

Bernstein dual-Petrov–Galerkin method: application to 2D time fractional diffusion equation

M. Jani¹ · S. Javadi¹ · E. Babolian¹ · D. Bhatta²

Received: 15 February 2017 / Revised: 22 April 2017 / Accepted: 7 May 2017 /

Published online: 15 May 2017

© SBMAC - Sociedade Brasileira de Matemática Aplicada e Computacional 2017

Abstract In this paper, we develop a Bernstein dual-Petrov–Galerkin method for the numerical simulation of a two-dimensional fractional diffusion equation. A spectral discretization is applied by introducing suitable combinations of dual Bernstein polynomials as the test functions and the Bernstein polynomials as the trial ones. We derive the exact sparse operational matrix of differentiation for the dual Bernstein basis which provides a matrix-based approach for the spatial discretization. It is shown that the method leads to banded linear systems that can be solved efficiently. The stability and convergence of the proposed method is discussed. Finally, some numerical examples are provided to support the theoretical claims and to show the accuracy and efficiency of the method.

Keywords Fractional PDEs · Bernstein polynomials · 2D subdiffusion · Dual-Petrov–Galerkin · Dual Bernstein basis · Operational matrix

Mathematics Subject Classification 41A10 · 65M22 · 35R11 · 76M22

Communicated by José Tenreiro Machado.

✉ M. Jani
mostafa.jani@gmail.com

S. Javadi
javadi@khu.ac.ir

E. Babolian
babolian@khu.ac.ir

D. Bhatta
dambaru.bhatta@utrgv.edu

¹ Department of Mathematics, Faculty of Mathematical Sciences and Computer, Kharazmi University, Tehran, Iran

² School of Mathematical and Statistical Sciences, The University of Texas Rio Grande Valley, 1201 West University Drive, Edinburg, TX, USA

1 Introduction

Bernstein polynomial basis plays an important role in computer graphics for geometric modeling, curve and surface approximation. Some interesting features have been investigated for this basis in the last decades; for instance, it is proven to be an optimal stable basis among nonnegative bases in a sense described in [Farouki and Goodman \(1996\)](#). Also, it provides some optimal shape-preserving features ([Carnicer and Pena 1993](#)). We refer to [Farin et al. \(2002\)](#), [Farouki and Rajan \(1988\)](#), [Farouki \(1991\)](#) for detailed properties and applications in computer-aided geometric design (CAGD).

Bernstein basis has also been used for the numerical solution of differential, integral, integro-differential and fractional differential equations; see e.g., [Behiry \(2014\)](#), [Javadi et al. \(2016a\)](#), [Javadi et al. \(2016b\)](#), [Maleknejad et al. \(2012\)](#), [Saadatmandi \(2014\)](#) and the references therein. However, it is not orthogonal and so leads to dense linear systems when using in numerical methods. Some numerical approaches implement the orthogonalized Bernstein basis. However, as we will see in the next section, it fails to keep some interesting properties of the Bernstein basis. Another approach uses the dual Bernstein polynomials (DBP) introduced by [Juttler \(1998\)](#). To the best of our knowledge, the DBP basis has been only discussed from the CAGD point of view (see the works of Lewanowicz and Wozny e.g. [Lewanowicz and Wozny 2011](#), [Wozny and Lewanowicz 2009](#)). So, it is of interest to explore some new aspects of this basis to facilitate the numerical methods for differential equations that are based on Bernstein polynomials and to present a method for time fractional diffusion equation in two dimensions.

Fractional partial differential equations (FPDEs) have been widely used for the description of some important physical phenomena in many applied fields including viscoelastic materials, control systems, polymer, electrical circuits, continuum and statistical mechanics; see, for instance [Dabiri and Butcher \(2017\)](#), [Goychuk \(2009\)](#), [Metzler and Klafter \(2004\)](#), [Moghaddam and Machado \(2017\)](#), 2016 and the references therein. The subdiffusion equation is a FPDE describing the behavior of anomalous diffusive systems with the probability density of particles diffusing proportional to the mean square displacement $\chi^2(t) \propto t^\alpha$ with $0 < \alpha < 1$ ([Gao et al. 2012](#)). Anomalous diffusion equations have been used for modeling transport dynamics, especially the continuous time random walk, the contamination in porous media, viscoelastic diffusion, etc ([Gao et al. 2012](#), 2015; [Goychuk 2009](#); [Metzler and Klafter 2004](#); [Wang and Wang 2016](#)). For the numerical solution of the one-dimensional problem, we refer to [Jin et al. \(2016\)](#), [Ren et al. \(2013\)](#), [Stokes et al. \(2015\)](#), [Zhou and Gong \(2016\)](#) and the references therein. Some classic numerical methods for PDEs have been developed for the simulation of two-dimensional subdiffusion equation, for example the finite difference schemes ([Gao et al. 2015](#); [Pang and Sun 2016](#); [Ren et al. 2013](#)), meshless methods ([Shirzadi et al. 2012](#); [Yang et al. 2015](#)), finite element method ([Zhao et al. 2016](#)) and alternating direction implicit methods ([Zhang 2011](#); [Zhang et al. 2012](#)).

In this paper, deriving some new aspects of DBPs, we present suitable combinations of these functions to develop a dual-Petrov–Galerkin method for solving the following 2D subdiffusion equation ([Wang and Wang 2016](#); [Yang et al. 2015](#); [Yang and Zhang 2014](#); [Zhang 2011](#); [Zhang et al. 2012](#); [Zhao et al. 2016](#)).

$$D_t^\alpha u(x, y, t) = \kappa \Delta u(x, y, t) + S(x, y, t), \quad (x, y, t) \in \Omega \times (0, T], \quad (1.1)$$

with the following initial and boundary conditions:

$$u(x, y, 0) = g(x, y), \quad (x, y) \in \Omega, \quad (1.2)$$

$$u(x, y, t) = 0, \quad (x, y, t) \in \partial\Omega \times (0, T], \quad (1.3)$$

where $\Omega = (0, 1)^2 \subset \mathbb{R}^2$, Δ is the Laplacian operator, $T > 0$, κ is the diffusion coefficient and S is the source function. Here, $D_t^\alpha u$ denotes the Caputo fractional derivative of order α , $0 < \alpha < 1$, with respect to t defined as

$$D_t^\alpha u(x, t) = \frac{1}{\Gamma(1 - \alpha)} \int_0^t \frac{1}{(t - s)^\alpha} \frac{\partial u(x, s)}{\partial s} ds, \quad 0 < \alpha < 1. \tag{1.4}$$

The main contribution of our work is the development of an accurate Bernstein dual-Petrov–Galerkin method and the application for the numerical simulation of the 2D subdiffusion equation. It is shown that the method leads to sparse linear systems. To give a matrix approach of the method, we present some results concerning the DBPs including a recurrence formula for the derivative, constructing a new basis using DBPs, deriving the operational matrix of differentiation and also providing the transformation matrices between the DBPs and the new basis.

The paper is organized as follows: Sect. 2 presents some new aspects of DBPs and provides modal basis functions and the associated transformation matrices between the bases. Section 3 is devoted to the Bernstein spectral formulation of the subdiffusion problem (1.1)–(1.3) and the stability and convergence results are discussed in Sect. 4. Numerical examples are provided in Sect. 5. The paper ends with some concluding remarks in Sect. 6.

2 Bernstein polynomials and DBPs

The Bernstein polynomials with degree N on the unit interval are defined by

$$\phi_i(x) = \binom{N}{i} x^i (1 - x)^{N-i}, \quad 0 \leq i \leq N.$$

The set $\{\phi_i(x) : i = 0, \dots, N\}$ forms a basis for \mathbb{P}_N , with the space of polynomials of degree not exceeding N .

These polynomials possess end-point interpolation property, i.e.,

$$\phi_i(0) = \delta_{i,0}, \quad \phi_i(1) = \delta_{i,N}, \quad 0 \leq i \leq N, \quad N > 0. \tag{2.1}$$

Also, the summation is one and the integral over the unit interval is constant, namely

$$\sum_{i=0}^N \phi_i(x) \equiv 1, \quad \int_0^1 \phi_i(x) dx = \frac{1}{N + 1}, \quad i = 0, 1, \dots, N. \tag{2.2}$$

The derivative enjoys the three-term recurrence relation (Jani et al. 2017)

$$\phi_i'(x) = (N - i + 1) \phi_{i-1}(x) - (N - 2i) \phi_i(x) - (i + 1) \phi_{i+1}(x), \quad 0 \leq i \leq N, \tag{2.3}$$

where we adopt the convention that $\phi_i(x) \equiv 0$ for $i < 0$ and $i > N$.

