and then present the results for parameter selections of the hyperbolic contour in [42]. Applications to one and two dimensional space-fractional diffusion equations are given in Section 3. In Section 4, numerical examples are provided and compared to traditional time-stepping methods. At last, we make some concluding remarks in Section 5.

2 Numerical contour integral method with the hyperbolic contour

2.1 Numerical contour integral method

First, we briefly review the numerical contour integral method. Consider a system of ordinary differential equations (ODEs)

$$\frac{\mathrm{d}\mathbf{u}(t)}{\mathrm{d}t} = A\mathbf{u}(t) + \mathbf{b}(t), \quad \mathbf{u}(0) = \mathbf{u}_0, \tag{2.1}$$

where $A \in \mathbb{R}^{n\times n}$, $\mathbf{u}(t)$, $\mathbf{u}_0$, and $\mathbf{b}(t) \in \mathbb{R}^n$. Recall that for a function $f(t)$, its Laplace transform is defined by

$$\hat{f}(z) = \int_0^\infty e^{-zt} f(t)\mathrm{d}t.$$

By taking the Laplace transform of (2.1), one obtains

$$\hat{\mathbf{u}}(z) = (zI - A)^{-1}\left[\mathbf{u}_0 + \hat{\mathbf{b}}(z)\right] \tag{2.2}$$

provided z is not part of the spectrum of A, where I is the identity matrix, $\hat{\mathbf{u}}(z)$ and $\hat{\mathbf{b}}(z)$ are the Laplace transforms of $\mathbf{u}(t)$ and $\mathbf{b}(t)$, respectively. Then by the Laplace inversion of (2.2), the solution of (2.1) can be expressed as

$$\mathbf{u}(t) = \frac{1}{2\pi\mathbf{i}} \int_\Gamma e^{zt}\hat{\mathbf{u}}(z)\mathrm{d}z, \tag{2.3}$$

where $\mathbf{i} \equiv \sqrt{-1}$, the contour Γ is a straight line parallel to the imaginary axis such that all the singularities of $\hat{\mathbf{u}}(z)$ lie in the left half plane of it. Finally, numerical quadratures are applied to the integration of (2.3) to approximate the solution $\mathbf{u}(t)$.

In practical computations, the curve Γ in (2.3) is often deformed to a special Hankel contour whose real part begins at the negative infinity in the third quadrant, then winds anticlockwise around the spectrum of A and all singularities of $\hat{\mathbf{b}}(z)$, and last terminates with the real part again going to negative infinity in the second quadrant [41,42]. The motivation of this contour deformation comes from the observation that the factor e^{zt} in the integrand decays rapidly as the real part of z goes to the negative infinity, and such rapidly decaying property makes the integral particularly suitable for approximation by the trapezoidal or midpoint rules [36].

Suppose the curve Γ in (2.3) is deformed to a suitable Hankel contour parameterized by

$$\Gamma: \quad z = z(\zeta), \quad -\infty < \zeta < \infty.$$

Then substituting this contour to (2.3) gives [6,41]

$$\mathbf{u}(t) = \frac{1}{2\pi \mathbf{i}} \int_{-\infty}^{+\infty} e^{z(\zeta)t} \hat{\mathbf{u}}(z(\zeta)) z'(\zeta) \mathrm{d}\zeta. \tag{2.4}$$

Discretization of this integral with uniform node spacing h yields

$$\mathbf{u}(t) \approx \frac{h}{2\pi \mathbf{i}} \sum_{k=-\infty}^{\infty} e^{z(\zeta_k)t} z'(\zeta_k) \hat{\mathbf{u}}(z(\zeta_k)), \tag{2.5}$$

where $\zeta_k \equiv kh$ for the trapezoidal rule, $\zeta_k \equiv (k + \frac{1}{2})h$ for the midpoint rule. In practical computations, the series should be truncated. Moreover, we follow [41] and prefer using the midpoint rule as it requires one less function evaluation (or saves one linear system solve) than the trapezoidal rule. If the contour Γ is further assumed to be symmetric with respect to the real axis, then (2.5) can be approximated by [6,41]

$$\mathbf{u}(t) \approx \frac{h}{\pi} \mathrm{Im} \left\{ \sum_{k=0}^{N-1} e^{z_k t} z_k' \hat{\mathbf{u}}_k \right\} \tag{2.6}$$

for the midpoint rule, with $z_k \equiv z(\zeta_k)$, $z_k' \equiv z'(\zeta_k)$, $\hat{\mathbf{u}}_k \equiv \hat{\mathbf{u}}(z(\zeta_k))$, N being the number of quadrature nodes. Here the vectors $\hat{\mathbf{u}}_k$ come from solving the N linear systems

$$(z_k I - A)\hat{\mathbf{u}}_k = \mathbf{u}_0 + \hat{\mathbf{b}}(z_k), \quad k = 0, 1, \ldots, N-1, \tag{2.7}$$

which generally represent the major computational cost of the method. In the present work, we exploit the hyperbolic contour.

2.2 Parameter selections of the hyperbolic contour

The hyperbolic contour can be parameterized by [14,42]

$$z = \mu[1 + \sin(\mathbf{i}\zeta - \theta)], \quad -\infty < \zeta < +\infty, \tag{2.8}$$

where parameters $\mu > 0$ and θ set the width and the asymptotic angle of the hyperbolic contour, respectively. The selection of parameters μ and θ, as well as the step-size h in (2.6) dramatically affects desired features of the numerical contour integral method. Weideman and Trefethen [42] carefully discussed the parameter selection for the hyperbolic contour and provided elegant and simple results. The analyses in [42] were based on a theorem of Martensen [17], which is originally stated for the real-valued function but can be trivially modified for the complex-valued function with unsymmetrically distributed singularities [6,42], namely,

$$\int_{-\infty}^{+\infty} g(\zeta)\mathrm{d}\zeta = h \sum_{k=-\infty}^{\infty} g(\zeta_k) + DE_+(h) + DE_-(h), \tag{2.9}$$

where the discretization errors, $DE_+(h)$ and $DE_-(h)$, are bounded by

$$|DE_+(h)| \le \frac{M(\nu_+)}{e^{2\pi\nu_+/h}-1} \quad \text{and} \quad |DE_-(h)| \le \frac{M(\nu_-)}{e^{-2\pi\nu_-/h}-1}, \tag{2.10}$$

respectively, with $\nu_+ > 0$, $\nu_- < 0$, and

$$M(\nu) \equiv \int_{-\infty}^{+\infty} |g(\zeta + \mathbf{i}\nu)| \mathrm{d}\zeta, \quad \nu = \nu_+(\text{ or } \nu_-). \tag{2.11}$$

The function $g(\zeta + \mathbf{i}\nu)$ is assumed to be analytic in the strip $\nu \in (\nu_-, \nu_+)$. Obviously, the wider this strip of analyticity, the quicker the convergence. If the infinite series in (2.9) is further truncated after $k = N$, then the truncation error can be approximated by the magnitude of the last term retained, i.e.,

$$TE = \mathcal{O}(|g(hN)|), \quad N \to \infty. \tag{2.12}$$

In [42], the authors investigated the problems with the integrand $g(\zeta)$ being scalar functions. Under the assumption that all the singularities of the integrand in (2.9) fall into a sectorial region

$$\Sigma_\delta = \left\{ z \in \mathbb{C} : |\arg(-z)| \le \delta,\ \delta \in \left(0, \frac{\pi}{2}\right) \right\}, \tag{2.13}$$

they derived the expressions of parameters h and μ by asymptotically matching DE_+, DE_-, and TE:

$$h = \frac{K(\theta)}{N}, \quad \mu = \frac{4\pi\theta - \pi^2 + 2\pi\delta}{K(\theta)} \frac{N}{t}, \tag{2.14}$$

where θ is regarded as the free variable and

$$K(\theta) = \cosh^{-1}\left(\frac{2\theta}{(4\theta - \pi + 2\delta)\sin\theta} \right). \tag{2.15}$$

With this choice the predicted convergence rate is given by

$$E_N = \left| \int_{-\infty}^{+\infty} g(\zeta)\mathrm{d}\zeta - h \sum_{k=-N}^{N-1} g(\zeta_k) \right| = \mathcal{O}\left(e^{-R(\theta)N} \right),\ R(\theta) = \frac{\pi^2 - 2\pi\theta - 2\pi\delta}{K(\theta)}. \tag{2.16}$$

According to (2.16), we see that once the semiangle δ of the sectorial region (2.13) is given, then one can find the optimal θ by maximizing the function $R(\theta)$. Consequently, the optimal parameters h and μ can be obtained from the formulas (2.14) and (2.15).

We note that the estimates in (2.10) and (2.12) can be extended in a straightforward manner to the case when $g(\zeta)$ takes value in a complex vector space. This is the situation for the integral (2.4). Nevertheless, in this case the quantity $M(\nu_+)$ in the error DE_+, which has been regarded as a constant in the analysis of [42], can become extremely large. As a result, the error term DE_+ may fail to give an accurate estimate of the error and the parameter estimates (2.14) would become unreliable. Indeed, if we apply (2.11) to (2.4), then we obtain

$$M(\nu_+) \le \frac{1}{2\pi} \max_{-\infty<\zeta<\infty} \left\{ \|(z(w)I - A)^{-1}\| \|\tilde{\mathbf{b}}(z(w))\| \right\} \int_{-\infty}^{\infty} |e^{z(w)t} z'(w)| \mathrm{d}\zeta, \tag{2.17}$$

where $w = \zeta + \mathbf{i}\nu_+$, $\tilde{\mathbf{b}}(z(w)) = \mathbf{u}_0 + \hat{\mathbf{b}}(z(w))$, and $\|\cdot\|$ is some matrix norm. To make the denominator in the error bound of DE_+ large, we should increase the value ν_+. However, the hyperbola $z(\zeta + \mathbf{i}\nu_+)$ will be close to the spectrum of A as ν_+ is increased. If the matrix A is nonnormal, the resolvent norm $\|(z(w)I - A)^{-1}\|$ in (2.17) may already be huge for much smaller values of ν_+. Therefore, for the vector-valued case (for instance the integral (2.4)), we should firstly determine a region such that not only all the singularities of the integrand are contained in this region, but also inside of which the resolvent norm in (2.17) can be large, while outside of which it is relatively small. Then based on this region, we choose the reliable parameters in the hyperbolic contour. In the following, we will carefully investigate those critical regions for one- and two-dimensional space-fractional diffusion equations.

3 Contour integral method for fractional diffusion equations

As it has been pointed out by [38–40], after the spatial discretization of space-fractional diffusion equations by the shifted Grünwald formula, the resulting coefficient matrices will possess the Toeplitz-like structure. In this section, we closely look at these coefficient matrices and identify the regions that their spectra lie in. Moreover, inside of these regions the resolvent norms of the coefficient matrices are large while outside of them the resolvent norms are relatively small. Based on the regions, we choose suitable parameters for the hyperbolic contour and then solve the equations efficiently by the numerical contour integral method. First, we introduce some background knowledge, which will come into use afterward.

3.1 Background knowledge

A matrix that has the following form

$$A = \begin{bmatrix} a_0 & a_{-1} & \cdots & a_{2-n} & a_{1-n} \\ a_1 & a_0 & a_{-1} & \cdots & a_{2-n} \\ \vdots & a_1 & a_0 & \ddots & \vdots \\ a_{n-2} & \cdots & \ddots & \ddots & a_{-1} \\ a_{n-1} & a_{n-2} & \cdots & a_1 & a_0 \end{bmatrix} \tag{3.1}$$

is called a Toeplitz matrix [2], whose entries are constant along its diagonals. If the diagonals of A are the Fourier coefficients of a function f, i.e.,

$$a_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\eta) e^{-\mathbf{i}k\eta} \mathrm{d}\eta,$$

then the function f is called the generating function of A.

