# Finite element methods for fractional-order diffusion problems with optimal convergence order

Gábor Maros [b], Ferenc Izsák [a],*

[a] *Department of Applied Analysis and Computational Mathematics & MTA ELTE NumNet Research Group, Eötvös Loránd University, Pázmány P. stny. 1C, 1117 Budapest, Hungary*
[b] *Department of Applied Analysis and Computational Mathematics, Eötvös Loránd University, Pázmány P. stny. 1C, 1117 Budapest, Hungary*

## ARTICLE INFO

## ABSTRACT

A convergence result is stated for the numerical solution of space-fractional diffusion problems. For the spatial discretization, an arbitrary family of finite elements can be used combined with the matrix transformation technique. The analysis covers the application of the implicit Euler method for time integration to ensure unconditional stability. The spatial convergence rate does not depend on the fractional power of the Laplacian operator. An efficient numerical implementation is developed avoiding the direct computation of matrix powers.

© 2020 Elsevier Ltd. All rights reserved.

## 1. Introduction

Many observations confirmed the presence of the fractional diffusion in the natural sciences. Tracking the motion of food-seeking animals, the presence of "anomalous" diffusion was reported earlier [1,2]. In concrete terms, assuming diffusion, the flight $x(t)$ of the individuals during time $t$ obeys a normal distribution. Consequently, the flight-length mean $\langle |x(t)| \rangle$ of the individuals should be proportional with $\sqrt{t}$.

Due to accurate measurements in the last decade, the dynamics of individual molecules could also be observed: in many cases, it exhibits a subdiffusive behavior such that instead of the above proportionality, the linear relation $\langle |x(t)| \rangle \sim \sqrt{t}^{\alpha}$ was detected. See a detailed overview of these measurements in [3]. A possible explanation of this dynamics can be found in [4].

A number of different mathematical models have been suggested to simulate this dynamics. Among the PDE models, in case of homogeneous Dirichlet boundary conditions, a possible choice is taking the fractional Laplacian on the entire space $\mathbb{R}^d$ (see [5]) and restricting to functions which are identically zero outside of the computational domain. The numerical analysis and implementation for the corresponding elliptic problems can be found in [6] and [7], respectively. In the one-dimensional case, an interesting, physically motivated approach is analyzed in [8].

Another conventional choice, which is used in the present work, is the spatial differential operator $-(-\Delta_{\mathcal{D}})^{\alpha}$, where $\Delta_{\mathcal{D}}$ denotes the Dirichlet Laplacian.

For a systematic comparison of the different approaches, we refer to [9,10] and [11].

For the numerical solution of problems with $-(-\Delta_{\mathcal{D}})^{\alpha}$, the fractional power can be applied also at the discrete level, approximating $(-\Delta_{\mathcal{D}})^{\alpha}$ with the power of the discretization of $-\Delta_{\mathcal{D}}$. This was first observed in [12] and analyzed in [13]

---

* Corresponding author.
  *E-mail addresses:* magabor@caesar.elte.hu (G. Maros), izsakf@caesar.elte.hu (F. Izsák).

for the corresponding elliptic problems. A similar analysis – based on a general integral form of operator powers – was performed in [14] using an efficient numerical integration technique.

Regarding the full discretization with an implicit time stepping and efficient implementation issues, three main approaches were proposed in the literature. In [15], the authors propose a direct approximation of the solution operator (which is a matrix-power exponential) based on the elliptic theory in [14]. The corresponding contour integral is recasted as a real improper integral and approximated using an exponentially convergent quadrature.

In the works [16] and [17], explicit and implicit time discretizations are used. In [17], again using an integral representation, an efficient simple quadrature is proposed to approximate time steps in an implicit method, which is shown to be unconditionally stable.

A related approach was recently developed in [18], where the best rational approximation (BURA) of the implicit time-step operators – including matrix powers – is constructed.

Dealing with the full matrices arising from the discretization of non-local operators is a challenging topic. One can also explore their structure, which makes possible to develop an efficient solution of the corresponding linear problems, see, *e.g.*, [19] and [20].

The aim of the present contribution is to propose an alternative approach, which has the following advances.

- The analysis for the full discretization error is simple.
- The spatial convergence order is independent of the power $\alpha$ and can ensure a higher-order accuracy.
- The numerical method is simple and applicable without any extra computations for all exponents $\alpha$.

To compare which the earlier achievements, note that in the literature, a spatial error of order $h^\alpha$ appears in the error estimates for any initial conditions, see, *e.g.*, Theorem 3.1 in [15]. We point out that assuming smooth initial condition, this can be changed to $h^k$, where $k$ is the polynomial order in the finite element approximation. The smoothness condition will be discussed after Theorem 2. Also, in the computationally most efficient BURA approach (see [18]), the determination of the coefficients for an arbitrary rational function $r^\alpha$ needs efforts.

The main ideas of our works are the following:

- For the analysis of the spatial discretization error we have used a weak form, which, for the FE discretization is sufficient.
- The computation algorithm is based on the matrix power–vector product in [21] combined with a conjugated gradient method such that the entire algorithm is composed of sparse matrix–vectors products.

## 2. Mathematical preliminaries

We investigate the numerical solution of space-fractional diffusion problems. Recall that the (negative) Dirichlet Laplacian operator $-\Delta_{\mathcal{D}} : L_2(\Omega) \to L_2(\Omega)$ is positive and has a compact inverse. The complete orthonormal system of its eigenfunctions and the corresponding eigenvalues will be denoted by $\{\phi_j\}_{j\in\mathbb{N}}$ and $\{\lambda_j\}_{j\in\mathbb{N}}$, respectively. Then the fractional Dirichlet Laplacian, which is investigated here, is defined with

$$\text{Dom}\,(-\Delta_{\mathcal{D}})^\alpha = \{u = \sum_{j=0}^\infty u_j\phi_j : \sum_{j=0}^\infty u_j^2 \cdot \lambda_j^{2\alpha} < \infty\}, \quad (-\Delta_{\mathcal{D}})^\alpha u = \sum_{j=0}^\infty u_j\lambda_j^\alpha\phi_j.$$

For more details, we refer to [22].