As we mentioned in the preceding section, the Bernstein basis is not orthogonal. The corresponding orthogonalized basis, obtained e.g., by the Gram–Schmidt process fails to keep some interesting aspects of the original basis. We will not consider this basis in the present work. Instead, we turn to the dual basis.

The DBPs are defined as

$$\tilde{\psi}_i(x) = \sum_{j=0}^N c_{i,j} \phi_j(x), \tag{2.4}$$

with the coefficients given by

$$c_{i,j} = \frac{(-1)^{i+j}}{\binom{N}{i}\binom{N}{j}} \sum_{r=0}^{\min(i,j)} (2r+1) \binom{N+r+1}{N-i} \binom{N-r}{N-i} \binom{N+r+1}{N-j} \binom{N-r}{N-j}. \tag{2.5}$$

It is verified that they satisfy the biorthogonal system (Juttler 1998, Theorem 3)

$$\int_0^1 \phi_i(x) \tilde{\psi}_j(x) dx = \delta_{ij}, \quad 0 \leq i, j \leq N. \tag{2.6}$$

It is worth noting that less than a quarter of the entries of transformation matrix between the Bernstein and dual Bernstein basis $C = [c_{i,j}]$, are to be computed directly by (2.5), for it is bisymmetric, i.e., symmetric about both of the main diagonal and antidiagonal.

Another property which is used later is that the sum of the entries for each row (column) is equal to the order of the matrix, i.e.,

$$\sum_{i=0}^N c_{i,j} = \sum_{j=0}^N c_{i,j} = N + 1. \tag{2.7}$$

In the next lemma, we present some properties of the DBPs.

Lemma 1 *Let N be a nonnegative integer. The following statements hold.*

- (i) For all $x \in [0, 1]$, $\tilde{\psi}_{N-i}(x) = \tilde{\psi}_i(1-x)$, $0 \leq i \leq N$.
- (ii) For all $x \in [0, 1]$, $\sum_{i=0}^N \tilde{\psi}_i(x) = N + 1$.
- (iii) The basis functions have the same definite integral, i.e., $\int_0^1 \tilde{\psi}_i(x) dx = 1$, $0 \leq i \leq N$.

Proof The first statement is an immediate consequence of the similar formula for Bernstein polynomials, i.e., $\phi_{N-i}(x) = \phi_i(1-x)$. From (2.4), (2.7) and (2.2), we have

$$\begin{aligned} \sum_{i=0}^N \tilde{\psi}_i(x) &= \sum_{i=0}^N \sum_{j=0}^N c_{i,j} \phi_j(x) \\ &= \sum_{j=0}^N \phi_j(x) \sum_{i=0}^N c_{i,j} = N + 1. \end{aligned}$$

Statement (iii) is also verified similarly. □

The property (i) implies that $\tilde{\psi}_i$, for $[\frac{N}{2}] + 1 \leq i \leq N$, need not to be computed directly by (2.4)–(2.5). It especially gives $\tilde{\psi}_i(0) = \tilde{\psi}_{N-i}(1)$.

2.1 Modal basis functions

One may choose a suitable compact combinations of orthogonal polynomials as the trial and test basis for the Galerkin and Petrov–Galerkin methods for BVPs in such a way leading to sparse linear systems for some special problems (see e.g., Goubet and Shen 2007; Yuan et al. 2008). Here, we use this idea for the non-orthogonal Bernstein polynomials to present a simple and accurate dual-Petrov–Galerkin spectral method for two-dimensional subdiffusion

equation. Following Shen et al. (2007, 2011, Section 1.3), we will refer to such basis functions as the modal basis functions.

Proposition 1 *Let $N \geq 2$ be an integer, $\{\tilde{\psi}_i : 0 \leq i \leq N\}$ be the dual Bernstein basis and $\mathbb{P}_N^0 = \{u \in \mathbb{P}_N : u(0) = 0, u(1) = 0\}$. Set*

$$\psi_i(x) = \tilde{\psi}_i(x) + a_i \tilde{\psi}_{i+1}(x) + b_i \tilde{\psi}_{i+2}(x), \tag{2.8}$$

for $0 \leq i \leq N - 2$, where

$$\begin{aligned} a_i &= \frac{2i + 4}{N - i + 1}, \\ b_i &= \frac{(i + 2)(i + 3)}{(N - i)(N - i + 1)}. \end{aligned} \tag{2.9}$$

Then, the polynomials $\tilde{\psi}_i(x)$ vanish at 0 and 1, so the set $\{\psi_i(x)\}_{i=0}^{N-2}$ forms a basis for \mathbb{P}_N^0 .

Proof By (2.4) and (2.1), we have

$$\tilde{\psi}_i(0) = \sum_{j=0}^N c_{i,j} \phi_j(0) = c_{i,0} = (-1)^i (N + 1) \binom{N + 1}{i + 1}. \tag{2.10}$$

From Lemma 1, we infer that

$$\tilde{\psi}_i(1) = \tilde{\psi}_{N-i}(0) = (-1)^{N-i} (N + 1) \binom{N + 1}{i}. \tag{2.11}$$

By (2.10) and (2.11), we obtain

$$\begin{aligned} \psi_i(0) &= (-1)^i (N + 1) \left(\binom{N + 1}{i + 1} - a_i \binom{N + 1}{i + 2} + b_i \binom{N + 1}{i + 3} \right) = 0, \\ \psi_i(1) &= (-1)^{N-i} (N + 1) \left(\binom{N + 1}{N - i + 1} - a_i \binom{N + 1}{N - i} + b_i \binom{N + 1}{N - i - 1} \right) = 0, \end{aligned}$$

for $0 \leq i \leq N - 2$. It is easy to see that $\{\psi_i(x)\}_{i=0}^{N-2}$ is linearly independent. Since $\dim \mathbb{P}_N^0 = N - 1$, this set is a basis for \mathbb{P}_N^0 . This completes the proof. \square

Figure 1 illustrates the DBPs and the modal basis functions for $6 \leq N \leq 8$. It is seen that the modal basis functions have less values than the corresponding dual functions on the unit interval, expecting less round-off errors.

2.2 Transformation matrices and the operational matrix for derivatives

For $N \geq 2$, consider the $(N + 1)$ -vector $\tilde{\Psi}$ and the $(N - 1)$ -vector Ψ consisting of dual functions given by (2.4) and the modal basis functions given by (2.8), respectively:

$$\tilde{\Psi}(\cdot) = [\tilde{\psi}_i(\cdot) : 0 \leq i \leq N]^T, \tag{2.12}$$

$$\Psi(\cdot) = [\psi_i(\cdot) : 0 \leq i \leq N - 2]^T. \tag{2.13}$$

For simplicity, we ignore the dependence of the vectors on the variable. First, note that

$$\Psi = \mathbf{G} \tilde{\Psi}, \tag{2.14}$$

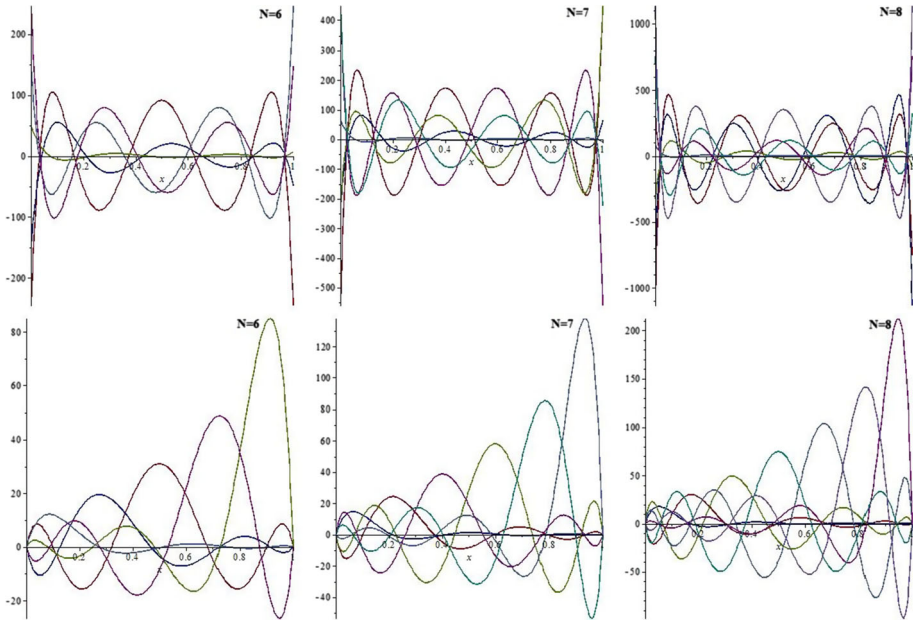


Fig. 1 Graphs of DBPs $\{\tilde{\psi}_i(x), 0 \leq i \leq N\}$ (top) and the modal basis functions $\{\psi_i(x), 0 \leq i \leq N - 2\}$ (bottom)

where $\mathbf{G} = [g_{i,j}]$ is an $(N - 1) \times (N + 1)$ matrix with three diagonals as

$$g_{i,j} = \begin{cases} 1, & i - j = 0, \\ a_i, & j = i + 1, \\ b_i, & j = i + 2, \end{cases} \quad 0 \leq i \leq N - 1, 0 \leq j \leq N.$$

To derive a formula for the derivative of the modal basis functions, we first prove the following result.