Definition 3.1 (see [29]) The intersection of all the convex sets containing a given set $\mathcal{S}$ is called the convex hull of $\mathcal{S}$ and is denoted by $\mathrm{conv}(\mathcal{S})$.

Definition 3.2 (see [5]) Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. The numerical range of A is defined as

$$\mathcal{W}(A) \equiv \{\mathbf{v}^* A\mathbf{v} \mid \mathbf{v} \in \mathbb{C}^n,\ \mathbf{v}^*\mathbf{v} = 1\}.$$

Let

$$\Omega(f) \equiv \{f(\eta) \mid \eta \in (-\pi, \pi)\}$$

denote the range of f. The following theorem reveals the relationship between the numerical range and the generating function of a Toeplitz matrix.

Theorem 3.1 (see [9,37]) *Suppose that the function $f(\eta)$ for $\eta \in (-\pi, \pi)$ is continuous. Let $A \in \mathbb{C}^{n\times n}$ be a Toeplitz matrix generated by $f(\eta)$. Then $\mathcal{W}(A)$ is a subset of the closure of* $\mathrm{conv}(\Omega(f))$*; i.e.,*

$$\mathcal{W}(A) \subseteq \overline{\mathrm{conv}(\Omega(f))}. \tag{3.2}$$

Another concept which is closely related to the numerical range is also useful in our following discussions, namely, the angular numerical range.

Definition 3.3 (see [5]) Let $A \in \mathbb{C}^{n\times n}$ be a square matrix. The angular numerical range of A is defined as

$$\mathcal{W}'(A) \equiv \{\mathbf{v}^* A\mathbf{v} \mid \mathbf{v} \in \mathbb{C}^n,\ \mathbf{v} \neq 0\}.$$

It is clear that $\mathcal{W}'(A)$ is a sectorial region of the complex plane that is anchored at the origin. The field angle is just the angular opening of the smallest angular sector that inclues $\mathcal{W}(A)$. For the angular numerical range, we have the following result.

Lemma 3.1 (see [5, Property 1.2.12]) *Let $A \in \mathbb{C}^{n\times n}$ be a square matrix and suppose that $D \in \mathbb{C}^{n\times n}$ is nonsingular. Then*

$$\mathcal{W}'(D^* AD) = \mathcal{W}'(A). \tag{3.3}$$

3.2 One dimensional space-fractional diffusion equations

We first consider the following one dimensional space-fractional diffusion equation [23,40]

$$\begin{cases} \dfrac{\partial u(x,t)}{\partial t} = d_1(x)\dfrac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha} + d_2(x)\dfrac{\partial^\alpha u(x,t)}{\partial_- x^\alpha} + b(x,t), & (x,t) \in (x_L, x_R)\times(0,T], \\ u(x_L,t) = u(x_R,t) = 0, & t \in (0,T], \\ u(x,0) = u_0(x), & x \in (x_L, x_R), \end{cases} \tag{3.4}$$

where $1 < \alpha < 2$, the diffusion coefficients satisfy $d_1(x) \geq 0$ and $d_2(x) \geq 0$ with $d_1(x) + d_2(x) > 0$, $b(x,t)$ is the forcing term. The left-sided $(+)$ and the right-sided $(-)$ fractional derivatives $\frac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha}$ and $\frac{\partial^\alpha u(x,t)}{\partial_- x^\alpha}$ can be defined in the Grünwald-Letnikov form [26]

$$\frac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha} = \lim_{\Delta x \to 0^+} \frac{1}{\Delta x^\alpha} \sum_{k=0}^{\lfloor (x-x_L)/\Delta x \rfloor} g_k^{(\alpha)} u(x - k\Delta x, t), \tag{3.5}$$

$$\frac{\partial^\alpha u(x,t)}{\partial_- x^\alpha} = \lim_{\Delta x \to 0^+} \frac{1}{\Delta x^\alpha} \sum_{k=0}^{\lfloor (x_R-x)/\Delta x \rfloor} g_k^{(\alpha)} u(x + k\Delta x, t), \tag{3.6}$$

where $\lfloor x \rfloor$ denotes the floor of x, and $g_k^{(\alpha)} = (-1)^k \binom{\alpha}{k}$ with $\binom{\alpha}{k}$ the fractional binomial coefficients.

Let $\Delta x = (x_R - x_L)/(n+1)$ with n being a positive integer. We define a spatial partition $x_i = x_L + i\Delta x$ for $i = 0, 1, \ldots, n+1$. Let $u_i(t) = u(x_i, t)$, $d_{1,i} = d_1(x_i)$, $d_{2,i} = d_2(x_i)$, and $b_i(t) = b(x_i, t)$. By the shifted Grünwald approximations [22,23]

$$\frac{\partial^\alpha u(x_i,t)}{\partial_+ x^\alpha} = \frac{1}{\Delta x^\alpha} \sum_{k=0}^{i+1} g_k^{(\alpha)} u_{i-k+1}(t) + \mathcal{O}(\Delta x),$$

$$\frac{\partial^\alpha u(x_i,t)}{\partial_- x^\alpha} = \frac{1}{\Delta x^\alpha} \sum_{k=0}^{n-i+2} g_k^{(\alpha)} u_{i+k-1}(t) + \mathcal{O}(\Delta x),$$

the fractional diffusion equation (3.4) can be discretized in the following form

$$\frac{\partial u_i(t)}{\partial t} = \frac{d_{1,i}}{\Delta x^\alpha} \sum_{k=0}^{i+1} g_k^{(\alpha)} u_{i-k+1}(t) + \frac{d_{2,i}}{\Delta x^\alpha} \sum_{k=0}^{n-i+2} g_k^{(\alpha)} u_{i+k-1}(t) + b_i(t), \quad i = 1, 2, \ldots, n.$$

Let $\mathbf{u}(t) = [u_1(t), u_2(t), \ldots, u_n(t)]^\intercal$ and $\mathbf{b}(t) = [b_1(t), b_2(t), \ldots, b_n(t)]^\intercal$. Then the above finite difference scheme can be expressed in the matrix form

$$\frac{\mathrm{d}\mathbf{u}(t)}{\mathrm{d}t} = A\mathbf{u}(t) + \mathbf{b}(t), \tag{3.7}$$

where

$$A = \frac{1}{\Delta x^\alpha}(D_1 G_\alpha + D_2 G_\alpha^\intercal) \tag{3.8}$$

with $D_1 = \mathrm{diag}(d_{1,1}, d_{1,2}, \ldots, d_{1,n})$, $D_2 = \mathrm{diag}(d_{2,1}, d_{2,2}, \ldots, d_{2,n})$, and

$$G_\alpha = \begin{bmatrix} g_1^{(\alpha)} & g_0^{(\alpha)} & 0 & \cdots & 0 & 0 \\ g_2^{(\alpha)} & g_1^{(\alpha)} & g_0^{(\alpha)} & 0 & \cdots & 0 \\ \vdots & g_2^{(\alpha)} & g_1^{(\alpha)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ g_{n-1}^{(\alpha)} & \ddots & \ddots & \ddots & g_1^{(\alpha)} & g_0^{(\alpha)} \\ g_n^{(\alpha)} & g_{n-1}^{(\alpha)} & \cdots & \cdots & g_2^{(\alpha)} & g_1^{(\alpha)} \end{bmatrix}. \tag{3.9}$$

For the ODEs (3.7), we can apply the numerical contour integral method with the hyperbolic contour as described in Section 2 to solve it. To make the contour integral method efficient, suitable parameters of the hyperbolic contour should be chosen. To this end, we need to first discuss the location of the spectrum of the coefficient matrix A.

Next, we consider the case that the diffusion coefficients $d_1(x)$ and $d_2(x)$ in (3.4) are proportional to each other, or equivalently, $d_l(x) = c_l \cdot d(x)$ for a certain function $d(x) > 0$ and some constants $c_l \geq 0$, $l = 1, 2$. Under this assumption, the coefficient matrix A in (3.8) reduces to

$$A = \frac{1}{\Delta x^\alpha} D(c_1 G_\alpha + c_2 G_\alpha^\intercal), \tag{3.10}$$

where $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ is positive definite with $d_i = d(x_i)$ for $1 \leq i \leq n$. Since c_1 and c_2 are constants, it follows from (3.9) that the matrix $\tilde{A} = c_1 G_\alpha + c_2 G_\alpha^\intercal$ is a Toeplitz matrix. The lemma below gives the generating function of $\tilde{A}$.

Lemma 3.2 *Let the matrix $\tilde{A} = c_1 G_\alpha + c_2 G_\alpha^\intercal$ with $c_1 \geq 0$ and $c_2 \geq 0$ being constants and G_α being given by (3.9). Then the generating function of $\tilde{A}$ is given by*

$$f(\eta) = \begin{cases} \left(2\sin\dfrac{\eta}{2}\right)^\alpha \left[(c_1+c_2)\cos\left(\dfrac{\alpha}{2}(\pi-\eta)+\eta\right)\right. \\ \qquad\qquad \left. +\mathbf{i}(c_2-c_1)\sin\left(\dfrac{\alpha}{2}(\pi-\eta)+\eta\right)\right], \quad \eta\in[0,\pi), \\ \left(2\sin\dfrac{-\eta}{2}\right)^\alpha \left[(c_1+c_2)\cos\left(\dfrac{\alpha}{2}(\pi+\eta)-\eta\right)\right. \\ \qquad\qquad \left. +\mathbf{i}(c_1-c_2)\sin\left(\dfrac{\alpha}{2}(\pi+\eta)-\eta\right)\right], \quad \eta\in(-\pi,0). \end{cases} \tag{3.11}$$

Proof According to (3.9), the generating function f of $\tilde{A}$ can be written as

$$\begin{aligned} f(\eta) &= c_1 \sum_{k=0}^{\infty} g_k^{(\alpha)} e^{\mathbf{i}(k-1)\eta} + c_2 \sum_{k=0}^{\infty} g_k^{(\alpha)} e^{-\mathbf{i}(k-1)\eta} \\ &= c_1 e^{-\mathbf{i}\eta} \sum_{k=0}^{\infty} g_k^{(\alpha)} e^{\mathbf{i}k\eta} + c_2 e^{\mathbf{i}\eta} \sum_{k=0}^{\infty} g_k^{(\alpha)} e^{-\mathbf{i}k\eta}. \end{aligned}$$

Note that $g_k^{(\alpha)}$ in fact are the coefficients of the power series of the function $(1-z)^\alpha$, i.e.,

$$(1-z)^\alpha = \sum_{k=0}^{\infty} g_k^{(\alpha)} z^k, \quad |z| \leq 1.$$

Therefore,

$$\begin{aligned} f(\eta) &= c_1 e^{-\mathbf{i}\eta}\left(1-e^{\mathbf{i}\eta}\right)^\alpha + c_2 e^{\mathbf{i}\eta}\left(1-e^{-\mathbf{i}\eta}\right)^\alpha \\ &= c_1 e^{-\mathbf{i}\eta}\left(2\mathbf{i}\sin\frac{-\eta}{2} e^{\mathbf{i}\frac{\eta}{2}}\right)^\alpha + c_2 e^{\mathbf{i}\eta}\left(2\mathbf{i}\sin\frac{\eta}{2} e^{-\mathbf{i}\frac{\eta}{2}}\right)^\alpha, \end{aligned}$$

from which the expression of (3.11) can be obtained. □

Using the generating function of $\tilde{A}$, we can further prove that the (angular) numerical range of the matrix $\tilde{A}$ is contained in a sectorial region.