With this, the space-fractional diffusion equation reads as

$$\begin{cases} \partial_t u(t,x) = -(-\Delta_{\mathcal{D}})^\alpha u(t,x) & x \in \Omega,\ t \in \mathbb{R}^+ \\ u(0,x) = u_0(x) & x \in \Omega, \end{cases} \tag{1}$$

where $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) is a Lipschitz domain. Note that the definition of the Dirichlet Laplacian involves the homogeneous Dirichlet boundary condition.

We use the notion of Sobolev spaces $H^k(\Omega)$ and $H_0^k(\Omega)$ with a non-negative index $k$; the corresponding inner product is denoted with $(\cdot;|\cdot)_k$ and the notation $\|\cdot\|_k$ will be used for the corresponding Sobolev norm. In case of $k = 0$, we usually omit the subscript. To depict clearly the matrix–vector operations in the practical computations, the Euclidian scalar product in $\mathbb{R}^N$ will be denoted with $(\cdot,\cdot)$.

In the estimates, the relation $r_1 \lesssim r_2$ means that $r_1 \leq cr_2$ is valid with a positive mesh-independent constant $c$.

The spatial discretization is performed using a generic finite element space $V_h^k \subset H_0^k(\Omega)$ and the corresponding elliptic projection $P_h : H_0^k(\Omega) \to V_h^k$, which satisfies

$$(\Delta_{\mathcal{D}} P_h v | v_h) = (\Delta_{\mathcal{D}} v | v_h) \quad \forall v_h \in V_h^k. \tag{2}$$

We assume that $V_h^k$ is chosen so that

$$\|v - P_h v\|_0 \lesssim h^k \|v\|_k \quad \forall v \in H_0^k(\Omega). \tag{3}$$

For the corresponding requirements, we refer to [23, Corollary 1.109]. Using a finite element basis $\{\varphi_j\}_{j=1,\ldots,N}$ of $V_h^k$, one can define the finite element mass matrix $M_h \in \mathbb{R}^{N \times N}$ and the stiffness matrix $S_h \in \mathbb{R}^{N \times N}$, the entries of which are given by

$$[M_h]_{i,j} = (\varphi_j|\varphi_i) \quad \text{and} \quad [S_h]_{i,j} = (\varphi_j|\varphi_i)_1 = (\nabla\varphi_j|\nabla\varphi_i). \tag{4}$$

In the discrete setting, we use the following expansion of $u_h \in V_h$:

$$u_h = u_1\varphi_1 + u_2\varphi_2 + \cdots + u_N\varphi_N \tag{5}$$

which defines a natural linear bijection

$$\pi_h : V_h \to \mathbb{R}^N, \quad \text{with} \quad \pi_h(u_h) = \mathbf{u} = (u_1, u_2, \ldots, u_N)^T. \tag{6}$$

For the approximation of the differential operator, we apply the so-called matrix transformation technique. According to this approach, the operation $(-\Delta_{\mathcal{D}})^\alpha v$ is approximated with $D_h^\alpha \pi_h(v_h)$, where $D_h \in \mathbb{R}^{N \times N}$ denotes a matrix corresponding to $-\Delta_{\mathcal{D}}$ in $V_h^k$. In case of finite difference discretization, this is straightforward [24], but for finite element methods combined with implicit time-stepping the definition of $D_h$ needs a special care.

## 3. Results

Using the variational principle together with the backward Euler time stepping for $\alpha = 1$, the full discretization of (1) can be given as

$$\left(\frac{u_h^{n+1} - u_h^n}{\delta}|\varphi_j\right) = -\left(\Delta_{\mathcal{D}}u_h^{n+1}|\varphi_j\right) \qquad j = 1, 2, \ldots, N,$$

where $\delta > 0$ is the time step and

$$u_h^n = u_1^n\varphi_1 + u_2^n\varphi_2 + \cdots + u_N^n\varphi_N \in H_0^1(\Omega)$$

is the approximation of $u(n\delta, \cdot) : \Omega \to \mathbb{R}$. According to (4), this can be recasted into the matrix–vector form

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\delta} = -D_h\mathbf{u}^{n+1},$$

where $D_h = M_h^{-1}S_h$. An important observation here, that for using the matrix transformation method, we need to take the power of $D_h$ instead of the stiffness matrix $S_h$. With this, the matrix transformation method for a general $\alpha \in \mathbb{R}^+$, combined with the time discretization above, can be given as

$$\frac{\mathbf{u}^{n+1} - \mathbf{u}^n}{\delta} = -D_h^\alpha\mathbf{u}^{n+1}. \tag{7}$$

We perform an error analysis for this approach.

### 3.1. Spatial discretization

For the error analysis of the spatial discretization, we need the following identities.

**Lemma 1.** *For any $\mathbf{a} \in \mathbb{R}^N$ and $v_h \in V_h$ we have*

$$(\pi_h^{-1}\mathbf{a}|v_h) = (M_h\mathbf{a}, \pi_h v_h) \tag{8}$$

*and*

$$(\nabla\pi_h^{-1}\mathbf{a}|\nabla v_h) = (S_h\mathbf{a}, \pi_h v_h). \tag{9}$$

**Proof.**
Using the definition of $\pi_h$, and the expansion in (5), we have

$$(\pi_h^{-1}\mathbf{a}|v_h) = (a_1\varphi_1 + \cdots + a_N\varphi_N|v_1\varphi_1 + \cdots + v_N\varphi_N) = M_h\mathbf{a} \cdot (v_1, v_2, \ldots, v_N)^T$$
$$= (M_h\mathbf{a}, \pi_h v_h),$$

as stated. The derivation of the second equality can be performed in a similar way. $\quad\square$

For the brevity, we also use the notation

$$(-\Delta_{\mathcal{D},h})^\alpha = \pi_h^{-1}D_h^\alpha\pi_h, \tag{10}$$

which gives an approximation of $(-\Delta_{\mathcal{D}})^\alpha$ on $V_h^k$.