Lemma 2 *The operational matrix for derivative of the DBPs, \mathbf{P} satisfies*

$$\tilde{\Psi}' = \mathbf{P}\tilde{\Psi}, \tag{2.15}$$

where the matrix $\mathbf{P} = [p_{i,j} : 0 \leq i, j \leq N]$ is given by

$$p_{i,j} = \begin{cases} -(-1)^i(N + 1)\binom{N+1}{i+1} + N\delta_{i,0} + \delta_{i,1}, & j = 0, \\ -PN-i,0 & j = N, \\ i, & j = i - 1, \quad j \neq 0, \\ N - 2i, & j = i, \quad j \neq 0, N, \\ -N + i, & j = i + 1, \quad j \neq N. \end{cases}$$

Proof The DBPs $\tilde{\Psi}$ is a basis for P_N , so we expand $\tilde{\psi}'_i(x)$ for $0 \leq i \leq N$ as

$$\tilde{\psi}'_i(x) = \sum_{j=0}^N p_{i,j}\tilde{\psi}_j(x).$$

Integration by parts and (2.3) imply that

$$\begin{aligned}
 p_{i,j} &= \int_0^1 \tilde{\psi}'_i(x) \phi_j(x) dx \\
 &= \tilde{\psi}_i(1) \delta_{j,N} - \tilde{\psi}_i(0) \delta_{j,0} - \int_0^1 \tilde{\psi}_i(x) ((N-j+1)\phi_{j-1}(x) \\
 &\quad - (N-2j)\phi_j(x) - (j+1)\phi_{j+1}(x)) dx.
 \end{aligned}$$

The biorthogonality (2.6) gives

$$p_{i,j} = \tilde{\psi}_i(1) \delta_{j,N} - \tilde{\psi}_i(0) \delta_{j,0} - ((N-j+1)\delta_{i,j-1} - (N-2j)\delta_{i,j} - (j+1)\delta_{i,j+1}).$$

Now, the result is proved by considering (2.10) and (2.11). □

Remark 1 The matrix \mathbf{P} is a sparse matrix of order $N + 1$ with $p_{i,j} = 0$ for $|i - j| > 1$, $j \neq 0, N$; for instance, see the matrix given below.

Corollary 1 Set $\alpha_{i,0} = -(-1)^i(N + 1)\binom{N+1}{i+1} + N\delta_{i,0} + \delta_{i,1}$ for $0 \leq i \leq N$. Then, from (2.15), we infer that the following five-term recurrence relation is deduced:

$$\begin{aligned}
 \tilde{\psi}'_i(x) &= \alpha_{i,0}\tilde{\psi}_0(x) + (1 - \delta_{i,1})i\tilde{\psi}_{i-1}(x) + (1 - \delta_{i,0})(1 - \delta_{i,N})(N - 2i)\tilde{\psi}_i(x) \\
 &\quad - (1 - \delta_{i,N-1})(N - i)\tilde{\psi}_{i+1}(x) - \alpha_{N-i,0}\tilde{\psi}_N(x),
 \end{aligned}$$

where we set $\tilde{\psi}_i \equiv 0$ for $i < 0$ and $i > N$.

We derive the transformation matrices that map the Bernstein and Chebyshev coefficients.

Now, we derive the transformation matrix that maps the derivative of modal basis functions to DBPs. This facilitates the use of the Galerkin method in the next section. In the following, (p, q) – band matrix stands for a matrix with p and q nonzero diagonals below and above the main diagonal, respectively.

Lemma 3 Let the vectors Ψ and $\tilde{\Psi}$ be defined as in (2.12) and (2.13), respectively. Then,

$$\Psi' = \mathbf{Q}\tilde{\Psi},$$

where \mathbf{Q} is an $(N - 1) \times (N + 1)$, $(1, 3)$ – band matrix given by $\mathbf{Q} = \mathbf{G}\mathbf{P}$.

Proof Combining (2.14) with (2.15) implies $\mathbf{Q} = \mathbf{G}\mathbf{P}$. To prove that \mathbf{Q} is a $(1, 3)$ – band matrix, it is sufficient to show that $q_{i,0} = 0$ for $i > 1$ and $q_{i,N} = 0$ for $i < N - 2$,

$$\begin{aligned}
 q_{i,0} &= (\mathbf{G}\mathbf{P})_{i,0} = p_{i,0} + a_i p_{i+1,0} + b_i p_{i+2,0} \\
 &= -\left(\tilde{\psi}_i(0) + a_i \tilde{\psi}_{i+1}(0) + b_i \tilde{\psi}_{i+2}(0)\right) = -\psi_i(0) = 0,
 \end{aligned}$$

and for $i < N - 2$, by (1)

$$\begin{aligned}
 q_{i,N} &= p_{i,N} + a_i p_{i+1,N} + b_i p_{i+2,N} = \tilde{\psi}_{N-i}(0) + a_i \tilde{\psi}_{N-i-1}(0) + b_i \tilde{\psi}_{N-i-2}(0) \\
 &= \tilde{\psi}_i(1) + a_i \tilde{\psi}_{i+1}(1) + b_i \tilde{\psi}_{i+2}(1) = \psi_i(1) = 0.
 \end{aligned}$$

Note that ψ_i 's vanish at the boundary values according to Proposition 1. The proof is complete.

To see the sparsity of the transformation matrices, **P**, **G** and **Q** for $N = 6$ are shown in the following

$$\mathbf{G} = \begin{bmatrix} 1 & \frac{4}{7} & \frac{1}{7} & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & \frac{2}{5} & 0 & 0 & 0 \\ 0 & 0 & 1 & \frac{8}{5} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & \frac{5}{2} & \frac{5}{2} & 0 \\ 0 & 0 & 0 & 0 & 1 & 4 & 7 \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} -43 & -6 & 0 & 0 & 0 & 0 & 7 \\ 148 & 4 & -5 & 0 & 0 & 0 & -49 \\ -245 & 2 & 2 & -4 & 0 & 0 & 147 \\ 245 & 0 & 3 & 0 & -3 & 0 & -245 \\ -147 & 0 & 0 & 4 & -2 & -2 & 245 \\ 49 & 0 & 0 & 0 & 5 & -4 & -148 \\ -7 & 0 & 0 & 0 & 0 & 6 & 43 \end{bmatrix},$$

$$\mathbf{Q} = \begin{bmatrix} \frac{46}{7} & -\frac{24}{7} & -\frac{18}{7} & -\frac{4}{7} & 0 & 0 & 0 \\ 1 & 6 & -\frac{9}{5} & -4 & -\frac{6}{5} & 0 & 0 \\ 0 & 2 & \frac{34}{5} & 0 & -\frac{34}{5} & -2 & 0 \\ 0 & 0 & 3 & 10 & \frac{9}{2} & -15 & -\frac{5}{2} \\ 0 & 0 & 0 & 4 & 18 & 24 & -46 \end{bmatrix}.$$

3 Variational formulation of the problem (1.1) and the spectral discretization

In this section, at first the problem (1.1)–(1.3) is discretized in time. Then we develop a matrix approach Bernstein dual-Petrov–Galerkin method using the results of the previous section.