Lemma 3.3 *Let the matrix $\tilde{A} = c_1 G_\alpha + c_2 G_\alpha^\intercal$ with $c_1 \geq 0$ and $c_2 \geq 0$ being constants and G_α being given by (3.9). Then the numerical range $\mathcal{W}(\tilde{A})$ of $\tilde{A}$ lies in a sectorial region Σ_δ such that*

$$\mathcal{W}(\tilde{A}) \subseteq \Sigma_\delta \equiv \left\{ z\in\mathbb{C} : |\arg(-z)| \leq \delta,\ \delta = \arctan\left(\left|\frac{c_2-c_1}{c_1+c_2}\tan\left(\frac{\alpha}{2}\pi\right)\right|\right)\right\}. \tag{3.12}$$

Furthermore,

$$\mathcal{W}'(\tilde{A}) \subseteq \Sigma_\delta. \tag{3.13}$$

Proof From (3.11), it is easy to show that the real part of $f(\eta)$,

$$\mathrm{Re}(f(\eta)) \leq 0 \qquad \text{for} \quad c_1, c_2 \geq 0,\ 1<\alpha<2.$$

When $\eta = 0$, we have $f(0) = 0$. When $\eta \neq 0$, we have $\mathrm{Re}(f(\eta)) < 0$ and

$$\frac{\mathrm{Im}(f(\eta))}{\mathrm{Re}(f(\eta))} = \begin{cases} \dfrac{c_2 - c_1}{c_1 + c_2} \tan\left(\dfrac{\alpha}{2}(\pi - \eta) + \eta\right), & \eta \in (0, \pi), \\ \dfrac{c_1 - c_2}{c_1 + c_2} \tan\left(\dfrac{\alpha}{2}(\pi + \eta) - \eta\right), & \eta \in (-\pi, 0), \end{cases}$$

where $\mathrm{Im}(f(\eta))$ denotes the imaginary part of $f(\eta)$. According to the above equation,

$$\left|\frac{\mathrm{Im}(f(\eta))}{\mathrm{Re}(f(\eta))}\right| \leq \left|\frac{c_2 - c_1}{c_1 + c_2} \tan\left(\frac{\alpha}{2}\pi\right)\right|, \quad \eta \in (-\pi, \pi) \text{ and } \eta \neq 0.$$

Thus, we obtain that all the values of $f(\eta)$ are located in a sectorial region Σ_δ as defined in (3.12). Apparently the sectorial region Σ_δ is a closed convex hull. Therefore,

$$\overline{\mathrm{conv}(\Omega(f))} \subseteq \Sigma_\delta.$$

By Theorem 3.1, we further have

$$\mathcal{W}(\tilde{A}) \subseteq \overline{\mathrm{conv}(\Omega(f))},$$

and hence

$$\mathcal{W}(\tilde{A}) \subseteq \Sigma_\delta.$$

Recall that $\mathcal{W}'(\tilde{A})$ is also a sectorial region whose angle is just the angular opening of the smallest angular sector that includes $\mathcal{W}(\tilde{A})$. Therefore we have

$$\mathcal{W}'(\tilde{A}) \subseteq \Sigma_\delta.$$

□

In order to study the location of the spectrum $\sigma(A)$ of A, the following results are required.

Lemma 3.4 (see [5, Properties 1.2.4 and 1.2.6] *For all $A \in \mathbb{C}^{n\times n}$ and $\alpha \in \mathbb{C}$, we have*

(i) $\mathcal{W}(\alpha A) = \alpha \mathcal{W}(A)$;
(ii) $\sigma(A) \subseteq \mathcal{W}(A)$.

Lemma 3.5 *Let $A, D \in \mathbb{C}^{n\times n}$. Suppose that D is positive definite, then*

$$\sigma(DA) \subseteq \mathcal{W}'(A).$$

Proof Since D is positive definite, we have by Lemmas 3.1 and 3.4 that

$$\sigma(DA) = \sigma(D^{\frac{1}{2}} A D^{\frac{1}{2}}) \subseteq \mathcal{W}(D^{\frac{1}{2}} A D^{\frac{1}{2}}) \subseteq \mathcal{W}'(D^{\frac{1}{2}} A D^{\frac{1}{2}}) = \mathcal{W}'(A),$$

where $D^{\frac{1}{2}}$ is the square root of D. The proof is completed. □

With the help of Lemma 3.5, we have the following theorem.

Theorem 3.2 *Suppose that the diffusion coefficients of* (3.4) *satisfy* $d_l(x) = c_l \cdot d(x)$ *for a certain function* $d(x) > 0$ *and some constants* $c_l \geq 0$, $l = 1, 2$. *Then the spectrum* $\sigma(A)$ *of the resulting coefficient matrix* A *given by* (3.10) *lies in the sectorial region* Σ_δ *defined by* (3.12), *i.e.,*

$$\sigma(A) \subseteq \Sigma_\delta = \left\{ z \in \mathbb{C} : |\arg(-z)| \leq \delta,\ \delta = \arctan\left(\left| \frac{c_2 - c_1}{c_1 + c_2} \tan\left(\frac{\alpha}{2}\pi\right) \right| \right) \right\}. \quad (3.14)$$

Proof It is clear by (3.10) that

$$A = \frac{1}{\Delta x^\alpha} D \tilde{A},$$

where $\tilde{A} = c_1 G_\alpha + c_2 G_\alpha^\intercal$ and $D = \mathrm{diag}(d_1, d_2, \ldots, d_n)$ with $d_i = d(x_i) > 0$, $1 \leq i \leq n$. Since the diagonal matrix D is positive definite, we get by Lemma 3.5 that

$$\sigma(A) \subseteq \mathcal{W}'(\tilde{A}).$$

Moreover, from Lemma 3.3 we have that

$$\mathcal{W}'(\tilde{A}) \subseteq \Sigma_\delta,$$

where the region Σ_δ is defined by (3.12). Thus,

$$\sigma(A) \subseteq \Sigma_\delta.$$

□

Theorem 3.2 provides a region in which the spectrum of the coefficient matrix A defined by (3.10) is contained. Next we show that the resolvent norm $\|(zI - A)^{-1}\|_{D^{-1}}$ is large for z inside of this region, while small for z outside of it. Here $\|\cdot\|_{D^{-1}}$ is a matrix D^{-1}-norm whose definition is given as follows. Let the D^{-1}-inner product and the corresponding vector norm be defined by

$$(\mathbf{v}, \mathbf{v})_{D^{-1}} = \mathbf{v}^* D^{-1} \mathbf{v}, \quad \|\mathbf{v}\|_{D^{-1}} = \sqrt{(\mathbf{v}, \mathbf{v})_{D^{-1}}}, \quad \forall \mathbf{v} \in \mathbb{C}^n.$$

Then the matrix D^{-1}-norm is induced by the vector D^{-1}-norm, i.e.,

$$\|B\|_{D^{-1}} = \max_{\mathbf{v} \neq 0} \frac{\|B\mathbf{v}\|_{D^{-1}}}{\|\mathbf{v}\|_{D^{-1}}}, \quad \forall B \in \mathbb{C}^{n \times n}.$$

From the above definitions, one can easily obtain that

$$\|\mathbf{v}\|_{D^{-1}} = \|D^{-\frac{1}{2}} \mathbf{v}\|_2, \quad \|B\|_{D^{-1}} = \|D^{-\frac{1}{2}} B D^{\frac{1}{2}}\|_2, \quad (3.15)$$

where $\|\cdot\|_2$ stands for the spectral norm. Obviously, when D is the identity matrix, the above D^{-1}-norm reduces to the spectral norm. Denote $\mathrm{dist}\big(z, \Sigma_\delta\big)$ the distance from z to Σ_δ, i.e.,

$$\mathrm{dist}\big(z, \Sigma_\delta\big) = \inf\left\{ |z - \xi| : \xi \in \Sigma_\delta \right\}.$$

Then we have the following result.

Theorem 3.3 *Under the assumption of Theorem 3.2, we have*

$$\|(zI - A)^{-1}\|_{D^{-1}} \leq \frac{1}{\mathrm{dist}\big(z, \Sigma_\delta\big)}, \quad z \notin \Sigma_\delta, \quad (3.16)$$

with Σ_δ *defined by* (3.12).

Proof According to (3.10) and (3.15), the resolvent norm

$$\begin{aligned}\|(zI-A)^{-1}\|_{D^{-1}} &= \|D^{-\frac{1}{2}}(zI-A)^{-1}D^{\frac{1}{2}}\|_2\\ &= \|(zI-\frac{1}{\Delta x}D^{\frac{1}{2}}\tilde{A}D^{\frac{1}{2}})^{-1}\|_2\\ &\le \frac{1}{\mathrm{dist}\left(z,\mathcal{W}\left(\frac{1}{\Delta x^\alpha}D^{\frac{1}{2}}\tilde{A}D^{\frac{1}{2}}\right)\right)}, \quad z\notin\mathcal{W}\left(\frac{1}{\Delta x^\alpha}D^{\frac{1}{2}}\tilde{A}D^{\frac{1}{2}}\right),\end{aligned}$$

the inequality is taken from Theorem 4.1 in [33]. By Lemmas 3.1 and 3.3, we have

$$\mathcal{W}\left(\frac{1}{\Delta x^\alpha}D^{\frac{1}{2}}\tilde{A}D^{\frac{1}{2}}\right)\subseteq\mathcal{W}'\left(\frac{1}{\Delta x^\alpha}D^{\frac{1}{2}}\tilde{A}D^{\frac{1}{2}}\right)=\mathcal{W}'(\tilde{A})\subseteq\Sigma_\delta.$$

Therefore,

$$\|(zI-A)^{-1}\|_{D^{-1}}\le\frac{1}{\mathrm{dist}\big(z,\Sigma_\delta\big)}, \quad z\notin\Sigma_\delta.$$

□

According to Theorem 3.3, the value $M(\nu_+)$ in (2.17) should be bounded for $\mathrm{dist}\big(z(w),\Sigma_\delta\big)\ge c>0$ with certain constant c.

We note that to choose an effective hyperbolic contour for (3.7), we still need the knowledge of the singularity distribution of the Laplace transform $\hat{\mathbf{b}}(z)$. If the singularities of $\hat{\mathbf{b}}(z)$ are also located in the sectorial region Σ_δ in (3.14), then we can determine the parameters of the hyperbolic contour according to the formulas (2.14)–(2.16) with δ given by (3.14). Nevertheless, the singularity distribution of $\hat{\mathbf{b}}(z)$ can be various. To make the discussion manageable, as in [42], here and in the following we simply assume that the singularities of $\hat{\mathbf{b}}(z)$ are located in the region that the spectrum of A is contained.

Finally, we remark that when the diffusion coefficients $d_1(x)$ and $d_2(x)$ in (3.7) are not proportional to each other, we have no similar result as in Theorems 3.2 and 3.3 at present. To apply the numerical contour integral method, one can first construct a matrix $\hat{A}$ by replacing the diagonal matrices in (3.8) by the mean of their corresponding diffusion coefficients. Then it is back to the constant coefficient case and we can apply Theorems 3.2 and 3.3 to $\hat{A}$ to estimate the parameters of the hyperbolic contour. Numerical experiments in Section 5 show that such strategy is efficient and acceptable.