To obtain a sharp estimate of this term, we use a weak formulation of Balakrishnan's representation [25] in the case of Hilbert space operators. Henceforth, in the article, we assume that $\alpha \in (0, 1)$.

**Theorem 1.** *Let $A$ denote a positive self-adjoint operator on a Hilbert space $\mathcal{H}$ and $\alpha \in (0,1)$ an arbitrary exponent. Then for all $u \in \operatorname{Dom} A$ and $v \in \mathcal{H}$ we have the following equality:*

$$\left(A^\alpha u | v\right) = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \left(s^{\alpha-1} A(sI+A)^{-1} u | v\right) \, ds. \tag{11}$$

Henceforth, we analyze the case $A = -\Delta_{\mathcal{D}}$ and $A_h = -\Delta_{\mathcal{D},h}$. The basis of the spatial error estimation is the following statement.

**Proposition 1.** *Using the above notations, we have the following inequality for each eigenfunction $\phi_j$ of $-\Delta_{\mathcal{D}}$:*

$$\left| \left( ((-\Delta_{\mathcal{D}})^\alpha - (-\Delta_{\mathcal{D},h})^\alpha P_h)\phi_j | v_h \right) \right| \lesssim h^k \lambda_j^\alpha \|\phi_j\|_k \|v\|.$$

**Proof.** Using Balakrishnan's representation in (11), we have

$$\left((A^\alpha - A_h^\alpha P_h)\phi_j | v_h\right) = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty s^{\alpha-1}\left((A(sI+A)^{-1} - A_h(sI+A_h)^{-1} P_h)\phi_j | v_h\right) \, ds. \tag{12}$$

Using the definition of $P_h$ in (2), then (9), (8) and (10), we obtain that for an arbitrary $w \in H_0^k(\Omega)$ the following identity is valid:

$$(Aw|v_h) = (AP_h w|v_h) = (\nabla P_h w|\nabla v_h) = (S_h \pi_h P_h w|\pi_h v_h)$$
$$= (M_h D_h \pi_h P_h w|\pi_h v_h) = \left(\pi_h^{-1} D_h \pi_h P_h w|v_h\right) = (A_h P_h w|v_h).$$

Using this with $w = (sI+A)^{-1}\phi_j$, we can rewrite the scalar product on the right hand side of (12) as follows:

$$\left((A(sI+A)^{-1} - A_h(sI+A_h)^{-1} P_h)\phi_j | v_h\right) = \left(A_h(P_h(sI+A)^{-1} - (sI+A_h)^{-1} P_h)\phi_j | v_h\right)$$
$$= \left((P_h(sI+A)^{-1} - (sI+A_h)^{-1} P_h)\phi_j | A_h v_h\right). \tag{13}$$

Inserting the identity

$$P_h(sI+A)^{-1} - (sI+A_h)^{-1} P_h = (sI+A_h)^{-1}[(sI+A_h)P_h - P_h(sI+A)](sI+A)^{-1}$$

into the last term of (13), using the self-adjoint property of $(sI+A_h)^{-1}$ and (12) again, we obtain that

$$\left((A(sI+A)^{-1} - A_h(sI+A_h)^{-1} P_h)\phi_j | v_h\right) = \left((sI+A_h)^{-1}[(sI+A_h)P_h - P_h(sI+A)](sI+A)^{-1}\phi_j | A_h v_h\right)$$
$$= \left((sI+A_h)^{-1}[A_h P_h - P_h A](sI+A)^{-1}\phi_j | A_h v_h\right) = \left([A_h P_h - P_h A](sI+A)^{-1}\phi_j | (sI+A_h)^{-1} A_h v_h\right) \tag{14}$$
$$= \left([I - P_h]A(sI+A)^{-1}\phi_j | (sI+A_h)^{-1} A_h v_h\right).$$

Applying the approximation property in (3) with (13) and (14) with the Cauchy–Schwarz inequality we finally have that

$$\left| \left((A(sI+A)^{-1} - A_h(sI+A_h)^{-1} P_h)\phi_j | v_h\right) \right| = \left| \left([I - P_h]A(sI+A)^{-1}\phi_j | (sI+A_h)^{-1} A_h v_h\right) \right|$$
$$\leq \|[I - P_h]A(sI+A)^{-1}\phi_j\| \|(sI+A_h)^{-1} A_h v_h\| \leq h^k \|A(sI+A)^{-1}\phi_j\|_k \|(sI+A_h)^{-1} A_h v_h\| \tag{15}$$
$$= h^k \|\lambda_j(sI+\lambda_j)^{-1}\phi_j\|_k \|(sI+A_h)^{-1} A_h v_h\|.$$

Since $A_h$ is a positive operator on $V_h$, we also have

$$\|(s+A_h)^{-1} A_h v_h\| \leq \|v_h\|$$

for every $s \geq 0$. Inserting this estimate with (15) into the right-hand side of (12), we get

$$\left|\left((A^\alpha - A_h^\alpha P_h)\phi_j | v_h\right)\right| \leq \frac{\sin(\pi\alpha)}{\pi} ch^k \int_0^\infty s^{\alpha-1}\lambda_j(s+\lambda_j)\|\phi_j\|_k \|v_h\| \, ds \lesssim h^k \lambda_j^\alpha \|\phi_j\|_k \|v_h\|,$$

as stated in the proposition. $\square$

Using the orthonormal system $\left\{\phi_j\right\}_{j\in\mathbb{N}}$, we have the following expansions in (1):

$$u(0, \cdot) = \sum_{j=1}^\infty u_{0j}\phi_j, \quad u(t, \cdot) = \sum_{j=1}^\infty u_j\phi_j, \tag{16}$$

where $\left\{u_{0j}\right\}_{j\in\mathbb{N}}, \left\{u_j\right\}_{j\in\mathbb{N}} \subset \mathbb{R}$, and in the second case, the $t$-dependence of the coefficients $u_j$ is not displayed.