3.1 Time discretization

Consider the subdiffusion Eq. (1.1) at $t = t_{k+1}$, $k \geq 0$ as

$$D_t^\alpha u(x, y, t_{k+1}) = \kappa \Delta u(x, y, t_{k+1}) + S(x, y, t_{k+1}). \tag{3.1}$$

Let u^k be an approximation of u at $t = t_k = k\tau$ for $k = 0, 1, \dots, M$, where $\tau = \frac{T}{M}$ is the time step length. The time fractional derivative can be approximated by definition (1.4) and using forward difference for the derivative inside as

$$D_t^\alpha u(x, y, t_{k+1}) = \mu \left(u^{k+1} - (1 - b_1)u^k - \sum_{j=1}^{k-1} (b_j - b_{j+1})u^{k-j} - b_k u^0 \right) + r_\tau^{k+1}, \quad k \geq 1, \tag{3.2}$$

where $\mu = \frac{1}{\tau^\alpha \Gamma(2-\alpha)}$ and $b_j = (j + 1)^{1-\alpha} - j^{1-\alpha}$ for $k \geq 0$ and $0 \leq j \leq k$. The error is bounded by

$$|r_\tau^{k+1}| \leq \tilde{c}_u \tau^{2-\alpha}, \tag{3.3}$$

where the coefficient \tilde{c}_u depends only on u (Deng 2008). The time discretization (3.2) is referred to as L1 approximation (see e.g., Deng 2008; Ramezani et al. 2015). Substituting from (3.2) into (3.1) and multiplying both sides by $\tau^\alpha \Gamma(2 - \alpha)$ and dropping (x, y) , the following time-discrete scheme is obtained:

$$\begin{aligned}
 u^{k+1} - \alpha_0 \Delta u^{k+1} &= f^{k+1}, \quad k \geq 0, \\
 f^{k+1} &:= (1 - b_1)u^k + \sum_{j=1}^{k-1} (b_j - b_{j+1})u^{k-j} + b_k u^0 + \frac{1}{\mu} S^{k+1},
 \end{aligned}
 \tag{3.4}$$

with $\alpha_0 = \frac{k}{\mu}$ and $u^0 = g$ is given by the initial condition (1.2) with the error

$$r^{k+1} \leq \tau^\alpha \Gamma (2 - \alpha) \left| r_\tau^{k+1} \right| \leq \tilde{c}_u \tau^2. \tag{3.5}$$

For $k = 0$, it reads as

$$u^1 - \alpha_0 \kappa \Delta u^1 = (1 - b_1)u^1 + b_1 u^0 + \frac{1}{\mu} S^1. \tag{3.6}$$

The boundary conditions for the semidiscrete problem is $u^{k+1} = 0$ on $\partial\Omega$.

3.2 Weak and spectral formulation

Consider the problem (3.4) with $\Omega = I^2$, $I = (0, 1)$ and the homogeneous Dirichlet boundary conditions $u^{k+1}|_{\partial\Omega} = 0$. We seek an approximate solution in the Sobolev space $H_0^1(\Omega) = \{u \in H^1(\Omega), u = 0, \text{ on } \partial\Omega\}$. A weak formulation of the problem (3.4) is to find $u^{k+1} \in H_0^1(\Omega)$ such that $\forall v \in H_0^1(\Omega)$:

$$(u^{k+1}, v) + \alpha_0 (\nabla u^{k+1}, \nabla v) = \left((1 - b_1)u^k + \sum_{j=1}^{k-1} (b_j - b_{j+1})u^{k-j} + b_k u^0 + \frac{1}{\mu} S^{k+1}, v \right). \tag{3.7}$$

Let \mathbb{P}_N be the space of polynomials over I with degree not exceeding N and $(\mathbb{P}_N^0)^2 = \{v \in (\mathbb{P}_N)^2 : v = 0, \text{ on } \partial\Omega\}$.

The Galerkin formulation of the (3.7) is to find $u_N^{k+1} \in (\mathbb{P}_N^0)^2$ such that $\forall v_N \in (\mathbb{P}_N^0)^2$:

$$\begin{aligned}
 &(u_N^{k+1}, v_N) + \alpha_0 (\nabla u_N^{k+1}, \nabla v_N) \\
 &= \left((1 - b_1)u_N^k + \sum_{j=1}^{k-1} (b_j - b_{j+1})u_N^{k-j} + b_k u_N^0 + \frac{1}{\mu} I_N S^{k+1}, v_N \right),
 \end{aligned}
 \tag{3.8}$$

with (f, g) being the standard inner product and I_N an interpolation operator.

3.2.1 Bernstein dual-Petrov–Galerkin method

Since $\dim \mathbb{P}_N^0 = N - 1$, and due to (2.1), we choose a basis for it by removing the first and last Bernstein polynomials of degree N , i.e.,

$$\Phi = [\phi_i(x) : 1 \leq i \leq N - 1]^T. \tag{3.9}$$

Using (2.3), it is easy to verify that

$$\Phi' = \mathbf{D}\Phi + \mathbf{d}, \tag{3.10}$$

where $\mathbf{D} = \text{tridiag}(N - i + 1, 2i - N, -(i + 1))$ is a tridiagonal matrix of order $N - 1$ and $\mathbf{d} = N [\phi_0, 0, \dots, 0, -\phi_N]^T$ is an $(N - 1)$ -vector.

Assuming $N_x = N_y = N$, we use the tensor product of the basis functions of \mathbb{P}_0^N as a basis for two-dimensional case, $\{\phi_i(x)\phi_j(y) : 1 \leq i, j \leq N - 1\}$ and consider an approximate solution of (3.4) as

$$u_N^{k+1} = \sum_{i,j=1}^{N-1} \hat{u}_{i,j}^{k+1} \phi_i(x)\phi_j(y) = \Phi^T(x) \mathbf{U}^{k+1} \Phi(y), \quad (x, y) \in \Omega, \tag{3.11}$$

where $\Phi(\cdot) = [\phi_i(\cdot) : 1 \leq i \leq N - 1]^T$ and $\mathbf{U}^{k+1} = [\hat{u}_{i,j}^{k+1}]$. Let us use the following notations.

$$\begin{aligned} a_{i,j} &= \int_I \phi_j'(x)\psi_i'(x)dx, \quad \mathbf{A} = [a_{i,j}], \\ b_{i,j} &= \int_I \phi_j(x)\psi_i(x)dx, \quad \mathbf{B} = [b_{i,j}], \\ f_{i,j}^{k+1} &= \int_{\Omega} I_N f^{k+1}(x, y)\psi_j(x)\psi_i(y)d\Omega, \end{aligned} \tag{3.12}$$

for $1 \leq j \leq N - 1, 0 \leq i \leq N - 2$.

Taking the test functions of (3.8) as $v = \psi_l(x)\psi_m(y)$ for $l, m = 0, 1, \dots, N - 2$, it is seen that the spectral form (3.8) is equivalent to the following linear system:

$$\mu_\tau^\alpha \mathbf{B}\mathbf{U}^{k+1}\mathbf{B}^T + \kappa (\mathbf{A}\mathbf{U}^{k+1}\mathbf{B}^T + \mathbf{B}\mathbf{U}^{k+1}\mathbf{A}^T) = \mathbf{F}^{k+1}, \quad k \geq 0,$$

which can be equivalently written as a Sylvester equation, but it requires computing the inverse of \mathbf{B} . Although \mathbf{B} has only three nonzero diagonals, it can be shown that its inverse is a dense matrix and so we avoid transforming to the Sylvester equation. Instead, we use the equivalent tensor product form:

$$(\mu_\tau^\alpha \mathbf{B} \otimes \mathbf{B} + \kappa (\mathbf{B} \otimes \mathbf{A} + \mathbf{A} \otimes \mathbf{B})) \mathbf{u}^{k+1} = \mathbf{f}^{k+1}, \tag{3.13}$$

with $\mathbf{f} = [f_{1,0,\dots}, f_{q,0}; f_{1,1,\dots}, f_{q,1}; \dots; f_{1,q-1,\dots}, f_{q,q-1}]^T, q = N - 1$. It is worth noting that the coefficient matrix of the linear system (3.13) is the same for all time steps, so it is to be evaluated just once for all $k \geq 0$.