3.3 Two dimensional space-fractional diffusion equations

We next consider the following two dimensional space-fractional diffusion equation [39]

$$\begin{cases}\dfrac{\partial u(x,y,t)}{\partial t} = d_1(x,y)\dfrac{\partial^\alpha u(x,y,t)}{\partial_+ x^\alpha}+d_2(x,y)\dfrac{\partial^\alpha u(x,y,t)}{\partial_- x^\alpha}+d_3(x,y)\dfrac{\partial^\beta u(x,y,t)}{\partial_+ y^\beta}\\ \qquad\qquad +d_4(x,y)\dfrac{\partial^\beta u(x,y,t)}{\partial_- y^\beta}+b(x,y,t), \qquad (x,y)\in\mathcal{D},\ t\in(0,T],\\ u(x,y,t)=0, \qquad (x,y)\in\partial\mathcal{D},\ t\in(0,T],\\ u(x,y,0)=u_0(x,y), \quad x\in\overline{\mathcal{D}},\end{cases}\tag{3.17}$$

where $1 < \alpha, \beta < 2$ and $\mathcal{D} = (x_L, x_R) \times (y_L, y_R)$ is a rectangular domain. The diffusion coefficients satisfy $d_l(x, y) \geq 0$ for $l = 1, 2, 3, 4$. The left-sided $(+)$ and right-sided $(-)$ fractional derivatives $\frac{\partial^\alpha u(x,y,t)}{\partial_+ x^\alpha}$, $\frac{\partial^\alpha u(x,y,t)}{\partial_- x^\alpha}$, $\frac{\partial^\alpha u(x,y,t)}{\partial_+ y^\alpha}$, and $\frac{\partial^\alpha u(x,y,t)}{\partial_- y^\alpha}$ are defined by the Grünwald-Letnikov form [26] as in (3.5) and (3.6).

Let $\Delta x = (x_R - x_L)/(m+1)$, $\Delta y = (y_R - y_L)/(n+1)$ with m, n being positive integers. We define a spatial partition $x_i = x_L + i\Delta x$ for $i = 0, 1, \ldots, m+1$, $y_j = y_L + j\Delta y$ for $j = 0, 1, \ldots, n+1$. Let $u_{i,j}(t) = u(x_i, y_j, t)$, $d_{i,j}^{(l)} = d_l(x_i, y_j)$, and $b_{i,j}(t) = b(x_i, y_j, t)$. By the shifted Grünwald approximations [23,35,38]

$$\frac{\partial^\alpha u(x_i, y_j, t)}{\partial_+ x^\alpha} = \frac{1}{\Delta x^\alpha} \sum_{k=0}^{i+1} g_k^{(\alpha)} u_{i-k+1,j}(t) + \mathcal{O}(\Delta x),$$

$$\frac{\partial^\alpha u(x_i, y_j, t)}{\partial_- x^\alpha} = \frac{1}{\Delta x^\alpha} \sum_{k=0}^{m-i+2} g_k^{(\alpha)} u_{i+k-1,j}(t) + \mathcal{O}(\Delta x),$$

$$\frac{\partial^\beta u(x_i, y_j, t)}{\partial_+ y^\beta} = \frac{1}{\Delta y^\beta} \sum_{k=0}^{j+1} g_k^{(\beta)} u_{i,j-k+1}(t) + \mathcal{O}(\Delta y),$$

$$\frac{\partial^\beta u(x_i, y_j, t)}{\partial_- y^\beta} = \frac{1}{\Delta y^\beta} \sum_{k=0}^{n-j+2} g_k^{(\beta)} u_{i,j+k-1}(t) + \mathcal{O}(\Delta y),$$

the fractional diffusion equation (3.17) can be discretized in the following form

$$\begin{aligned}\frac{\partial u_{i,j}(t)}{\partial t} = {} & \frac{d_{i,j}^{(1)}}{\Delta x^\alpha} \sum_{k=0}^{i+1} g_k^{(\alpha)} u_{i-k+1,j}(t) + \frac{d_{i,j}^{(2)}}{\Delta x^\alpha} \sum_{k=0}^{m-i+2} g_k^{(\alpha)} u_{i+k-1,j}(t) + \frac{d_{i,j}^{(3)}}{\Delta y^\beta} \sum_{k=0}^{j+1} g_k^{(\beta)} u_{i,j-k+1}(t) \\ & + \frac{d_{i,j}^{(4)}}{\Delta y^\beta} \sum_{k=0}^{n-j+2} g_k^{(\beta)} u_{i,j+k-1}(t) + b_{i,j}(t), \qquad 1 \leq i \leq m,\ 1 \leq j \leq n. \quad (3.18)\end{aligned}$$

Let

$$\mathbf{u}(t) = [u_{1,1}(t), \ldots, u_{1,n}(t), u_{2,1}(t), \ldots, u_{2,n}(t), \ldots, u_{m,1}(t), \ldots, u_{m,n}(t)]^\intercal,$$

$$\mathbf{b}(t) = [b_{1,1}(t), \ldots, b_{1,n}(t), b_{2,1}(t), \ldots, b_{2,n}(t), \ldots, b_{m,1}(t), \ldots, b_{m,n}(t)]^\intercal.$$

Then the finite difference scheme (3.18) can be rewritten in the matrix form

$$\frac{\mathrm{d}\mathbf{u}(t)}{\mathrm{d}t} = A_{mn}\mathbf{u}(t) + \mathbf{b}(t) \tag{3.19}$$

with

$$A_{mn} = \frac{1}{\Delta x^\alpha}\Big[D_1(G_\alpha \otimes I_n) + D_2(G_\alpha^\intercal \otimes I_n)\Big] + \frac{1}{\Delta y^\beta}\Big[D_3(I_m \otimes G_\beta) + D_4(I_m \otimes G_\beta^\intercal)\Big], \tag{3.20}$$

where

$$D_l = \operatorname{diag}\left(d_{1,1}^{(l)}, \ldots, d_{1,n}^{(l)}, d_{2,1}^{(l)}, \ldots, d_{2,n}^{(l)}, \ldots, d_{m,1}^{(l)}, \ldots, d_{m,n}^{(l)}\right) \quad \text{for} \quad l = 1, 2, 3, 4,$$

G_α and G_β are Toeplitz matrices and have forms given by (3.9), I_m and I_n are identity matrices with orders m and n, respectively.

Analogous to the one dimensional case, we also employ the numerical contour integral method with the hyperbolic contour to solve (3.19). In order to choose the suitable parameters for the hyperbolic contour, we should carefully investigate the spectrum and the resolvent norm of the resulting coefficient matrix A_{mn}. Indeed, determining the location of the spectrum of the resulting coefficient matrix A_{mn} in general is not easy. Here we focus on two cases:

case (1) all the diffusion coefficients of (3.17) are proportional to each other, or equivalently, $d_l(x,y) = c_l \cdot d(x,y)$ for a certain function $d(x,y) > 0$ and some constants $c_l \geq 0$, $l = 1,2,3,4$;

case (2) the diffusion coefficients of (3.17) satisfy that

- $d_l(x,y) = c_l \cdot d_\alpha(x)$ for a certain function $d_\alpha(x) > 0$ and some constants $c_l \geq 0$, $l = 1,2$,
- $d_k(x,y) = c_k \cdot d_\beta(y)$ for a certain function $d_\beta(y) > 0$ and some constants $c_k \geq 0$, $k = 3,4$.

We first consider the **case (1)**. In this case, the resulting coefficient matrix A_{mn} in (3.20) reduces to the form

$$A_{mn} = D\left[\frac{1}{\Delta x^\alpha}\left(c_1 G_\alpha + c_2 G_\alpha^\intercal\right) \otimes I_n + I_m \otimes \frac{1}{\Delta y^\beta}\left(c_3 G_\beta + c_4 G_\beta^\intercal\right)\right], \tag{3.21}$$

where

$$D = \operatorname{diag}\left(d_{1,1}, \ldots, d_{1,n}, d_{2,1}, \ldots, d_{2,n}, \ldots, d_{m,1}, \ldots, d_{m,n}\right)$$

with $d_{i,j} = d(x_i, y_j) > 0$ for $1 \leq i \leq m$ and $1 \leq j \leq n$. The following properties are useful for our studies.

Lemma 3.6 (see [5, Properties 1.2.7 and 1.2.8, Theorem 4.2.16, Corollary 4.3.10])

(i) *For all* $A, B \in \mathbb{C}^{n\times n}$, $\mathcal{W}(A+B) \subseteq \mathcal{W}(A) + \mathcal{W}(B)$;
(ii) *For all* $A, U \in \mathbb{C}^{n\times n}$ *with* U *unitary,* $\mathcal{W}(U^*AU) = \mathcal{W}(A)$;
(iii) *For all* $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, *if* A *is normal, then* $\mathcal{W}(A\otimes B) = \operatorname{conv}\left(\mathcal{W}(A)\mathcal{W}(B)\right)$;
(iv) *For all* $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, *there exists a permutation matrix* P *such that*

$$A \otimes B = P^\intercal (B \otimes A) P \quad \textit{with} \quad P^\intercal = P^{-1}.$$

Note that a permutation matrix is unitary. Then from Lemma 3.6 (ii), (iii), and (iv), we immediately draw the following conclusion.

Corollary 3.1 *For all* $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, *if* B *is normal, then*

$$\mathcal{W}(A \otimes B) = \operatorname{conv}\left(\mathcal{W}(A)\mathcal{W}(B)\right).$$

Lemma 3.7 *For all* $A \in \mathbb{C}^{m\times m}$, $B \in \mathbb{C}^{n\times n}$, *we have*

$$\mathcal{W}(A \otimes I_n + I_m \otimes B) \subseteq \mathcal{W}(A) + \mathcal{W}(B).$$

Proof It is well known that the identity matrix is normal. Therefore by Lemma 3.6 (i) and (iii), and Corollary 3.1, we obtain that

$$\begin{aligned}\mathcal{W}(A \otimes I_n + I_m \otimes B) &\subseteq \mathcal{W}(A \otimes I_n) + \mathcal{W}(I_m \otimes B)\\ &= \operatorname{conv}\left(\mathcal{W}(A)\mathcal{W}(I_n)\right) + \operatorname{conv}\left(\mathcal{W}(I_m)\mathcal{W}(B)\right)\\ &= \mathcal{W}(A) + \mathcal{W}(B).\end{aligned}$$

The proof is completed. □

The following theorem gives the location of the spectrum of A_{mn} defined by (3.21).