**Theorem 2.** *Assume that $\sum_{j=0}^\infty |\lambda_j^{k/2+\alpha} u_{0j}| = C_0 < \infty$. Then the following estimate holds for a general $u = u(t, \cdot)$ in (1):*

$$\left|\left((A^\alpha - A_h^\alpha P_h)u | v_h\right)\right| \lesssim C_0 h^k \|v_h\|.$$

**Proof.**

The solution of (1) can be given as

$$u(t, \cdot) = \sum_{j=0}^{\infty} e^{-\lambda_j^\alpha \cdot t} u_{0,j} \phi_j,$$

see [26], such that $|u_j| < |u_{0j}|$ and the assumption imply that

$$\sum_{j=0}^{\infty} |\lambda_j^{k/2+\alpha} u_j| < C_0 \tag{17}$$

uniformly for all $t \in \mathbb{R}^+$.

We also recall a growth condition for the $H^k(\Omega)$-norm of the Dirichlet-Laplacian eigenfunctions:

$$\|\phi_j\|_k \lesssim \lambda_j^{\frac{k}{2}}. \tag{18}$$

Note also that by the assumption, $u_j \lambda_j \to 0$. Hence, there is an index $j_0$ such that for all $j \geq j_0$, the following estimate is valid:

$$\left(u_j \lambda_j^\alpha\right)^2 \leq |u_j \lambda_j^\alpha| \leq \lambda_j^{\frac{k}{2}} |u_j \lambda_j^\alpha|.$$

Accordingly, we have $\sum_{j=0}^{\infty} \left(u_j \lambda_j^\alpha\right)^2 < \infty$, and therefore, $u \in \mathcal{D}(A^\alpha)$ with

$$A^\alpha u = \sum_{j=0}^{\infty} u_j \lambda_j^\alpha \phi_j = \sum_{j=0}^{\infty} A^\alpha \left(u_j \phi_j\right). \tag{19}$$

In other words, (19) means for $\alpha = 1$ that

$$\lim_{K \to \infty} A \sum_{j=0}^{K} u_j \phi_j = \lim_{K \to \infty} \sum_{j=0}^{K} u_j A \phi_j = \sum_{j=0}^{\infty} u_j A \phi_j = A \sum_{j=0}^{\infty} u_j \phi_j = Au.$$

In this way, the limit

$$\lim_{K \to \infty} \sum_{j=0}^{K} u_j \phi_j = u = \sum_{j=0}^{\infty} u_j \phi_j$$

is also valid in the $H^1(\Omega)$-norm so that using the $H^1(\Omega)$-orthogonality of the projection $P_h$ and (16), we finally have

$$P_h \sum_{j=0}^{K} u_j \phi_j = P_h u = \sum_{j=0}^{\infty} P_h \left(u_j \phi_j\right) = \lim_{K \to \infty} \sum_{j=0}^{K} P_h u_j \phi_j \tag{20}$$

Applying the continuous linear operator $A_h^\alpha : V_h \to H_0^1(\Omega)$ in (20), using (19), the estimate in Proposition 1, the estimate in (18) and finally the assumption in (17), we have

$$\left| \left( (A^\alpha - A_h^\alpha P_h) u | v_h \right) \right| = \left| \left( (A^\alpha - A_h^\alpha P_h) \sum_{j=0}^{\infty} u_j \phi_j | v_h \right) \right|$$

$$= \left| \left( A^\alpha \sum_{j=0}^{\infty} u_j \phi_j - A_h^\alpha P_h \sum_{j=0}^{\infty} u_j \phi_j | v_h \right) \right| = \left| \left( \sum_{j=0}^{\infty} u_j A^\alpha \phi_j - \sum_{j=0}^{\infty} u_j A_h^\alpha P_h \phi_j | v_h \right) \right|$$

$$= \left| \left( \sum_{j=0}^{\infty} u_j (A^\alpha - A_h^\alpha P_h) \phi_j | v_h \right) \right| \leq \sum_{j=0}^{\infty} \left| \left( u_j (A^\alpha - A_h^\alpha P_h) \phi_j | v_h \right) \right|$$

$$\lesssim \|v_h\| \sum_{j=0}^{\infty} |u_j| h^k \lambda_j^\alpha \|\phi_j\|_k \lesssim h^k \|v_h\| \sum_{j=0}^{\infty} |u_j| \lambda_j^{\alpha + \frac{k}{2}} \lesssim h^k \|v_h\|,$$

which proves the statement. $\square$

**Remark.** If the fractional-order diffusion is observed from $t_0 > 0$, i.e. $u(0, \cdot) = u^*(t_0, \cdot)$ for some $u^*$, the exponential decay of the coefficients $u_j$ implies that $\sum_{j=0}^{\infty}(j+1)^K u_{0j} < \infty$ for any power $K$. Therefore, the relation $\lambda_j \approx j^{\frac{d}{2}}$ immediately gives $\sum_{j=0}^{\infty} \lambda_j^K u_{0j} < \infty$, as we have assumed in Theorem 2.

### 3.2. Time discretization

Now we can prove the convergence of the full discretization. Since for the solution $u$ of the problem in (1), we have $u(t, \cdot) \in C^{\infty}(\Omega)$ for every $t \in (0, T]$ (see Proposition 1 in [9]), we only need the assumption in Theorem 2 for the initial data. Note that regarding the stability, a related result was established in [16] including also second order time discretizations with a possible source term. Here, we also give the convergence rate explicitly.