In terms of the trial vector (3.9), and test vector (2.13), we may write

$$\mathbf{A} = \int_I \Psi' \Phi'^T dx, \quad \mathbf{B} = \int_I \Psi \Phi^T.$$

To facilitate the computations, in what follows, these matrices are related to the transformation matrices introduced in Sect. 2.2. First, note that by the biorthogonality (2.6), we have

$$\int_I \tilde{\Psi} \Phi^T dx = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix} =: \tilde{\mathbf{I}}. \tag{3.14}$$

Now from (2.14), and writing \mathbf{G} as $\mathbf{G} = [\mathbf{g}_0, \mathbf{g}_1, \dots, \mathbf{g}_N]$, we get

$$\mathbf{B} = \int_I \mathbf{G} \tilde{\Psi} \Phi^T dx = \mathbf{G} \tilde{\mathbf{I}} = [\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_{N-1}].$$

So \mathbf{B} is a tridiagonal matrix whose entries are given by

$$b_{i,j} = \begin{cases} 1, & j = i + 1, \\ a_i, & j = i, \\ b_i, & j = i - 1, \\ 0, & \text{otherwise,} \end{cases} \tag{3.15}$$

where a_i s and b_i s are easily computed by (2.9). On the other hand, from Lemma 3, the Bernstein operational matrix of differentiation (3.10) and (3.14), we obtain

$$\begin{aligned} \mathbf{A} &= \int_I \mathbf{Q} \tilde{\Psi} (\Phi^T D^T + \mathbf{d}^T) dx \\ &= \mathbf{Q} \tilde{\mathbf{I}} D^T + N[\mathbf{q}_0, \mathbf{0}, \dots, \mathbf{0}, \mathbf{q}_N] \\ &= [\mathbf{q}_1, \dots, \mathbf{q}_{N-1}] \mathbf{D}^T + N[\mathbf{q}_0, \mathbf{0}, \dots, \mathbf{0}, \mathbf{q}_N], \end{aligned} \tag{3.16}$$

where $\mathbf{Q} = [\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_N]$ is a $(1, 3)$ – band matrix introduced Lemma 3. Hence, $\mathbf{Q} \tilde{\mathbf{I}}$ is a pentadiagonal matrix and \mathbf{A} is the product of a pentadiagonal and a tridiagonal matrix plus a sparse matrix. From Lemma 3 and (3.16), it is seen that \mathbf{A} is a seven-diagonal matrix.

Notice that the solution of linear system (3.13) requires the matrices \mathbf{A} and \mathbf{B} . \mathbf{A} is obtained by a sparse matrix–matrix multiplication (3.16) and entries of \mathbf{B} are given by (3.15).

Remark 2 Since the coefficient matrix of the linear system (3.13) remains intact for a fixed τ , only the RHS vector needs to be computed for different time steps, $k = 0, 1, \dots$ up to a desired time. So it is efficient to use a band-LU factorization for solving the system. It is remarkable that for a $(2p + 1)$ -band matrix, the LU-factorization can be done with $O(Np^2)$ flops and backward substitutions require $O(Np)$ flops Golub and Ortega (1992, Section 4.3).

4 Stability and convergence analysis

For the error analysis, we assume the problem (1.1) to be homogeneous, $S = 0$.

For $\alpha \geq 0$, the bilinear form $a(u, v) = (\nabla u, \nabla v) + \alpha(u, v)$ in (3.8) is continuous and coercive in $H_0^1(\Omega) \times H_0^1(\Omega)$. The existence and uniqueness of the solution for both the weak form (3.8) and the Galerkin form (3.8) is guaranteed by the well-known Lax–Milgram lemma.

We define the following inner product and the associated energy norm on $H_0^1(\Omega)$:

$$(u, v) = \int_{\Omega} uv d\Omega, \quad (u, v)_1 = (u, v) + \alpha_0(\nabla u, \nabla v), \quad \|u\|_1 = (u, u)_1^{\frac{1}{2}}. \tag{4.1}$$

Theorem 1 *The weak form (3.7) is unconditionally stable:*

$$\|u^k\|_1 \leq \|u^0\|, \quad k = 1, \dots, M. \tag{4.2}$$

Proof Let $v = u^1$ in (3.7). Then,

$$(u^1, u^1) + \alpha_0(\nabla u^1, \nabla u^1) = (u^0, u^1),$$

giving (4.2) for $k = 1$, by the definition (4.1), the Schwarz inequality and the inequality $\|v\| \leq \|v\|_1$. By mathematical induction, assume (4.2) holds for $k = 0, \dots, n$. Let $v = u^{n+1}$ in (3.7), i.e.,

$$\begin{aligned} & (u^{n+1}, u^{n+1}) + \alpha_0(\nabla u^{n+1}, \nabla u^{n+1}) \\ &= (1 - b_1)(u^n, u^{n+1}) + \sum_{j=1}^{n-1} (b_j - b_{j+1})(u^{n-j}, u^{n+1}) + b_n(u^0, u^{n+1}). \end{aligned}$$

It is easy to see that the RHS coefficients in (3.4) are positive. So we obtain

$$\begin{aligned} \|u^{n+1}\|_1 &\leq (1 - b_1)\|u^n\| + \sum_{j=1}^{n-1} (b_j - b_{j+1})\|u^{n-j}\| + b_n\|u^0\| \\ &\leq \left((1 - b_1) + \sum_{j=1}^{n-1} (b_j - b_{j+1}) + b_n \right) \|u^0\| = \|u^0\|. \end{aligned}$$

So the proof is done. □

Theorem 2 *Let u be the solution of the equation (1.1) with conditions (1.2)–(1.3) and u^k be the solution of the the semidiscrete problem (3.4). Then,*

$$\|u(t_k) - u^k\|_1 \leq \frac{c_u}{1 - \alpha} T^\alpha \tau^{2-\alpha}, \quad 0 < \alpha < 1, \tag{4.3}$$

$$\|u(t_k) - u^k\|_1 \leq c_u T \tau, \quad \text{as } \alpha \rightarrow 1. \tag{4.4}$$

Proof The idea of the proof comes from Lin and Chuanju (2007). We first prove

$$\|u(t_k) - u^k\|_1 \leq \frac{c_u}{b_{k-1}} \tau^2, \quad k = 1, \dots, M. \tag{4.5}$$

By (1.1) and (3.6), we have

$$(e^1, v) + \alpha_0(\nabla e^1, \nabla v) = (e^0, v) + (r^1, v), \quad \forall v \in H_0^1(\Omega),$$

in which $e^k := u(t_k) - u^k$. For $v = e^1$ and by using $e^0 = 0, \|v\| \leq \|v\|_1$ and (3.5), we get

$$\|e^1\|_1 \leq c_u \tau^2, \tag{4.6}$$

i.e., (4.5) holds for $k = 1$. By induction, assume (4.5) holds for $k \leq n$. Using (1.1) and (3.4), we get

$$\begin{aligned} & (e^{n+1}, v) + \alpha_1(\nabla e^{n+1}, \nabla v) \\ &= (1 - b_1)(e^n, v) + \sum_{j=1}^{n-1} (b_j - b_{j+1})(e^{n-j}, v) + b_n(e^0, v) + (r^{n+1}, v), \quad \forall v \in H_0^1(\Omega). \end{aligned}$$

For $v = e^{n+1}$, it reads as

$$\begin{aligned} \|e^{n+1}\|_1^2 &\leq (1 - b_1)\|e^n\| \|e^{n+1}\|_1 + \sum_{j=1}^{n-1} (b_j - b_{j+1})\|e^{n-j}\| \|e^{n+1}\|_1 + \|r^{n+1}\| \|e^{n+1}\|_1, \\ \Rightarrow \|e^{n+1}\|_1 &\leq (1 - b_1) \frac{c_u}{b_{n-1}} \tau^2 + \sum_{j=1}^{n-1} (b_j - b_{j+1}) \frac{c_u}{b_{n-j-1}} \tau^2 + c_u \tau^2 \\ &\leq \left((1 - b_1) + \sum_{j=1}^{n-1} (b_j - b_{j+1}) + b_n \right) \frac{c_u}{b_n} \tau^2 = \frac{c_u}{b_n} \tau^2, \end{aligned}$$

proving (4.5) for $k = n + 1$ that completes the proof of (4.5).

Consider $f(t) = t^{1-\alpha}$, then there exists a ξ , $k - 1 < \xi < k \leq M$ such that

$$\begin{aligned}
 b_{k-1}\tau^{-\alpha} &= \frac{(k\tau)^{1-\alpha} - (\tau(k-1))^{1-\alpha}}{\tau} = (1-\alpha)(\xi\tau)^{-\alpha} \\
 &\geq (1-\alpha)(k\tau)^{-\alpha} \geq (1-\alpha)(T)^{-\alpha},
 \end{aligned}$$

which gives

$$\frac{c_u}{b_{k-1}}\tau^2 \leq \frac{c_u}{1-\alpha}T^\alpha\tau^{2-\alpha}. \tag{4.7}$$

Now using this along with (4.5) proves (4.3).