Theorem 3.4 *Suppose that the diffusion coefficients of* (3.17) *satisfy* $d_l(x,y) = c_l \cdot d(x,y)$ *for a certain function* $d(x,y) > 0$ *and some constants* $c_l \geq 0$, $l = 1,2,3,4$. *Then the spectrum* $\sigma(A_{mn})$ *of the resulting coefficient matrix* A_{mn} *given by* (3.21) *lies in a sectorial region* Σ_δ *such that*

$$\sigma(A_{mn}) \subseteq \Sigma_\delta \equiv \{z \in \mathbb{C} : |\arg(-z)| \leq \delta,\ \delta = \max\{\delta_1, \delta_2\}\}, \tag{3.22}$$

where

$$\delta_1 = \arctan\left(\left|\frac{c_2 - c_1}{c_1 + c_2}\tan\left(\frac{\alpha}{2}\pi\right)\right|\right) \quad \text{and} \quad \delta_2 = \arctan\left(\left|\frac{c_4 - c_3}{c_3 + c_4}\tan\left(\frac{\beta}{2}\pi\right)\right|\right).$$

Proof Denote

$$\tilde{A}_\alpha = \frac{1}{\Delta x^\alpha}(c_1 G_\alpha + c_2 G_\alpha^\intercal), \qquad \tilde{A}_\beta = \frac{1}{\Delta y^\beta}(c_3 G_\beta + c_4 G_\beta^\intercal).$$

It is clear that both $\tilde{A}_\alpha$ and $\tilde{A}_\beta$ satisfy the conditions of Lemma 3.3. Then we have by Lemma 3.3 that

$$\mathcal{W}(\tilde{A}_\alpha) \subseteq \Sigma_{\delta_1} \equiv \left\{z \in \mathbb{C} : |\arg(-z)| \leq \delta_1,\ \delta_1 = \arctan\left(\left|\frac{c_2 - c_1}{c_1 + c_2}\tan\left(\frac{\alpha}{2}\pi\right)\right|\right)\right\}$$

and

$$\mathcal{W}(\tilde{A}_\beta) \subseteq \Sigma_{\delta_2} \equiv \left\{z \in \mathbb{C} : |\arg(-z)| \leq \delta_2,\ \delta_2 = \arctan\left(\left|\frac{c_4 - c_3}{c_3 + c_4}\tan\left(\frac{\beta}{2}\pi\right)\right|\right)\right\}.$$

Let

$$\tilde{A}_{mn} = \tilde{A}_\alpha \otimes I_n + I_m \otimes \tilde{A}_\beta.$$

It follows from Lemma 3.7 that

$$\mathcal{W}(\tilde{A}_{mn}) \subseteq \mathcal{W}(\tilde{A}_\alpha) + \mathcal{W}(\tilde{A}_\beta) \subseteq \Sigma_{\delta_1} + \Sigma_{\delta_2} = \Sigma_\delta$$

with Σ_δ defined by (3.22). Obviously, Σ_δ is a sectorial region with vertex at the origin, therefore we have that

$$\mathcal{W}'(\tilde{A}_{mn}) \subseteq \Sigma_\delta. \tag{3.23}$$

By (3.21), we know that

$$A_{mn} = D\tilde{A}_{mn},$$

where D is a diagonal matrix with diagonal entries being positive. Then by Lemma 3.5,

$$\sigma(A_{mn}) \subseteq \mathcal{W}'(\tilde{A}_{mn}).$$

The proof is completed from (3.23). □

Note that the diagonal matrix D defined in (3.21) is positive definite, following the similar proof of Theorem 3.3, we can also derive that

$$\|(zI - A_{mn})^{-1}\|_{D^{-1}} \leq \frac{1}{\operatorname{dist}\big(z, \Sigma_\delta\big)}, \quad z \notin \Sigma_\delta,$$

where Σ_δ is defined by (3.22). This inequality shows that the resolvent norm of the coefficient matrix A_{mn} will not be very large for z outside of the region Σ_δ, which implies that $M(\nu_+)$ in (2.17) should be bounded and the parameter estimates (2.14) are reliable.

Next, we consider the **case (2)**. For this case, the coefficient matrix A_{mn} in (3.20) reduces to the form

$$A_{mn} = \left[\frac{1}{\Delta x^\alpha} D_\alpha (c_1 G_\alpha + c_2 G_\alpha^\intercal)\right] \otimes I_n + I_m \otimes \left[\frac{1}{\Delta y^\beta} D_\beta (c_3 G_\beta + c_4 G_\beta^\intercal)\right], \quad (3.24)$$

where $D_\alpha = \operatorname{diag}(d_{\alpha,1}, d_{\alpha,2}, \ldots, d_{\alpha,m})$ and $D_\beta = \operatorname{diag}(d_{\beta,1}, d_{\beta,2}, \ldots, d_{\beta,n})$ with $d_{\alpha,i} = d_\alpha(x_i) > 0$ and $d_{\beta,j} = d_\beta(y_j) > 0$, $1 \leq i \leq m$, $1 \leq j \leq n$. The following properties are useful for our discussions.

Lemma 3.8 (see [5, Theorem 4.4.5]) *Let $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ be given. If $\sigma(A) = \{\lambda_1, \ldots, \lambda_m\}$ and $\sigma(B) = \{\mu_1, \ldots, \mu_n\}$, then*

$$\sigma\big((A \otimes I_n) + (I_m \otimes B)\big) = \{\lambda_i + \mu_j \mid i = 1, 2, \ldots, m,\ j = 1, 2, \ldots, n\}.$$

Lemma 3.9 (see [5, Lemma 4.2.10 and Corollary 4.2.11])

(i) *Let $A, C \in \mathbb{C}^{m\times m}$ and $B, D \in \mathbb{C}^{n\times n}$ be given, then*

$$(A \otimes B)(C \otimes D) = AC \otimes BD;$$

(ii) *If $A \in \mathbb{C}^{m\times m}$ and $B \in \mathbb{C}^{n\times n}$ are nonsingular, then so is $A \otimes B$, and*

$$(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}.$$

The theorem below shows that the spectrum of the resulting coefficient matrix A_{mn} given by (3.24) is also located in a sectorial region.

Theorem 3.5 *Suppose the diffusion coefficients of* (3.17) *satisfy that $d_l(x, y) = c_l \cdot d_\alpha(x)$ for a certain function $d_\alpha(x) > 0$ and some constants $c_l \geq 0$, $l = 1, 2$, and $d_k(x, y) = c_k \cdot d_\beta(y)$ for a certain function $d_\beta(y) > 0$ and some constants $c_k \geq 0$, $k = 3, 4$. Then the spectrum $\sigma(A_{mn})$ of the resulting coefficient matrix A_{mn} given by* (3.24) *lies in a sectorial region Σ_δ such that*

$$\sigma(A_{mn}) \subseteq \Sigma_\delta \equiv \{z \in \mathbb{C} : |\arg(-z)| \leq \delta,\ \delta = \max\{\delta_1, \delta_2\}\}, \quad (3.25)$$

where

$$\delta_1 = \arctan\left(\left|\frac{c_2 - c_1}{c_1 + c_2} \tan\left(\frac{\alpha}{2}\pi\right)\right|\right) \quad \text{and} \quad \delta_2 = \arctan\left(\left|\frac{c_4 - c_3}{c_3 + c_4} \tan\left(\frac{\beta}{2}\pi\right)\right|\right).$$

Proof Let

$$A_\alpha = \frac{1}{\Delta x^\alpha} D_\alpha (c_1 G_\alpha + c_2 G_\alpha^\intercal), \quad A_\beta = \frac{1}{\Delta y^\beta} D_\beta (c_3 G_\beta + c_4 G_\beta^\intercal).$$

Then both A_α and A_β have the same form as that of (3.10). Thus, in light of Theorem 3.2,

$$\sigma(A_\alpha) \subseteq \Sigma_{\delta_1} \equiv \left\{ z \in \mathbb{C} : |\arg(-z)| \le \delta_1,\ \delta_1 = \arctan\left(\left| \frac{c_2 - c_1}{c_1 + c_2} \tan\left(\frac{\alpha}{2}\pi\right) \right| \right) \right\} \tag{3.26}$$

and

$$\sigma(A_\beta) \subseteq \Sigma_{\delta_2} \equiv \left\{ z \in \mathbb{C} : |\arg(-z)| \le \delta_2,\ \delta_2 = \arctan\left(\left| \frac{c_4 - c_3}{c_3 + c_4} \tan\left(\frac{\beta}{2}\pi\right) \right| \right) \right\}. \tag{3.27}$$

From (3.24), we know that

$$A_{mn} = A_\alpha \otimes I_n + I_m \otimes A_\beta.$$

Therefore, we have by Lemma 3.8,

$$\sigma(A_{mn}) \subseteq \Sigma_{\delta_1} + \Sigma_{\delta_2} = \Sigma_\delta$$

with Σ_δ being defined by (3.25). □

Denote

$$D = D_\alpha \otimes D_\beta, \tag{3.28}$$

with D_α and D_β given in (3.24). Then D is a diagonal matrix with positive diagonal entries. In the following, we prove that the resolvent norm of A_{mn} defined by (3.24) is also bounded outside of the sectorial region Σ_δ given in (3.25).

Theorem 3.6 *Under the assumptions of Theorem 3.5, we have*

$$\|(zI - A_{mn})^{-1}\|_{D^{-1}} \le \frac{1}{\operatorname{dist}\big(z, \Sigma_\delta\big)}, \quad z \notin \Sigma_\delta, \tag{3.29}$$

where Σ_δ is defined by (3.25).

Proof By (3.15), we know that

$$\begin{aligned} \|(zI - A_{mn})^{-1}\|_{D^{-1}} &= \|D^{-\frac{1}{2}}(zI - A_{mn})^{-1} D^{\frac{1}{2}}\|_2 \\ &= \|(zI - D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}})^{-1}\|_2 \\ &\le \frac{1}{\operatorname{dist}\big(z, \mathcal{W}(D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}})\big)}, \quad z \notin \mathcal{W}(D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}}), \end{aligned} \tag{3.30}$$

where A_{mn} is defined by (3.24) and D is given in (3.28).

Note that the matrix A_{mn} can be written as

$$A_{mn} = (D_\alpha \tilde{A}_\alpha) \otimes I_n + I_m \otimes (D_\beta \tilde{A}_\beta)$$

with $\tilde{A}_\alpha = 1/\Delta x^\alpha(c_1 G_\alpha + c_2 G_\alpha^\intercal)$ and $\tilde{A}_\beta = 1/\Delta y^\beta(c_3 G_\beta + c_4 G_\beta^\intercal)$. According to Lemma 3.9, we obtain

$$\begin{aligned} D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}} &= (D_\alpha^{-\frac{1}{2}} \otimes D_\beta^{-\frac{1}{2}}) \left[(D_\alpha \tilde{A}_\alpha) \otimes I_n + I_m \otimes (D_\beta \tilde{A}_\beta) \right] (D_\alpha^{\frac{1}{2}} \otimes D_\beta^{\frac{1}{2}}) \\ &= (D_\alpha^{\frac{1}{2}} \tilde{A}_\alpha D_\alpha^{\frac{1}{2}}) \otimes I_n + I_m \otimes (D_\beta^{\frac{1}{2}} \tilde{A}_\beta D_\beta^{\frac{1}{2}}). \end{aligned}$$

Then, by Lemmas 3.7 and 3.1, we have

$$\begin{aligned} \mathcal{W}(D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}}) &\subseteq \mathcal{W}(D_\alpha^{\frac{1}{2}} \tilde{A}_\alpha D_\alpha^{\frac{1}{2}}) + \mathcal{W}(D_\beta^{\frac{1}{2}} \tilde{A}_\beta D_\beta^{\frac{1}{2}}) \\ &\subseteq \mathcal{W}'(D_\alpha^{\frac{1}{2}} \tilde{A}_\alpha D_\alpha^{\frac{1}{2}}) + \mathcal{W}'(D_\beta^{\frac{1}{2}} \tilde{A}_\beta D_\beta^{\frac{1}{2}}) \\ &= \mathcal{W}'(\tilde{A}_\alpha) + \mathcal{W}'(\tilde{A}_\beta). \end{aligned} \tag{3.31}$$

Notice that both $\tilde{A}_\alpha$ and $\tilde{A}_\beta$ satisfy the conditions of Lemma 3.3. Therefore, we get that

$$\mathcal{W}'(\tilde{A}_\alpha) \subseteq \Sigma_{\delta_1} \quad \text{and} \quad \mathcal{W}'(\tilde{A}_\beta) \subseteq \Sigma_{\delta_2} \tag{3.32}$$

with Σ_{δ_1} and Σ_{δ_2} being defined by (3.26) and (3.27), respectively. It follows from (3.31) and (3.32) that

$$\mathcal{W}(D^{-\frac{1}{2}} A_{mn} D^{\frac{1}{2}}) \subseteq \Sigma_{\delta_1} + \Sigma_{\delta_2} = \Sigma_\delta,$$

where Σ_δ is defined by (3.25). Hence, we have by (3.30) that

$$\|(zI - A_{mn})^{-1}\|_{D^{-1}} \leq \frac{1}{\operatorname{dist}\big(z, \Sigma_\delta\big)}, \quad z \notin \Sigma_\delta.$$

□

By now, we have considered two cases of (3.17). In both cases, the spectra of resulting coefficient matrices are contained in some sectorial regions. Moreover, the resolvent norms of coefficient matrices are bounded outside of those regions. Therefore, when the numerical contour integral method with hyperbolic contour is applied, we can parameterize the hyperbolic contour according to the formulas (2.14)–(2.16) with δ given by (3.22) or (3.25), which may lead to a fast convergence rate. Nevertheless, for the general situation we have no similar results for the resulting coefficient matrix. In practical computations, following the strategy proposed in one dimension case, we can firstly construct a matrix $\hat{A}_{mn}$ by replacing the diagonal matrices in (3.20) by the mean of their corresponding diffusion coefficients. Then it reduces to the constant coefficient case and we can utilize Theorem 3.4 (or Theorem 3.5) to estimate the parameters of the hyperbolic contour. We remark that this treatment is acceptable in practical computation providing that the diffusion coefficients are very smoothly; see examples in section 4. The theoretical analysis will remain for the future study.