**Theorem 3.** *Using the assumption in Theorem 2 for $u(0, \cdot)$ in (1), the full discretization in (7) is convergent of order $O(\delta + h^k)$, where $k$ is the approximation order of the finite element discretization.*

**Proof.** Rewriting (7) in $V_h^k$ gives the scheme

$$1/\delta\big(u_h^{j+1} - u_h^j | v_h\big) + \big(A_h^{\alpha} u_h^{j+1} | v_h\big) = 0 \qquad \forall v_h \in V_h^k, \quad j = 0, 1, \dots, \tag{21}$$

where $u_h^j$ is the numerical solution at $t = j\delta$. $\quad\square$

To analyze this scheme, we use the notation $u^j = u(j\delta, \cdot)$ with $u$ the analytic solution of (1) and the elliptic projection $P_h$ introduced in (2) to obtain the following equality:

$$\begin{aligned}
&1/\delta\big(P_h u^{j+1} - P_h u^j | v_h\big) + \big(A_h^{\alpha} P_h u^{j+1} | v_h\big) \\
&= 1/\delta\big(u^{j+1} - u^j | v_h\big) + \big(A^{\alpha} u^{j+1} | v_h\big) + 1/\delta\big((P_h u^{j+1} - u^{j+1}) - (P_h u^j - u^j) | v_h\big) \\
&\quad + \big(A_h^{\alpha} P_h u^{j+1} - A^{\alpha} u^{j+1} | v_h\big) \\
&= \big(1/\delta(u^{j+1} - u^j) - \partial_t u^{j+1} | v_h\big) + 1/\delta\big((P_h u^{j+1} - u^{j+1}) - (P_h u^j - u^j) | v_h\big) \\
&\quad + \big(A_h^{\alpha} P_h u^{j+1} - A^{\alpha} u^{j+1} | v_h\big) \\
&:= (z^j | v_h),
\end{aligned} \tag{22}$$

where $z^j$ can be recognized as a consistency error, which we estimate termwise.

Obviously, using the general mean value theorem, and the smoothness of the analytic solution, we obtain

$$\begin{aligned}
\|1/\delta(u^{j+1} - u^j) - \partial_t u^{j+1}\| &= \left\| 1/\delta \int_0^{\delta} \partial_t u(j\delta + s, \cdot) - \partial_t u((j+1)\delta, \cdot) \, ds \right\| \\
&\leq 1/\delta \int_0^{\delta} \|\partial_t u(j\delta + s, \cdot) - \partial_t u((j+1)\delta, \cdot)\| \, ds \leq 1/\delta \int_0^{\delta} \delta \max_{\xi \in [0,\delta]} \|\partial_{tt} u(j\delta + \xi, \cdot)\| \, ds \lesssim \delta.
\end{aligned}$$

Similarly, the general mean value theorem and the approximation property in (3) imply

$$\|1/\delta((P_h u^{j+1} - u^{j+1}) - (P_h u^j - u^j))\| \leq ch^k \|\partial_t u\|_{L^{\infty}([jk,(j+1)k];H^k(\Omega))}.$$

Combining these estimates with Theorem 2, we get

$$\big|(z^j | v_h)\big| \leq (O(\delta) + O(h^{k+1})) \|v_h\|. \tag{23}$$

Using the notation $y_h^j = P_h u^j - u_h^j$ in the difference of (21) and (22), we obtain that

$$1/\delta\big(y_h^{j+1} - y_h^j | v_h\big) + \big(A_h^{\alpha} y_h^{j+1} | v_h\big) = \big(z^j | v_h\big) \qquad \forall v_h \in V_h.$$

Taking here $v_h = y_h^{j+1}$ and rearranging the equality, we have

$$\|y_h^{j+1}\|^2 + \delta\big(A_h^{\alpha} y_h^{j+1} | y_h^{j+1}\big) = \big(y_h^j + \delta z^j | y_h^{j+1}\big),$$

and therefore, using the positivity of $A_h^{\alpha}$ and (23), we get

$$\|y_h^{j+1}\|^2 \leq \|y_h^j\| \|y_h^{j+1}\| + \delta\big(z^j | y_h^{j+1}\big) \leq \|y_h^j\| \|y_h^{j+1}\| + \delta \cdot O(\delta + h^k) \|y_h^{j+1}\|,$$

which results in the following estimate:

$$\|y_h^{j+1}\| \leq \|y_h^j\| + \delta(O(\delta + h^k)).$$

A consecutive application of this inequality gives

$$\max_{0 \leq j \leq M} \|y_h^j\| \leq \|y_h^0\| + T \cdot O(\delta + h^k),$$

which, together with (3) can be used to get the final error estimator

$$\max_{0 \le j \le M} \|u^j - u_h^j\| \le \max_{0 \le j \le M} \left( \|u^j - P_h u^j\| + \|P_h u^j - u_h^j\| \right) = \max_{0 \le j \le M} (\|u^j - P_h u^j\| + \|y_h^j\|) = O(\delta + h^k),$$

as stated in the theorem.

## 4. Numerical method

For the numerical solution of problem (1), we need to use an efficient method avoiding the direct computation of the $\alpha$-th power of the matrices. Such an approach was first proposed in [21], which can be applied immediately only for explicit time stepping. Here we propose an algorithm to approximate the implicit time stepping $\mathbf{u}^{j+1} = (I + \delta A_h^\alpha)^{-1}\mathbf{u}^j$, i.e. $\mathbf{u}^n = (I + \delta A_h^\alpha)^{-n}\mathbf{u}^0$ without computing matrix powers or solving linear systems with a dense matrix. We also assume here that $A_h$ is symmetric, which will be satisfied for the finite element space in the numerical experiments.
In a general situation, it is also satisfied if the finite element basis functions are translations of each other. In practice, we can use this if the domain is approximated using a square grid.

In a general situation, for arbitrary finite elements, an $L^2$-orthogonal basis implies $M_h = I_h$ such that $A_h = I_h^{-1}S_h$ becomes symmetric. This, however, leads to a full matrix $A_h$, which will slow down our algorithm.

For this, we will combine the algorithm in [21] with a conjugate gradient method.

### 4.1. The algorithm

The proposed method consists of the following steps.