To derive (4.4), we first prove

$$\|u(t_k) - u^k\|_1 \leq c_u k \tau^2, \quad k = 1, \dots, M. \tag{4.8}$$

By (4.6), the inequality (4.8) holds for $k = 1$. Assume that (4.8) holds for $k = 1, \dots, n$, $n \leq M - 1$. Then, from (1.1), (3.4) and (3.5), we obtain

$$\begin{aligned}
 \|e^{n+1}\|_1 &\leq (1 - b_1)\|e^n\| + \sum_{j=1}^{n-1} (b_j - b_{j+1})\|e^{n-j}\| + \|r^{n+1}\| \\
 &\leq \left((1 - b_1)\frac{n}{n+1} + \sum_{j=1}^{n-1} (b_j - b_{j+1})\frac{n-j}{n+1} + \frac{1}{(n+1)} \right) c_u(n+1)\tau^2 \\
 &\leq \left((1 - b_1)\frac{n}{n+1} + (b_1 - b_n)\frac{n}{n+1} - (b_1 - b_n)\frac{1}{n+1} + \frac{1}{(n+1)} \right) c_u(n+1)\tau^2 \\
 &= \left(1 - b_n\frac{n}{n+1} - (b_1 - b_n)\frac{1}{n+1} \right) c_u(n+1)\tau^2 \leq c_u(n+1)\tau^2.
 \end{aligned}$$

So (4.8) holds for $k = n + 1$. From $k\tau \leq T$ and (4.8), we get (4.4). □

4.1 Convergence of the full discretization scheme

Let $\pi_N^{1,0}$ be the H^1 -orthogonal projection operator from $H_0^1(\Omega)$ into $(\mathbb{P}_N^0)^2$ associated with the energy norm $\|\cdot\|_1$ defined in (4.1). Due to the equivalence of this norm with the standard H^1 norm, we have the following error estimation Lin and Chuanju (2007, Relation (4.3))

$$\left\| u - \pi_N^{1,0} u \right\|_1 \leq cN^{1-m}\|u\|_m, \quad u \in H_0^m(\Omega) \cap H_0^1(\Omega), \quad m \geq 1. \tag{4.9}$$

The idea of the proof for the following result comes from the paper (Lin and Chuanju 2007).

Theorem 3 *Let $u^k, k = 0, \dots, M$ be the solution of the variational formulation (3.7) and u_N^k be the solution of the scheme (3.8), assuming $u^0 = \pi_N^{1,0} u^0$ and $u^k \in H^m(\Omega) \cap H_0^1(\Omega)$ for some $m > 1$. Then,*

$$\begin{aligned}
 \left\| u^k - u_N^k \right\|_1 &\leq \frac{c}{1-\alpha} \tau^{-\alpha} N^{1-m} \max_{0 \leq j \leq k} \|u^j\|_m, \quad 0 < \alpha < 1, \\
 \left\| u^k - u_N^k \right\|_1 &\leq cN^{1-m} \sum_{j=0}^k \|u^j\|_m, \quad \alpha \rightarrow 1,
 \end{aligned} \tag{4.10}$$

for $k = 1, \dots, M$, where c depends only on T^α .

Proof We have $(u^{k+1} - \pi_N^{1,0} u^{k+1}, v_N)_1 = 0, \forall v_N \in (\mathbb{P}_N^0)^2$ by the projection operator. By definition of the norm (4.1), we get

$$\begin{aligned} & (\pi_N^{1,0} u^{k+1}, v_N) + \alpha_1 (\nabla \pi_N^{1,0} u^{k+1}, \nabla v_N) \\ &= (u^{k+1}, v_N) + \alpha_1 (\nabla u^{k+1}, \nabla v_N), \quad \forall v_N \in (\mathbb{P}_N^0)^2. \end{aligned}$$

By the weak form (3.7), the RHS of the above equation is replaced as

$$\begin{aligned} & (\pi_N^{1,0} u^{k+1}, v_N) + \alpha_1 (\nabla \pi_N^{1,0} u^{k+1}, \nabla v_N) = (1 - b_1) (u^k, v_N) \\ &+ \sum_{j=1}^{k-1} (b_j - b_{j+1}) (u^{k-j}, v_N) + b_k (u^0, v_N), \quad \forall v_N \in (\mathbb{P}_N^0)^2. \end{aligned} \tag{4.11}$$

Subtracting (4.11) from (3.8), we have

$$\begin{aligned} & (\tilde{e}_N^{k+1}, v_N) + \alpha_1 \left(\frac{\partial \tilde{e}_N^{k+1}}{\partial x}, \frac{\partial v_N}{\partial x} \right) \\ &= (1 - b_1) (e_N^k, v_N) + \sum_{j=1}^{k-1} (b_j - b_{j+1}) (e_N^{k-j}, v_N) + b_k (e_N^0, v_N), \quad \forall v_N \in (\mathbb{P}_N^0)^2, \end{aligned}$$

where $e_N^{k+1} = u^{k+1} - u_N^{k+1}$ and $\tilde{e}_N^{k+1} = \pi_N^{1,0} u^{k+1} - u_N^{k+1}$. Let $v_N = \tilde{e}_N^{k+1}$, then

$$\| \tilde{e}_N^{k+1} \|_1 \leq (1 - b_1) \| e_N^k \| + \sum_{j=1}^{k-1} (b_j - b_{j+1}) \| e_N^{k-j} \| + b_k \| e_N^0 \|.$$

With $\| e_N^{k+1} \|_1 \leq \| \tilde{e}_N^{k+1} \|_1 + \| u^{k+1} - \pi_N^{1,0} u^{k+1} \|_1$, we obtain

$$\| e_N^{k+1} \|_1 \leq (1 - b_1) \| e_N^k \| + \sum_{j=1}^{k-1} (b_j - b_{j+1}) \| e_N^{k-j} \| + b_k \| e_N^0 \| + cN^{1-m} \| u^{k+1} \|.$$

As in the proof of Theorem 2, it is first proved by induction that:

$$\begin{aligned} \| e_N^{k+1} \|_1 &\leq \frac{1}{b_{k-1}} \max_{0 \leq j \leq k} \| u^j - \pi_N^{1,0} u^j \|_1, \quad 0 < \alpha < 1, \\ \| e_N^{k+1} \|_1 &\leq \sum_{j=0}^k \| u^j - \pi_N^{1,0} u^j \|_1, \quad \alpha \rightarrow 1, \end{aligned}$$

for $0 \leq k \leq M$. Then, by using (4.7) and the projection error (4.9) the desired result is derived. □

The following theorem is obtained by the triangle inequality $\| u(\cdot, t_k) - u_N^k \|_1 \leq \| u(\cdot, t_k) - u^k \|_1 + \| u^k - u_N^k \|_1$ along with the inequalities (4.3) and (4.10).

Theorem 4 *Let u be the solution of the problem (1.1) with the initial and boundary conditions given by (1.2)–(1.3) and u_N^k be the solution of the scheme (3.8). Then, assuming $u_N^0 = \pi_N^{1,0} u^0$ and $u \in H^m(\Omega) \cap H_0^1(\Omega)$, we have*

$$\begin{aligned} \|u(t_k) - u_N^k\|_1 &\leq \frac{CT^\alpha}{1-\alpha} \left(c_u \tau^{2-\alpha} + c\tau^{-\alpha} N^{1-m} \sup_{0 < t < T} \|u(x, t)\|_m \right), \quad k \leq M, \quad 0 < \alpha < 1, \\ \|u(t_k) - u_N^k\|_1 &\leq T^\alpha \left(c_u \tau + c\tau^{-1} N^{1-m} \sup_{0 < t < T} \|u(x, t)\|_m \right), \quad k \leq M, \quad \alpha \rightarrow 1. \end{aligned} \tag{4.12}$$

The constants C and c are independent of τ, T, N .

It is seen that the method has the so-called spectral convergence in space and the order of convergence $O(\tau^{2-\alpha})$ in time.

5 Numerical examples

Here, some numerical experiments are provided to show the accuracy of the proposed method. For the computations, we use Maple 18 on a laptop with CPU core i3 1.9 GHz and RAM 4 running Windows 8.1 platform. To compute the errors, we use the discrete L^2 and L^∞ errors defined as

$$\begin{aligned} L^2 &\approx \left(\frac{1}{N^2} \sum_{i,j=0}^{N-1} |u(x_i, y_j, t_m) - u_N^m(x_i, y_j)|^2 \right)^{1/2}, \\ L^\infty &\approx \max_{0 \leq i,j \leq N} |u(x_i, y_j, t_m) - u_N^m(x_i, y_j)|, \end{aligned}$$

respectively, where u is the exact solution of the problem (1.1)–(1.3), u_N^m is the approximation solution (3.11) at $t = t_m = m\tau, x_i = \frac{i}{N}, y_j = \frac{j}{N}$ and $N = 100$. Also, the convergence rates in space and time are, respectively, computed by

$$\text{rate}_{N_i} = \frac{\log \frac{E(N_i, \tau)}{E(N_{i-1}, \tau)}}{\log \frac{N_{i-1}}{N_i}}, \quad \text{rate}_{\tau_i} = \frac{\log \frac{E(N, \tau_i)}{E(N, \tau_{i-1})}}{\log \frac{\tau_i}{\tau_{i-1}}},$$

where $E(h, \tau)$ is the error with $h = 1/N$, where N stands for the dimension of basis and τ is the time-step size. However, as it is common in the literature, we will show the spectral convergence of the proposed method by logarithmic scaled error plots.