Before the end of this section, we would like to emphasize that application of the numerical contour integral method finally results in solving N shifted linear systems as given by (2.7), which represents the main workload of the method. Recall that the resulting coefficient matrices of space-fractional diffusion equations possess Toeplitz-like structures [38,40]; see also (3.8) and (3.20). Many efficient strategies have been proposed and used for fast solving the resulting systems emerged in the time-stepping scheme, such as the fast conjugate gradient squared

method [38], multigrid method [25], preconditioned Krylov subspace methods with circulant-type preconditioners [10,24,27], and band preconditioners [12], etc. In our numerical experiments, we adopt the preconditioning strategies proposed in [10] to solve the resulting shifted linear systems. Finally, we remark that these shifted linear systems can be further solved in parallel, but we would not like to pursue this here.

4 Numerical experiments

In this section, we employ the numerical contour integral method with hyperbolic contour as described in Section 2 to solve a variety of space-fractional diffusion equations. Parameters in the hyperbolic contour are selected based on the results in Section 3. As comparisons, we also apply the implicit Euler method [22,23, 38] to solve these fractional diffusion equations. All experiments are carried out in MATLAB 7.10 (R2010a) on a PC with configuration: Intel(R)Xeon(R)CPU E3-1230 V2 @ 3.30GHz and 32.0GB RAM.

For all tables, "contour integral" represents the numerical contour integral method and "time-stepping" stands for the time-stepping method; "m" and "n" denote the spatial grid numbers; "N" refers to the number of quadrature nodes in the contour integral method and "M" is the number of time steps in the time-stepping method; "avegiter" means the average iteration numbers for solving corresponding linear systems; "CPU" refers to the total CPU time with unit in second; "Error" represents the maximum error $\|u(\cdot,T)-U(\cdot,T)\|_\infty$ at time T, where $u(\cdot,T)$ is the exact solution and $U(\cdot,T)$ is the numerical solution. In all numerical examples we set numerical quadrature error $E_N \leq \Delta x/20$ to ensure that the temporal error does not contaminate the spatial discretization error. The number of quadrature nodes N is chosen according to this setting. In addition, we solve the resulting linear systems by the preconditioned GMRES method with Strang's circulant preconditioner as proposed in [10] and stop the iteration processes if

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq 10^{-8},$$

where r_k is the residual vector after the k-th iteration.

Example 1. Consider the fractional diffusion equation

$$\begin{aligned}
&\frac{\partial u(x,t)}{\partial t} = d_1(x)\frac{\partial^\alpha u(x,t)}{\partial_+ x^\alpha} + d_2(x)\frac{\partial^\alpha u(x,t)}{\partial_- x^\alpha} + b(x,t), \quad (x,t)\in(0,1)\times(0,1], \quad (4.1)\\
&u(0,t) = u(1,t) = 0, \quad t\in[0,1],\\
&u(x,0) = 32x^3(1-x)^3, \quad x\in[0,1],
\end{aligned}$$

with the diffusion coefficients

$$d_1(x) = c_1\cdot(1+x)^\alpha(2-x)^\alpha, \quad d_2(x) = c_2\cdot(1+x)^\alpha(2-x)^\alpha,$$

and the source term

$$b(x,t) = -32e^{-t}\left\{x^3(1-x)^3 + \left[\frac{\Gamma(4)}{\Gamma(4-\alpha)}\left(c_1x^{3-\alpha} + c_2(1-x)^{3-\alpha}\right)\right.\right.$$
$$-\frac{3\Gamma(5)}{\Gamma(5-\alpha)}\left(c_1x^{4-\alpha} + c_2(1-x)^{4-\alpha}\right) + \frac{3\Gamma(6)}{\Gamma(6-\alpha)}\left(c_1x^{5-\alpha} + c_2(1-x)^{5-\alpha}\right)$$
$$\left.\left.-\frac{\Gamma(7)}{\Gamma(7-\alpha)}\left(c_1x^{6-\alpha} + c_2(1-x)^{6-\alpha}\right)\right](1+x)^{\alpha}(2-x)^{\alpha}\right\}.$$

where c_1 and c_2 are positive constants. The exact solution is

$$u(x,t) = 32e^{-t}x^3(1-x)^3.$$

This is a one-dimensional fractional diffusion equation and the diffusion coefficients are proportional to each other. According to Theorem 3.2, the spectrum of the spatial discretized matrix will be contained in a sectorial region Σ_δ with δ given by (3.14). Substituting δ into (2.16), we obtain the optimal θ by maximizing the function $R(\theta)$. Using this optimal θ, we can determine the quadrature node number N by setting $e^{-R(\theta)N} \leq \Delta x/20$ and the parameters h and μ by the formulas (2.14) and (2.15). Numerical results for solving Example 1 are presented in Table 1.

Table 1 Numerical results for Example 1 with $c_1 = 1$ and $c_2 = 10$.

		contour integral				time-stepping			
α	n	N	avegiter	CPU	Error	M	avegiter	CPU	Error
1.5	64	8	7.1	0.040	2.67E-3	32	6.0	0.077	2.64E-3
	128	9	7.1	0.044	1.33E-3	64	6.0	0.164	1.33E-3
	256	9	7.1	0.045	6.67E-4	128	6.0	0.389	6.66E-4
	512	10	7.1	0.066	3.36E-4	256	6.0	1.005	3.34E-4
	1024	11	8.1	0.094	1.68E-4	512	5.0	2.688	1.69E-4
1.8	64	5	7.2	0.020	8.79E-4	32	7.0	0.080	8.75E-4
	128	5	7.2	0.025	4.40E-4	64	6.0	0.163	4.39E-4
	256	5	7.2	0.026	2.19E-4	128	6.0	0.388	2.20E-4
	512	6	7.2	0.033	1.10E-4	256	6.0	1.005	1.10E-4
	1024	6	7.2	0.049	5.42E-5	512	5.0	2.706	5.51E-5

It is easy to see from the "Error" columns that setting the numerical quadrature error $E_N \leq \Delta x/20$ is small enough to match the spatial discretization error. By comparison, we observe that to achieve the similar accuracy, the numbers N of the quadrature nodes are much smaller than the numbers M of time steps. In another word, there are much fewer linear systems needed to be solved by the numerical contour integral method than that by the time-stepping method to achieve the similar accuracy. Correspondingly, the total CPU times required by the numerical contour method are much less than that by the time-stepping method. Moreover, for both methods the average iteration numbers of solving resulting linear systems are small and almostly not increasing with the increase of spatial grid numbers, which show the efficiency of the preconditioners.

It would be worth noting that in the above discussion, the contour integral method just provides the solution at $t = T$. In the event that the solution needs to

be computed in an interval $[t_0, t_1]$ with $t_1 = \Lambda t_0$, one can modify the parameters μ and $K(\theta)$ in the formulas (2.14) and (2.15) by [42]

$$\mu = \frac{4\pi\theta - \pi^2 + 2\pi\delta}{K(\theta)} \frac{N}{t_1}, \quad K(\theta) = \cosh^{-1}\left(\frac{(\pi - 2\theta - 2\delta)\Lambda + (4\theta - \pi + 2\delta)}{(4\theta - \pi + 2\delta)\sin\theta}\right).$$

In this case the accuracy of the contour integral method should not suffer unduly even for a large Λ (say $\Lambda = 50$, see [15,42]) if the same contour is used for any t in the range $t_0 \le t \le t_1$. This approach will be much cheaper than using a different (optimal) contour for each of many different values of t in this range.

To gain further insight into the convergence of the numerical quadrature, we next only measure the temporal error which is defined as the difference between the exact solution vaule $\mathbf{u}(T)$ of the semidiscrete ODE system and its approximation obtained by the contour integral method. Recall that the semidiscrete system of (4.1) has the form (3.7) with $\mathbf{b}(t) = e^{-t}\tilde{\mathbf{b}}$, where $\tilde{\mathbf{b}}$ is independent of t. By the variation-of-constants formula, the exact solution of the semidiscrete system of (4.1) at time T is given by

$$\begin{aligned}\mathbf{u}(T) &= e^{TA}\mathbf{u}_0 + \int_0^T e^{(T-\tau)A} e^{-\tau}\tilde{\mathbf{b}}\mathrm{d}\tau \\ &= e^{TA}\left[\mathbf{u}_0 + (A+I)^{-1}\tilde{\mathbf{b}}\right] - e^{-T}(A+I)^{-1}\tilde{\mathbf{b}}\end{aligned}$$

provided that $A + I$ is invertible. We fix the spatial grid number $n = 256$ and compute the reference solution $\mathbf{u}(T)$ using the `expm` command in Matlab to obtain the matrix e^{TA} and "\" command to get the vector $(A+I)^{-1}\tilde{\mathbf{b}}$.

Figure 1 plots the temporal errors of our contour integral method for a variety of α, c_1, and c_2. The solid curve represents the actual error and the dash line refers to their theoretical estimate. It clearly shows the good agreement between the theoretical convergence and the practical one. Also observe from Figure 1 is that when $c_1 = c_2$, the contour integral method always has rapid convergence rate, say $e^{-2.3157N}$. This is because when $c_1 = c_2$, the semiangle $\delta = 0$ by Theorem 3.2 and thus the convergence rate of the contour integral method is equal to $e^{-2.3157N}$ according to (2.16). When $c_1 \neq c_2$, the convergence rate of the contour integral method becomes worse and worse as the value of α tends to 1. Since as the value of α decreases, the value of δ will increase by (3.14) and then in light of (2.16) the corresponding convergence rate may decrease.