(i) Following the method in [21], we compute the $k_1$ smallest and $k_2$ largest eigenvalues. In an increasing order with $\bar{k} = k_1 + k_2$ these and the corresponding eigenvectors are denoted by $\lambda_1, \lambda_2, \ldots, \lambda_{\bar{k}}$ of $A_h$ and $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{\bar{k}}$, respectively.

(ii) Let $X = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{\bar{k}}] \in \mathbb{R}^{n \times \bar{k}}$ denote the matrix composed of these eigenvectors, and $Q_{\bar{k}} = XX^T$ the orthogonal projection matrix to the subspace span $\{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{\bar{k}}\}$. In this case, $I - Q_{\bar{k}}$ is the orthogonal projection matrix to the complementary subspace.

(iii) The problem is divided then into two parts:

$$(I + \delta A_h^\alpha)^{-n}\mathbf{u}^j = (I + \delta A_h^\alpha)^{-n}Q_{\bar{k}}\mathbf{u}^j + (I + \delta A_h^\alpha)^{-n}(I - Q_{\bar{k}})\mathbf{u}^j \tag{24}$$

(a) We can directly compute the first part, since we already know the corresponding eigenvalues:

$$(I + \delta A_h^\alpha)^{-n}Q_{\bar{k}}\mathbf{u}^j = X\Lambda^n X^T\mathbf{u}^j,$$

where $\Lambda$ denotes the diagonal matrix consisting of the elements $\frac{1}{1+\delta\lambda_1^\alpha}, \frac{1}{1+\delta\lambda_2^\alpha}, \ldots, \frac{1}{1+\delta\lambda_{\bar{k}}^\alpha}$.

(b) To approximate $(I + \delta A_h^\alpha)^{-n}(I - Q_{\bar{k}})\mathbf{u}^j$, we apply a conjugate gradient method $CG$, see, *e.g.* [], such that

$$CG\, \mathbf{w} \approx (I + A_h^\alpha)^{-1}\mathbf{w}.$$

In the steps of the conjugate gradient algorithm, we use an approximation of the matrix–vector products $(I + \delta A_h^\alpha)\mathbf{w}$ without computing $A_h^\alpha$. This is performed using the Taylor series approach

$$\left(\frac{2A_h}{\sigma(A)}\right)^\alpha \mathbf{w} = \sum_{n=0}^{\infty}\binom{\alpha}{n}\left(\frac{2A_h}{\sigma(A)} - I\right)^n\mathbf{w} \approx \sum_{n=0}^{K}\binom{\alpha}{n}\left(\frac{2A_h}{\sigma(A)} - I\right)^n\mathbf{w} := \left(\frac{2}{\sigma(A)}\right)^\alpha T(A_h, \alpha)\mathbf{w}, \tag{25}$$

where $\sigma(A)$ denotes the spectral radius of $A$.

**Remark.** The conjugate gradient algorithm is suitable here as we have a symmetric positive definite matrix $I + A_h^\alpha$. The stopping criterion (or tolerance) for this procedure is given by discussing the numerical experiments.

An important parameter in the approximation in (25) is the parameter $K$. An estimation of this is given in Section 2.3 in [21], which motivated our choice in Section 4.3.

The operations of the conjugate gradient method are invariant to the subspace ran$(I - Q_{\bar{k}})$, thus the Taylor-series method will converge quickly in every time step. Also, in the Taylor series approach, we use only sparse matrix–vector products such that beyond the eigenvalue approximation, the entire algorithm involves only these very quick operations.

### 4.2. Error analysis of the algorithm

Recall that in Theorem 3, we estimated the difference between the analytic solution and the numerical solution based on the implicit Euler time stepping. In the practice, however, according to (iii)(b) in the above algorithm, we do not apply directly the implicit time steps. In concrete terms, we compute $CG^n\mathbf{w}^0$ instead of $(I + \delta A_h^\alpha)^{-n}\mathbf{w}^0$, where $\mathbf{w}^0 = (I - Q_{\bar{k}})\mathbf{u}^0$.

To estimate the extra error term arising from this approximation, we also use the notation $T(A_h, \alpha)$ for the Taylor series approximation of $A_h^\alpha$. We assume that $T(A_h, \alpha)$ gives also a positive definite matrix with the minimal eigenvalue $\lambda_{1,T}$, which is satisfied for any reasonable approximation. Since $A_h^\alpha$ is also positive definite, for all exponents $n \in \mathbb{N}$, we have

$$\|(I + \delta T(A_h, \alpha))^{-1}\| \leq \frac{1}{1 + \lambda_{1,T}} \leq 1 \quad \text{and} \quad \|((I + \delta A_h^\alpha))^{-1}\| \leq 1. \tag{26}$$

With these, a triangle inequality gives

$$\begin{aligned}
&\|(I + \delta A_h^\alpha)^{-n}\mathbf{w}^0 - CG^n\mathbf{w}^0\| \\
&\leq \|(I + \delta A_h^\alpha)^{-n}\mathbf{w}^0 - (I + \delta T(A_h, \alpha))^{-n}\mathbf{w}^0\| + \|CG^n\mathbf{w}^0 - (I + \delta T(A_h, \alpha))^{-n}\mathbf{w}^0\|,
\end{aligned} \tag{27}$$

where the first term, using (26) can be estimated further as

$$\begin{aligned}
&\|(I + \delta A_h^\alpha)^{-n}\mathbf{w}^0 - (I + \delta T(A_h, \alpha))^{-n}\mathbf{w}^0\| \\
&\leq \|(I + \delta A_h^\alpha)^{-n}[(I + \delta A_h^\alpha)^n - (I + \delta T(A_h, \alpha))^n](I + \delta T(A_h, \alpha))^{-n}\| \cdot \|\mathbf{w}^0\| \\
&\leq \delta \cdot \|\mathbf{w}^0\| \cdot \|(I + \delta A_h^\alpha)^{-n}\| \cdot \|(I + \delta T(A_h, \alpha))^{-n}\| \\
&\quad \cdot \|(A_h^\alpha - T(A_h, \alpha)) \cdot ((I + \delta A_h^\alpha)^{n-1} + (I + \delta A_h^\alpha)^{n-2}(I + \delta T(A_h, \alpha)) + \cdots + (I + \delta T(A_h, \alpha))^{n-1})\| \\
&\leq n\delta \cdot \|\mathbf{w}^0\| \cdot \|[A_h^\alpha - T(A_h, \alpha)]\| = T \cdot \|\mathbf{w}^0\|\|A_h^\alpha - T(A_h, \alpha)\|.
\end{aligned} \tag{28}$$

Note that a computable upper bound for the Taylor's remainder term $\|A_h^\alpha - T(A_h, \alpha)\|$ was developed in [21].