It is worth mentioning that as we derived the operational matrices in Sect. 2.2 with special structures, and the proposed method finally leads to the linear system (3.13) that is sparse and banded. To see the sparsity and bandedness, the nonzero entries of the coefficient matrix of (3.13) are depicted in Fig. 2 using the *sparsematrixplot* command of Maple. It is seen that the density decreases rapidly as N increases.

Example 1 Consider the problem (1.1) with $\kappa = 1$ and the exact solution $u(x, y, t) = \sin(\pi x) \sin(\pi y)t^2$. Table 1 shows the convergence of the method in space for $\tau = \Delta t = 0.01$ for fractional orders $\alpha = 0.25, 0.50, 0.75$. N_x and N_y stand for the number of basis in the x and y direction. Figure 3 demonstrates the logarithmic scale error plot in terms of H^1 -norm for $\alpha = 1/2$ for some t s. It is seen that the method preserves the spectral convergence at different time rows $t < 1$ and $t > 1$. Table 2 reports the convergence in time by considering $N_x = N_y = N = 8$ as τ decreases at time rows $t = 0.1$ and $t = 1$. It verifies the $O(\tau^{2-\alpha})$ temporal rate of convergence.

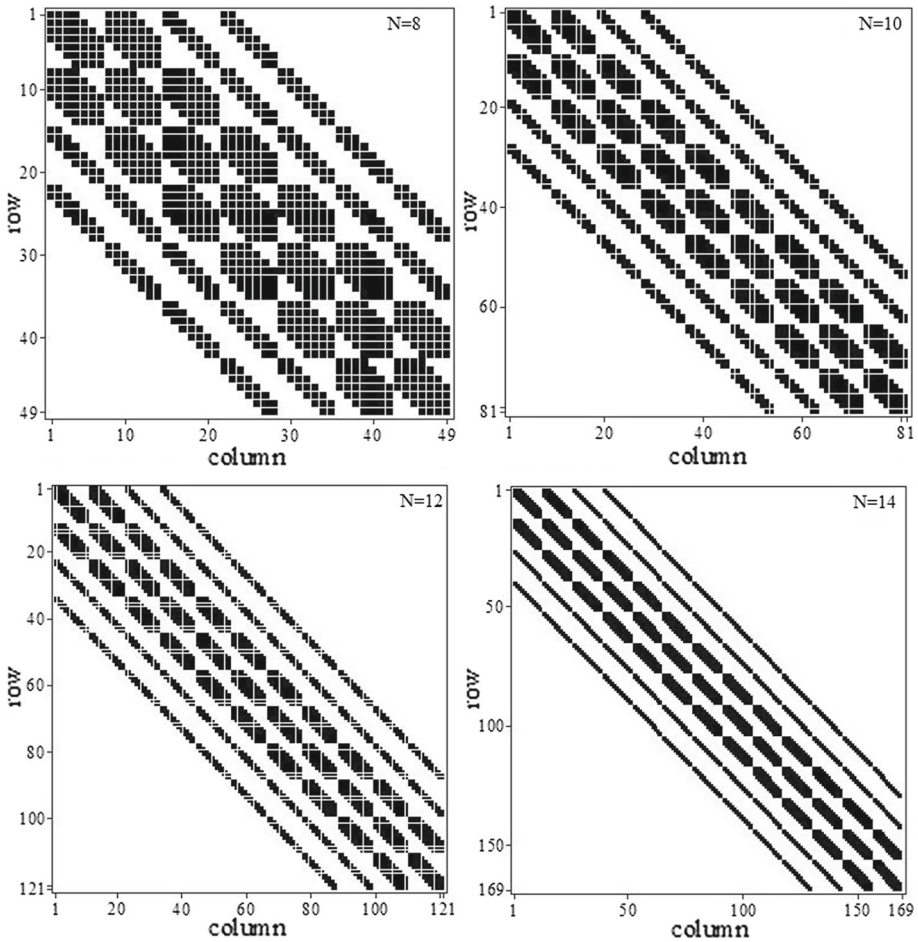


Fig. 2 The coefficient matrix of the linear system (3.13)

Table 1 Convergence in space at $t = 1$ for Example 1

$N_x = N_y$	$\alpha = 0.25$		$\alpha = 0.5$		$\alpha = 0.75$	
	L^∞	H^1	L^∞	H^1	L^∞	H^1
2	7.53E-02	2.81E-01	7.52E-02	2.81E-01	7.49E-02	2.81E-01
4	1.74E-03	8.91E-03	1.72E-03	8.91E-03	1.63E-03	8.91E-03
6	1.78E-05	1.34E-04	3.01E-05	1.43E-04	9.98E-05	2.86E-04
8	3.67E-06	8.75E-06	2.25E-05	5.15E-05	2.79E-06	2.54E-04

Example 2 To see whether the method works for the case in which there is no source term, consider the problem (1.1)–(1.3) with the initial condition $u(x, y, 0) = x(x - 1) \sin(2\pi y)$, $\kappa = 1$ and no source term (Yang and Zhang 2014).

Fig. 3 Convergence of the spectral method (3.13) in space with $\tau = 0.001$

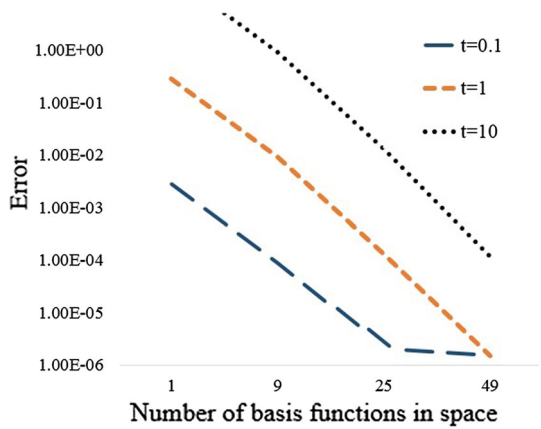
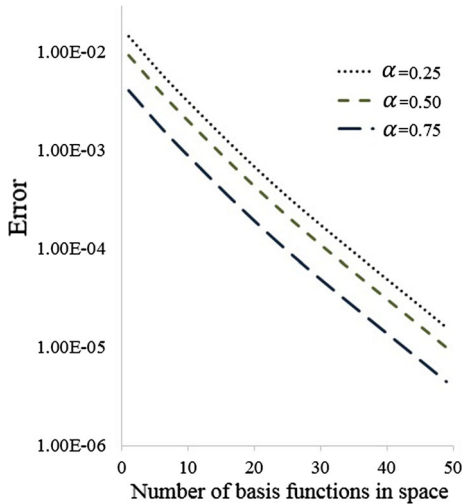


Table 2 Error and temporal rate of convergence at $t = 1$ for Example 1

M	$\alpha = 0.25$				$\alpha = 0.5$				$\alpha = 0.75$			
	t = 0.1		t = 1		t = 0.1		t = 1		t = 0.1		t = 1	
	H^1	Rate	H^1	Rate	H^1	Rate	H^1	Rate	H^1	Rate	H^1	Rate
10	2.90E-04		4.21E-04		1.16E-03		1.55E-03		3.27E-03		4.46E-03	
20	9.93E-05	1.55	1.32E-04	1.67	4.60E-04	1.33	5.61E-04	1.47	1.54E-03	1.09	1.89E-03	1.24
40	3.27E-05	1.60	4.12E-05	1.68	1.73E-04	1.41	2.01E-04	1.48	6.87E-04	1.16	7.95E-04	1.25
80	1.05E-05	1.64	1.27E-05	1.70	6.35E-05	1.45	7.18E-05	1.49	2.97E-04	1.21	3.35E-04	1.25
160	3.31E-06	1.67	4.05E-06	1.65	2.30E-05	1.47	2.56E-05	1.49	1.26E-04	1.24	1.41E-04	1.25

Fig. 4 Convergence in space for some fractional orders with $\tau = 0.01$



The spectral convergence in space is seen from Fig. 4 in which the time step length is considered to be $\tau = 0.01$. The solution with $N_x = N_y = 10$ is treated as the exact solution. The errors are reported at $t = 1$ with the H^1 -norm.