Example 2. (see [23,40]) Consider the fractional diffusion equation

$$\begin{aligned}&\frac{\partial u(x,t)}{\partial t} = d_1(x)\frac{\partial^{1.8}u(x,t)}{\partial_+ x^{1.8}} + d_2(x)\frac{\partial^{1.8}u(x,t)}{\partial_- x^{1.8}} + b(x,t), \quad (x,t) \in (0,2)\times(0,1], \\ &u(0,t) = u(2,t) = 0, \quad t \in [0,1], \\ &u(x,0) = 4x^2(2-x)^2, \quad x \in [0,2],\end{aligned}$$

with coefficients

$$d_1(x) = \Gamma(1.2)x^{1.8}, \quad d_2(x) = \Gamma(1.2)(2-x)^{1.8},$$

and the source term

$$b(x,t) = -32e^{-t}\Big\{x^2 + (1+0.125x^2)(2-x)^2 - 2.5[x^3 + (2-x)^3] + \frac{25}{22}[x^4 + (2-x)^4]\Big\}.$$

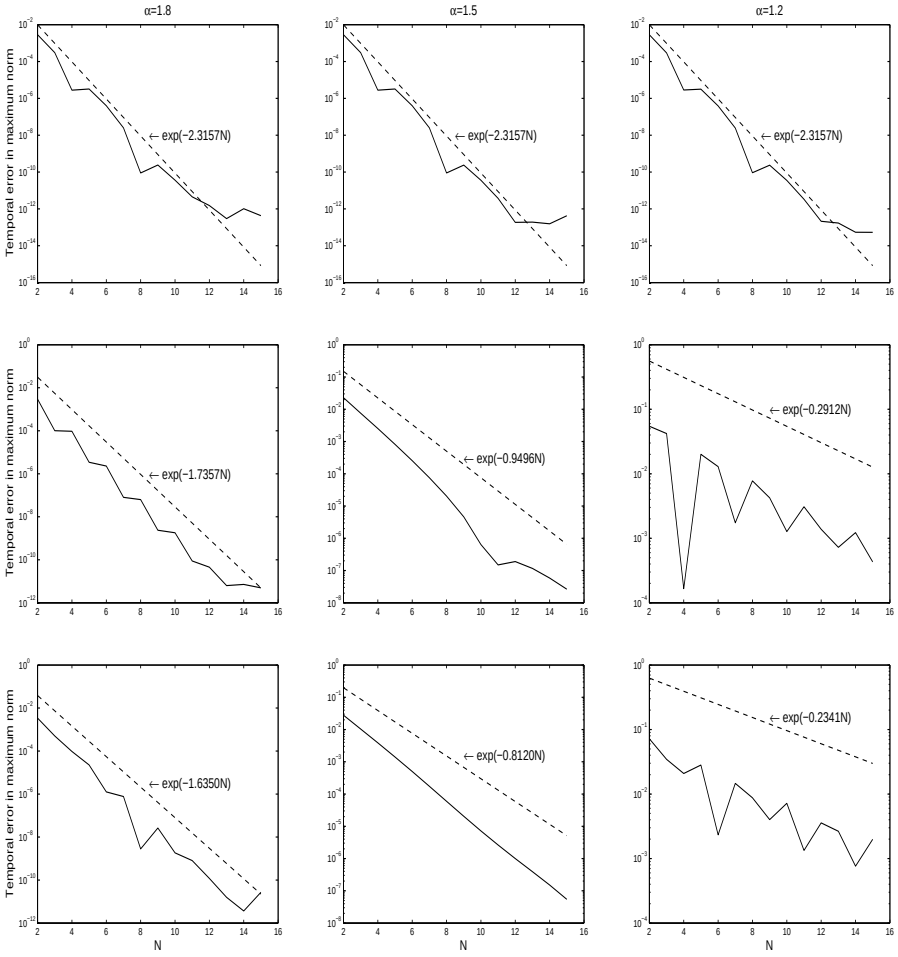

Fig. 1 Convergence history of the contour integral method. Solid curves represent actual error and dash lines refer to their theoretical estimates. The first row is for $c_1 = 1$ and $c_2 = 1$, the second row is for $c_1 = 1$ and $c_2 = 10$, and the third row is for $c_1 = 1$ and $c_2 = 100$. The first column, second column, and third column are corresponding to $\alpha = 1.8$, $\alpha = 1.5$, and $\alpha = 1.2$, respectively.

The exact solution is

$$u(x,t) = 4e^{-t}x^2(2-x)^2.$$

For this example, the diffusion coefficients are not proportional to each other anymore. As mentioned in Section 3.2, there is no similar result as in Theorems 3.2 and 3.3 for such fractional diffusion equation. In practical computations, we replace the variable coefficients by their means. Then it is back to the constant coefficient case and we employ Theorems 3.2 and 3.3 again to determine the parameters in the contour integral method. Table 2 gives the numerical results of solving Example 2 by the numerical contour integral method and the time-stepping method. Clearly, the contour integral method still works very well and outperforms

the time-stepping method both from the number of the linear system solve and the CPU time.

Table 2 Numerical results for Example 2.

	contour integral				time-stepping			
n	N	avegiter	CPU	Error	M	avegiter	CPU	Error
64	3	11.7	0.020	1.06E-2	64	8.0	0.187	1.47E-2
128	3	12.7	0.018	4.35E-3	128	7.0	0.355	7.00E-3
256	4	13.0	0.028	2.85E-3	256	7.0	0.834	3.41E-3
512	4	13.0	0.033	1.41E-3	512	6.0	2.121	1.68E-3
1024	4	13.0	0.051	7.11E-4	1024	5.0	5.572	8.37E-4

Next, we consider two dimensional fractional diffusion equations.

Example 3. Consider the fractional diffusion equation

$$\begin{aligned}
&\frac{\partial u(x,y,t)}{\partial t} = d_1 \frac{\partial^\alpha u(x,y,t)}{\partial_+ x^\alpha} + d_2 \frac{\partial^\alpha u(x,y,t)}{\partial_- x^\alpha} + d_3 \frac{\partial^\beta u(x,y,t)}{\partial_+ y^\beta} + d_4 \frac{\partial^\beta u(x,y,t)}{\partial_- y^\beta} + f(x,y,t),\\
&u(0,y,t) = u(x,0,t) = u(2,y,t) = u(x,2,t) = 0, \qquad t \in [0,1],\\
&u(x,y,0) = 16x^2(2-x)^2y^2(2-y)^2, \qquad (x,y) \in [0,2]\times[0,2],
\end{aligned}$$

with source term

$$\begin{aligned}
f(x,y,t) = &-16e^{-t}\left\{y^2(2-y)^2\left[x^2(2-x)^2 + \frac{\Gamma(5)}{\Gamma(5-\alpha)}\left(d_1x^{4-\alpha} + d_2(2-x)^{4-\alpha}\right)\right.\right.\\
&-\frac{\Gamma(5)}{\Gamma(4-\alpha)}\left(d_1x^{3-\alpha} + d_2(2-x)^{3-\alpha}\right) + \frac{4\Gamma(3)}{\Gamma(3-\alpha)}\left(d_1x^{2-\alpha} + d_2(2-x)^{2-\alpha}\right)\Big]\\
&+\left[\frac{\Gamma(5)}{\Gamma(5-\beta)}\left(d_3y^{4-\beta} + d_4(2-y)^{4-\beta}\right) - \frac{\Gamma(5)}{\Gamma(4-\beta)}\left(d_3y^{3-\beta} + d_4(2-y)^{3-\beta}\right)\right.\\
&\left.\left.+\frac{4\Gamma(3)}{\Gamma(3-\beta)}\left(d_3y^{2-\beta} + d_4(2-y)^{2-\beta}\right)\right]x^2(2-x)^2\right\}.
\end{aligned}$$

The exact solution of Example 3 is

$$u(x,y,t) = 16e^{-t}x^2(2-x)^2y^2(2-y)^2.$$

This is a two dimensional fractional diffusion equation with constant coefficients which satisfy the conditions in both Theorems 3.4 and 3.5. So the spectrum of the resulting coefficient matrix is contained in a sectorial region given by (3.22) or (3.25). Consequently, we can determine the parameters of the hyperbolic contour in the contour integral method by substituting the semiangle δ of the sectorial region to the formulas (2.14)–(2.16).

Note that the resulting coefficient matrix in this event is a block Toepltiz matrix with Toeplitz blocks (BTTB) [2]. Then we suggest to use the block circulant matrix with circulant blocks (BCCB) [2] as preconditioners to solve the corresponding resulting linear systems. Numerical results with fractional orders $\alpha = 1.8$, $\beta = 1.6$ and diffusion coefficients $d_1 = 3$, $d_2 = 5$, $d_3 = 2$, $d_4 = 7$ are presented in Table 3. We can see from the table that for this two dimensional case the contour integral method works well and is more efficient than the time-stepping method.

Table 3 Numerical results for Example 3 with $\alpha = 1.8$, $\beta = 1.6$, and $d_1 = 3$, $d_2 = 5$, $d_3 = 2$, $d_4 = 7$.

	contour integral				time-stepping			
(m,n)	N	avegiter	CPU	Error	M	avegiter	CPU	Error
(16,16)	3	15.0	0.040	1.48E-1	8	13.0	0.059	1.12E-1
(32,32)	4	18.0	0.080	7.13E-2	16	15.0	0.223	6.89E-2
(64,64)	4	21.0	0.231	3.77E-2	32	16.0	1.229	3.76E-2
(128,128)	5	24.4	1.414	2.00E-2	64	16.9	10.862	1.96E-2
(256,256)	5	29.0	6.546	9.77E-3	128	17.0	80.164	9.99E-3
(512,512)	6	33.0	41.614	5.63E-3	256	17.0	665.650	5.04E-3

Example 4. Consider the fractional diffusion equation

$$\begin{aligned}\frac{\partial u(x,y,t)}{\partial t} &= d_1(x,y)\frac{\partial^\alpha u(x,y,t)}{\partial_+ x^\alpha} + d_2(x,y)\frac{\partial^\alpha u(x,y,t)}{\partial_- x^\alpha} + d_3(x,y)\frac{\partial^\beta u(x,y,t)}{\partial_+ y^\beta}\\ &\quad + d_4(x,y)\frac{\partial^\beta u(x,y,t)}{\partial_- y^\beta} + f(x,y,t), \qquad (4.2)\\ u(0,y,t) &= u(x,0,t) = u(2,y,t) = u(x,2,t) = 0, \qquad t \in [0,1],\\ u(x,y,0) &= 16x^2(2-x)^2y^2(2-y)^2, \qquad (x,y) \in [0,2]\times[0,2],\end{aligned}$$

with coefficients

$$\begin{aligned} d_1(x,y) &= \Gamma(3-\alpha)(1+x)^\alpha(1+y)^2, \qquad d_2(x,y) = \Gamma(3-\alpha)(3-x)^\alpha(3-y)^2,\\ d_3(x,y) &= \Gamma(3-\beta)(1+x)^2(1+y)^\beta, \qquad d_4(x,y) = \Gamma(3-\beta)(3-x)^2(3-y)^\beta, \end{aligned}$$

and the source term

$$\begin{aligned} f(x,y,t) = &-16e^{-t}\Big\{x^2(2-x)^2y^2(2-y)^2\\ &+\Big[8x^{2-\alpha} - \frac{24x^{3-\alpha}}{3-\alpha} + \frac{24x^{4-\alpha}}{(4-\alpha)(3-\alpha)}\Big](1+x)^\alpha(1+y)^2y^2(2-y)^2\\ &+\Big[8(2-x)^{2-\alpha} - \frac{24(2-x)^{3-\alpha}}{3-\alpha} + \frac{24(2-x)^{4-\alpha}}{(4-\alpha)(3-\alpha)}\Big](3-x)^\alpha(3-y)^2y^2(2-y)^2\\ &+\Big[8y^{2-\beta} - \frac{24y^{3-\beta}}{3-\beta} + \frac{24y^{4-\beta}}{(4-\beta)(3-\beta)}\Big](1+y)^\beta(1+x)^2x^2(2-x)^2\\ &+\Big[8(2-y)^{2-\beta} - \frac{24(2-y)^{3-\beta}}{3-\beta} + \frac{24(2-y)^{4-\beta}}{(4-\beta)(3-\beta)}\Big](3-y)^\beta(3-x)^2x^2(2-x)^2\Big\}. \end{aligned}$$

The exact solution is

$$u(x,y,t) = 16e^{-t}x^2(2-x)^2y^2(2-y)^2.$$

This example is a two dimensional fractional diffusion equation with variable coefficients. Moreover, the coefficients are not proportional to each other. For this situation, we have no similar results as in Theorems 3.4 and 3.5 to be utilized to select the parameters of the hyperbolic contour. Following the suggestion in Section 3.3, in practical computations we first replace the coefficients of (4.2) by their means and then employ Theorem 3.4 (or Theorem 3.5) again to parametrize the hyperbolic contour in the contour integral method.