To estimate the second term in (27), we recall that a time step in (iii)(b) has the form

$$\mathbf{w}^{j+1} = CG\mathbf{w}^j,$$

such that we define the computational error $\mathbf{e}^{j+1}$ regarding the CG method as

$$\mathbf{e}^{j+1} = (I + \delta T(A_h, \alpha))^{-1}\mathbf{w}^j - CG\mathbf{w}^j = (I + \delta T(A_h, \alpha))^{-1}\mathbf{w}^j - \mathbf{w}^{j+1}, \qquad j = 0, 1, \ldots, N - 1,$$

which implies

$$(I + \delta T(A_h, \alpha))^{-1}\mathbf{w}^j = \mathbf{w}^{j+1} + \mathbf{e}^{j+1}, \qquad j = 0, 1, \ldots, N - 1. \tag{29}$$

Note that the error term, using again (26), can be controlled easily based on the estimation

$$\begin{aligned}
\|\mathbf{e}^{j+1}\| &\leq \|(I + T(A_h, \alpha))^{-1}\|\|(I + \delta T(A_h, \alpha))\mathbf{e}^{j+1}\| \\
&= \|(I + T(A_h, \alpha))^{-1}\|\|(I + \delta T(A_h, \alpha))((I + \delta T(A_h, \alpha))^{-1}\mathbf{w}^j - \mathbf{w}^{j+1})\| \leq \|\mathbf{w}_j - (I + \delta T(A_h, \alpha))\mathbf{w}^{j+1}\|,
\end{aligned}$$

where the right-hand side can be computed using a sparse matrix–vector product.

Using (29) in the second term of (27), we obtain

$$\begin{aligned}
(I + \delta T(A_h, \alpha))^{-n}\mathbf{w}^0 - CG^n\mathbf{w}^0 &= (I + \delta T(A_h, \alpha))^{-(n-1)}\mathbf{w}^1 + (I + \delta T(A_h, \alpha))^{-(n-1)}\mathbf{e}^1 - CG^{n-1}\mathbf{w}^1 \\
&= (I + \delta T(A_h, \alpha))^{-(n-2)}\mathbf{w}^2 + (I + \delta T(A_h, \alpha))^{-(n-2)}\mathbf{e}^2 + (I + \delta T(A_h, \alpha))^{-(n-1)}\mathbf{e}^1 - CG^{n-2}\mathbf{w}^2 \\
\cdots &= \mathbf{w}^n + \mathbf{e}^n + (I + \delta T(A_h, \alpha))^{-1}\mathbf{e}^{n-1} + \cdots + (I + \delta T(A_h, \alpha))^{-(n-1)}\mathbf{e}^1 - \mathbf{w}^n.
\end{aligned}$$

Accordingly, using (26), we obtain the estimate

$$\begin{aligned}
\|CG^n\mathbf{w}^0 - (I + \delta T(A_h, \alpha))^{-n}\mathbf{w}^0\| &\leq \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \sum_{j=0}^{n-1} \|I + \delta T(A_h, \alpha)\|^{-j} \\
&= \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \sum_{j=0}^{n-1} \frac{1}{(1 + \lambda_{1,T})^j} \leq \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \frac{1 - \frac{1}{(1+\lambda_{1,T})^n}}{1 - \frac{1}{1+\lambda_{1,T}}} \leq \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \frac{1}{1 - \frac{1}{1+\lambda_{1,T}}} \\
&= \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \left(1 + \frac{1}{\lambda_{1,T}}\right).
\end{aligned} \tag{30}$$

Summarized, the inserting (28) and (30) into (27) gives the following upper bound for the computational error:

$$\|(I + \delta A_h^\alpha)^{-n}\mathbf{w}^0 - CG^n\mathbf{w}^0\| \leq T \cdot \|\mathbf{w}^0\|\|A_h^\alpha - T(A_h, \alpha)\| + \max_{j=1,2,\ldots,N} \|\mathbf{e}^j\| \cdot \left(1 + \frac{1}{\lambda_{1,T}}\right). \tag{31}$$

Here, as mentioned, all of the terms can be controlled, moreover, the estimate is independent of the number of time steps.

### 4.3. Numerical experiments

We investigate the model problem

$$
\begin{cases}
\partial_t u(t, y) = -10^{-3}(-\Delta_{\mathcal{D}})^{0.75} u(t, x, y) & t \in (0, 1], \ (x, y) \in (0, 1) \times (0, 1) \\
u(t, x, y) = 0 & t \in (0, 1], \ (x, y) \in \partial\Omega \\
u(0, x, y) = \displaystyle\sum_{j=1}^{100} \frac{1}{j^2} \sin(7\pi jx) \cdot \sin(9\pi jy) & (x, y) \in (0, 1) \times (0, 1),
\end{cases}
$$

which has the analytic solution

$$
u(t, x, y) = \sum_{j=1}^{100} \frac{1}{j^2} \sin(7\pi jx) \sin(9\pi jy) \exp\left[-10^{-3}(t((7j)^2 + (9j)^2)\pi^2)^{0.75}\right].
$$

The computational domain was tessellated uniformly and locally, $\mathbf{Q}_1$ elements were used to constitute the finite element space. Since they contain all piecewise linear functions, by means of Corollary 1.109 in [23], the spatial approximation order is $k = 2$ in (3). Therefore, according to 3, we expect second-order convergence spatially. Also, this choice ensures that $A_h$ is symmetric.