The numerical examples present the spectral convergence in space and fixed convergence of $O(\tau^{2-\alpha})$ in time, confirming the theoretical claims. It should be noted that we have used the eight-point Gauss–Legendre quadrature rule to perform the integrals (3.12) in the right hand side of the linear system (3.13).

6 Conclusion

In this paper, some new aspects of the dual Bernstein polynomials have been discussed. A suitable compact combinations of these polynomials has been derived for developing a dual-Petrov–Galerkin variational formulation for the numerical simulation of the two-dimensional subdiffusion equation. It is shown that the method leads to sparse linear systems. The illustrated numerical examples have been provided to show the accuracy of the method. It is important to note that the transformation matrices and the operational matrix for differentiation of dual Bernstein polynomials that have been obtained in this work can be used similarly for developing Bernstein-based dual-Petrov–Galerkin Galerkin methods for other fractional partial differential equations on bounded domains.

References

- Behiry SH (2014) Solution of nonlinear Fredholm integro-differential equations using a hybrid of block pulse functions and normalized Bernstein polynomials. *J Comput Appl Math* 260:258–265
- Carnicer JM, Pena JM (1993) Shape preserving representations and optimality of the Bernstein basis. *Adv Comput Math* 1:173–196
- Dabiri A, Butcher EA (2017) Efficient modified Chebyshev differentiation matrices for fractional differential equations. *Commun Nonlinear Sci Numer Simul*. doi:[10.1016/j.cnsns.2017.02.009](https://doi.org/10.1016/j.cnsns.2017.02.009)
- Deng W (2008) Finite element method for the space and time fractional Fokker–Planck equation. *SIAM J Numer Anal* 47:204–226
- Farin GE, Hoschek J, Kim MS (2002) *Handbook of computer aided geometric design*. Elsevier, Amsterdam
- Farouki RT, Rajan VT (1988) Algorithms for polynomials in Bernstein form. *Comput Aided Geom Des* 5:1–26
- Farouki RT (1991) On the stability of transformations between power and Bernstein polynomial forms. *Comput Aided Geom Des* 8:29–36
- Farouki RT, Goodman TNT (1996) On the optimal stability of the Bernstein basis. *Math Comput* 64:1553–1566
- Gao GH, Sun ZZ, Zhang YN (2012) A finite difference scheme for fractional sub-diffusion equations on an unbounded domain using artificial boundary conditions. *J Comput Phys* 231:2865–2879
- Gao GH, Sun HW, Sun ZZ (2015) Stability and convergence of finite difference schemes for a class of time-fractional sub-diffusion equations based on certain superconvergence. *J Comput Phys* 280:510–528
- Golub GH, Ortega JM (1992) *Scientific computing and differential equations: an introduction to numerical methods*. Academic Press, San Diego
- Goubet O, Shen J (2007) On the dual Petrov–Galerkin formulation of the KDV equation on a finite interval. *Adv Differ Equ* 12:221–239
- Goychuk I (2009) Viscoelastic subdiffusion: from anomalous to normal. *Phys Rev E* 80:046125
- Jani M, Babolian E, Javadi S, Bhatta D (2017) Banded operational matrices for Bernstein polynomials and application to the fractional advection-dispersion equation. *Numer Algorithms*. doi:[10.1007/s11075-016-0229-1](https://doi.org/10.1007/s11075-016-0229-1)
- Javadi S, Babolian E, Taheri Z (2016) Solving generalized pantograph equations by shifted orthonormal Bernstein polynomials. *J Comput Appl Math* 303:1–14
- Javadi S, Jani M, Babolian E (2016) A numerical scheme for space-time fractional advection-dispersion equation. *Int J Nonlinear Anal Appl* 7:331–343
- Jin B, Lazarov R, Zhou Z (2016) Two fully discrete schemes for fractional diffusion and diffusion-wave equations with nonsmooth data. *SIAM J Sci Comput* 38:A146–A170
- Juttler B (1998) The dual basis functions for the Bernstein polynomials. *Adv Comput Math* 8:345–352
- Lewanowicz S, Wozny P (2011) Bezier representation of the constrained dual Bernstein polynomials. *Appl Math Comput* 218:4580–4586

- Lin Y, Chuanju X (2007) Finite difference/spectral approximations for the time-fractional diffusion equation. *J Comput Phys* 225:1533–1552
- Maleknejad K, Basirat B, Hashemizadeh E (2012) A Bernstein operational matrix approach for solving a system of high order linear Volterra-Fredholm integro-differential equations. *Math Comput Modell* 55:1363–1372
- Metzler R, Klafter J (2004) The restaurant at the end of the random walk: recent developments in the description of anomalous transport by fractional dynamics. *J Phys A Math Gen* 37:R161–R208
- Moghaddam BP, Machado JAT (2017) A stable three-level explicit spline finite difference scheme for a class of nonlinear time variable order fractional partial differential equations. *Comput Math Appl*. doi:[10.1016/j.camwa.2016.07.010](https://doi.org/10.1016/j.camwa.2016.07.010)
- Mustapha K, Nour M, Cockburn B (2016) Convergence and superconvergence analyses of HDG methods for time fractional diffusion problems. *Adv Comput Math* 42:377–393
- Pang HK, Sun HW (2016) Fourth order finite difference schemes for time—space fractional sub-diffusion equations. *Comput Math Appl* 71:287–1302
- Ramezani M, Mojtabaei M, Mirzaei D (2015) DMLPG solution of the fractional advection—diffusion problem. *Eng Anal Bound Elem* 59:36–42
- Ren J, Sun ZZ, Zhao X (2013) Compact difference scheme for the fractional sub-diffusion equation with Neumann boundary conditions. *J Comput Phys* 232:456–467
- Saadatmandi A (2014) Bernstein operational matrix of fractional derivatives and its applications. *Appl Math Model* 38:1365–1372
- Shen J, Tang T, Wang LL (2011) *Spectral methods: algorithms, analysis and applications*. Springer, Berlin
- Shirzadi A, Ling L, Abbasbandy S (2012) Meshless simulations of the two-dimensional fractional-time convection—diffusion—reaction equations. *Eng Anal Bound Elem* 36:1522–1527
- Stokes PW, Philippa B, Read W, White RD (2015) Efficient numerical solution of the time fractional diffusion equation by mapping from its Brownian counterpart. *J Comput Phys* 282:334–344
- Wang T, Wang YM (2016) A compact LOD method and its extrapolation for two-dimensional modified anomalous fractional sub-diffusion equations. *Comput Math Appl* 71:147–170
- Wozny P, Lewanowicz S (2009) Multi-degree reduction of Bezier curves with constraints, using dual Bernstein basis polynomials. *Comput Aided Geom Des* 26:566–579
- Yang X-J, Tenreiro Machado JA, Srivastava HM (2016) A new numerical technique for solving the local fractional diffusion equation: two-dimensional extended differential transform approach. *Appl Math Comput* 274:143–151
- Yang JY, Zhao YM, Liu N, Bu WP, Xu TL, Tang YF (2015) An implicit MLS Meshless method for 2D time dependent fractional diffusion-wave equation. *Appl Math Model* 39:1229–1240
- Yang X, Zhang H, Da Xu (2014) Orthogonal spline collocation method for the two-dimensional fractional sub-diffusion equation. *J Comput Phys* 256:824–837
- Yuan JM, Shen J, Wu J (2008) A dual-Petrov-Galerkin method for the Kawahara-type equations. *J Sci Comput* 34:48–63
- Zhang YN, Sun ZZ (2011) Alternating direction implicit schemes for the two-dimensional fractional sub-diffusion equation. *J Comput Phys* 230:8713–8728
- Zhang YN, Sun ZZ, Zhao X (2012) Compact alternating direction implicit scheme for the two-dimensional fractional diffusion-wave equation. *SIAM J Numer Anal* 50:1535–1555
- Zhao Y, Zhang Y, Shi D, Liu F, Turner I (2016) Superconvergence analysis of nonconforming finite element method for two-dimensional time fractional diffusion equations. *Appl Math Lett* 59:38–47
- Zhou Z, Gong W (2016) Finite element approximation of optimal control problems governed by time fractional diffusion equation. *Comput Math Appl* 71:301–318