Note that the resulting coefficient matrix for this example is a BTTB-like matrix with the form given by (3.20) but not an exact BTTB matrix. To solve the resulting linear systems efficiently, we follow the ideas introduced in [10] to design preconditioners. More precisely, we first construct a BTTB matrix $\hat{A}$ by replacing the diagonal matrices D_1, D_2, D_3, D_4 in (3.20) by the means of their corresponding diffusion coefficients, and then construct Strang's BCCB preconditioner [2] for resulting linear systems based on $\hat{A}$.

Numerical results with fractional orders $\alpha = 1.8$ and $\beta = 1.6$ are displayed in Table 4. From Table 4, we see that the contour integral method still works well even for this general diffusion coefficients case which indicates the validity of the strategy for parameter selections. In addition, the average iteration numbers for solving resulting linear systems are not large and increase slowly with respect to the refinement of spatial grid mesh which implies the efficiency of the preconditioners. By comparison of the two methods, we also see that the contour integral method is much superior to the time-stepping method both from the aspects of the number of the linear system solve and the CPU time.

Table 4 Numerical results for Example 4 with $\alpha = 1.8$, $\beta = 1.6$.

	contour integral				time-stepping			
(m, n)	N	avegiter	CPU	Error	M	avegiter	CPU	Error
(16,16)	2	17.0	0.035	9.82E-2	8	15.1	0.052	1.25E-1
(32,32)	3	21.0	0.063	5.40E-2	16	17.9	0.245	5.71E-2
(64,64)	3	25.7	0.221	2.50E-2	32	20.9	1.361	2.74E-2
(128,128)	3	31.7	1.122	1.16E-2	64	24.0	13.202	1.35E-2
(256,256)	4	37.8	8.139	6.65E-3	128	26.0	104.19	6.69E-3
(512,512)	4	46.3	55.252	3.33E-3	256	28.6	1247.40	3.34E-3

5 Concluding remarks

We have employed the numerical contour integral method with hyperbolic contour to solve the space-fractional diffusion equations. The spatial discretized matrices are Toeplitz-like. We have proved by taking advantage of Toeplitz-like structures and its properties that under some conditions, the spectra of resulting matrices are contained in a certain sectorial region. In addition, the resolvent norms of resulting matrices are bounded outside of it. Therefore we can parameterize the hyperbolic contour according to this sectorial region. To reduce the computational cost of the numerical contour integral method, we have utilized the preconditioned GMRES method with circulant preconditioners to solve the resulting shifted linear systems. A variety of fractional diffusion equations have been tested to demonstrate the efficiency of our method.

Nevertheless, we note that when the order of fractional derivatives is closed to 1, the semiangle δ in (3.14), (3.22), and (3.25) will get large, which may lead to a slow convergence rate for the numerical contour integral method in light of (2.16); see also Figure 1. With further research the contour method is likely to be improved for this situation. In addition, we comment that the resulting linear systems arising from the contour integral method are shifted linear systems with

complex shifts. We simply apply the preconditioning strategy proposed in [10] to solve them. How to design a more efficient solver by their special structures also deserves a deeper investigation

Acknowledgements The authors would like to thank the anonymous referees for their valuable comments and suggestions. The first author was supported by the National Natural Science Foundation of China under grant 11201192, the Natural Science Foundation of Jiangsu Province under grant BK2012577, and the Natural Science Foundation for Colleges and Universities in Jiangsu Province under grant 12KJB110004. The second author was supported by research grants 105/2012/A3 from FDCT of Macao and MYRG206(Y3-L4)-FST11-SHW from University of Macau.

References

1. Benson D., Schumer R., Meerschaert M., Wheatcraft S.: Fractional dispersion, Lévy motion, and the MADE tracer tests. Transport Porous Med. 42, 211–240 (2001)
2. Chan R., Jin X.: An Introduction to Iterative Toeplitz Solvers. SIAM, Philadelphia (2007)
3. Gavrilyuk I. P., Makarov V. L.: Exponentially convergent algorithms for the operator exponential with applications to inhomogeneous problems in Banach spaces. SIAM J. Numer. Anal. 43, 2144–2171 (2005)
4. Hilfer R.: Applications of Fractional Calculus in Physics. Word Scientific, Singapore (2000)
5. Horn R., Johnson C.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
6. in 't Hout K. J., Weideman J. A. C.: A contour integral method for the Black-Scholes and Heston equations. SIAM J. Sci. Comput. 33, 763–785 (2011)
7. Kirchner J. W., Feng X., Neal C.: Fractal stream chemistry and its implications for containant transport in catchments. Nature 403, 524–526 (2000)
8. Lee H., Lee J., Sheen D.: Laplace transform method for parabolic problems with time-dependent coefficients. SIAM J. Numer. Anal. 51, 112–125 (2013)
9. Lee S., Pang H., Sun H.: Shift-invert Arnoldi approximation to the Toeplitz matrix exponential. SIAM J. Sci. Comput. 32, 774–792 (2010)
10. Lei S., Sun H.: A circulant preconditioner for fractional diffusion equations. J. Comput. Phys. 242, 715–725 (2013)
11. Li X., Xu C.: Existence and uniqueness of the weak solution of the space-time fractional diffusion equation and a spectral method approximation. Commun. Comput. Phys. 8, 1016–1051 (2010)
12. Lin F., Yang S., Jin X.: Preconditioned iterative methods for fractional diffusion equation. J. Comput. Phys. 256, 109–117 (2014)
13. Liu F., Anh V., Turner I.: Numerical solution of the space fractional Fokker-Planck equation. J. Comput. Appl. Math. 166, 209–219 (2004)
14. López-Fernández M., Palencia C.: On the numerical inversion of the Laplace transform of certain holomorphic mapping. Appl. Numer. Math. 51, 289–303 (2004)
15. López-Fernández M., Palencia C., Schädle A.: A spectral order method for inverting sectorial Laplace transforms. SIAM J. Numer. Anal. 44, 1332–1350 (2006)
16. Magin R. L.: Fractional Calculus in Bioengineering. Begell House Publishers (2006)
17. Martensen E.: Zur numerischen Auswertung uneigentlicher Integrale. ZAMM Z. Angew. Math. Mech. 48, T83–T85 (1968)
18. Mclean W., Sloan I. H., Thomée V.: Time discretization via Laplace transformation of an integro-differential equation of parabolic type. Numer. Math. 102, 497–522 (2006)
19. Mclean W., Thomée V.: Time discretization of an evolution equation via Laplace transformation. IMA J. Numer. Anal. 24, 439–463 (2004)
20. Mclean W., Thomée V.: Numerical solution via Laplace transformation of a fractional-order evolution equation. J. Integr. Equ. Appl. 22, 57–94 (2010)
21. Mclean W., Thomée V.: Maximum-norm error analysis of a numerical solution via Laplace transformation and quadrature of a fractional-order evolution equation. IMA J. Numer. Anal. 30, 208–230 (2010)
22. Meerschaert M. M., Tadjeran C.: Finite difference approximations for fractional advection-diffusion flow equations. J. Comput. Appl. Math. 172, 65–77 (2004)

23. Meerschaert M. M., Tadjeran C.: Finite difference approximations for two-sided space-fractional partial differential equations. Appl. Numer. Math. 56, 80–90 (2006)
24. Pan J., Ke R., Ng M., Sun H.: Preconditioning techniques for diagonal-times-Toeplitz matrices in fractional diffusion equations. SIAM J. Sci. Comput. 36, A2698–A2719 (2014)
25. Pang H., Sun H.: Multigrid method for fractional diffusion equations. J. Comput. Phys. 231, 693–703 (2012)
26. Podlubny I.: Fractional Differential Equations. Academic Press, New York (1999)
27. Qu W., Lei S., Vong S.: Circulant and skew-circulant splitting iteration for fractional advection-diffusion equations. Int. J. Comput. Math. 91, 2232–2242 (2014)
28. Raberto M., Scalas E., Mainardi F.: Waiting-times and returns in high-frequency financial data: an empirical study. Physica 314, 749–755 (2002)
29. Rockafellar R.: Convex Analysis. Princeton University Press, Princeton, NJ (1970)
30. Shen S., Liu F.: Error analysis of an explicit finite difference approximation for the space fractional diffusion. ANZIAM J. 46, 871–887 (2005)
31. Sheen D., Sloan I. H., Thomée V.: A parallel method for time discretization of parabolic problems based on contour integral representation and quadrature. Math. Comput. 69, 177–195 (1999)
32. Sheen D., Sloan I. H., Thomée V.: A parallel method for time discretization of parabolic equations based on Laplace transformation and quadrature. IMA. J. Numer. Anal. 23, 269–299 (2003)
33. Spijker M. N.: Numerical ranges and stability estimates. Appl. Numer. Math. 13, 241–249 (1993)
34. Tadjeran C., Meerschaert M. M., Scheffler H. P.: A second-order accurate numerical approximation for the fractional diffusion equation. J. Comput. Phys. 213, 205–213 (2006)
35. Tadjeran C., Meerschaert M. M.: A second-order accurate numerical method for the two-dimensional fractional diffusion equation, J. Comput. Phys. 220, 813–823 (2007)
36. Talbot A.: The accurate numerical inversion of Laplace transforms. J. Inst. Math. Appl. 23, 97–120 (1979)
37. Tilli P.: Singular values and eigenvalues of non-Hermitian block Toeplitz matrices. Linear Algebra Appl. 272, 59–89 (1998)
38. Wang H., Basu T.: A fast finite difference method for two-dimensional space-fractional diffusion equations. SIAM J. Sci. Comput. 34, A2444–A2458 (2012)
39. Wang H., Wang K.: An $O(N \log^2 N)$ alternating-direction finite difference method for two-dimensional fractional diffusion equations. J. Comput. Phys. 230, 7830–7839 (2011)
40. Wang H., Wang K., Sircar T.: A direct $O(N \log^2 N)$ finite difference method for fractional diffusion equations. J. Comput. Phys. 229, 8095–8104 (2010)
41. Weideman J. A. C.: Improved contour integral methods for parabolic PDEs. IMA J. Numer. Anal. 30, 334–350 (2010)
42. Weideman J. A. C., Trefethen L. N.: Parabolic and hyperbolic contours for computing the Bromwich integral. Math. Comput. 76, 1341–1356 (2007)
43. Zhou H., Tian W., Deng W.: Quasi-compact finite difference schemes for space fractional diffusion equations. J. Sci. Comput. 56, 45–66 (2013)