In the conjugate gradient algorithm, we have used the tolerance $\max_{j=1,2,\dots,N} \|e_j\| = 10^{-6}$ and for a fine spatial grid, K=100, 200 and 300 terms in the Taylor expansion. To accelerate the convergence of this, we have computed 10 eigenvalues in step (i) of our algorithm. Corresponding to (31), these are the chief parameters in our estimation.

In each case, the computational error was computed in a discrete $L^2$ norm and we have displayed the computing time in seconds.

After the consecutive refinements, we have estimated the convergence rate, which is shortly called rate in the tables.

In the first table, the result for different time discretizations is given estimating also the temporal convergence rate by fixing $h = 0.005$.

| $\delta$ | 0.1 | 0.05 | 0.025 | 0.02 | 0.0166 | 0.0125 |
|---|---|---|---|---|---|---|
| error | 0.0018 | $7.305 \cdot 10^{-4}$ | $2.011 \cdot 10^{-4}$ | $1.038 \cdot 10^{-4}$ | $6.3449 \cdot 10^{-5}$ | $1.017 \cdot 10^{-4}$ |
| time | 39.6 | 53.97 | 77.46 | 89.24 | 99.08 | 120.74 |
| rate | 1.0923 | 1.2847 | 1.8612 | 2.3187 | 2.5675 | 0.9832 |

In a second series of experiments, we investigated the spatial convergence rate using for several spatial discretization parameters $h$ by fixing $\delta = 0.01$ such that, practically, the accuracy depends mainly on the spatial discretization parameter $h$. To depict the efficiency of the algorithm, we have also displayed the number of unknowns (DOF) in the corresponding linear system and the effect of choosing different number $K$ of terms in the Taylor expansion. An extra step in our algorithm compared to any other approaches is the computation of some eigenvalues of the stiffness matrix. Therefore, the corresponding computing time is also shown in the following table.

| $h$ | 0.025 | 0.0166 | 0.0125 | 0.01 | 0.00833 | 0.00714 | 0.00625 | 0.005 |
|---|---|---|---|---|---|---|---|---|
| DOF | 1521 | 3481 | 6241 | 9801 | 14161 | 19321 | 25281 | 39601 |
| eig. time | 0.026 | 0.065 | 0.117 | 0.207 | 0.325 | 0.411 | 0.678 | 1.133 |
| time, K=100 | 1.19 | 2.97 | 5.36 | 12.14 | 17.96 | 26.61 | 34.17 | 56.04 |
| time, K=200 | 2.28 | 5.64 | 10.45 | 23.52 | 33.47 | 48.4 | 63.52 | 96.58 |
| time, K=300 | 3.34 | 8.36 | 15.58 | 34.76 | 48.86 | 69.44 | 90.17 | 139.03 |
| error, K=100 | 0.0101 | 0.0042 | 0.0022 | 0.0013 | $7.96 \cdot 10^{-4}$ | $5.19 \cdot 10^{-4}$ | $6.58 \cdot 10^{-4}$ | 0.0019 |
| error, K=200 | 0.0101 | 0.0042 | 0.0022 | 0.0013 | $7.84 \cdot 10^{-4}$ | $5.13 \cdot 10^{-4}$ | $3.46 \cdot 10^{-4}$ | $2.61 \cdot 10^{-4}$ |
| error, K=300 | 0.0101 | 0.0042 | 0.0022 | 0.0013 | $7.84 \cdot 10^{-4}$ | $5.12 \cdot 10^{-4}$ | $3.41 \cdot 10^{-4}$ | $1.52 \cdot 10^{-4}$ |
| rate, K=100 | 1.713 | 2.043 | 2.218 | 2.329 | 2.406 | 2.346 | 1.723 | −0.558 |
| rate, K=200 | 1.713 | 2.043 | 2.219 | 2.331 | 2.428 | 2.528 | 2.649 | 2.269 |
| rate, K=300 | 1.713 | 2.043 | 2.219 | 2.331 | 2.428 | 2.531 | 2.670 | 3.052 |

One can observe that in this case, an optimal choice of $K$ is a few hundreds, which ensures already the desired convergence rate. It is also clear, that we really need "long" Taylor approximations. An error analysis for this can be found in [21]. It also clear that the computation of the eigenvalues requires minimal extra efforts.

Also, the convergence results in the tables are in a good accordance with our theoretical results: the spatial convergence order is above two and the temporal order is about one. Finally, the computational time remains proportional with the number of unknowns. This indicates that only operations with sparse matrices were used in the algorithm.

### CRediT authorship contribution statement

**Gábor Maros:** Derivation of error estimates, Numerical experiments. **Ferenc Izsák:** Conceptualization, Corrections, Editing, Supervision.

## Acknowledgments

## References

[1] A.M. Edwards, et al., Revisiting Lévy flight search patterns of wandering albatrosses, bumblebees and deer, Nature 449 (2007) 1044–1048, http://dx.doi.org/10.1038/nature06199.

[2] N. Humphries, et al., Environmental context explains Lévy and Brownian movement patterns of marine predators, Nature 465 (2010) 1066–1069, http://dx.doi.org/10.1038/nature09116.

[3] F. Höfling, T. Franosch, Anomalous transport in the crowded world of biological cells, Rep. Progr. Phys. 76 (4) (2013) 046602, http://dx.doi.org/10.1088/0034-4885/76/4/046602.

[4] G. Pagnini, Fractional kinetics in random/complex media, in: Handbook of Fractional Calculus with Applications, vol. 5, De Gruyter, Berlin, 2019, pp. 183–205.

[5] M. Kwaśnicki, Ten equivalent definitions of the fractional Laplace operator, Fract. Calc. Appl. Anal. 20 (1) (2017) 7–51, http://dx.doi.org/10.1515/fca-2017-0002.

[6] G. Acosta, J.P. Borthagaray, A fractional Laplace equation: regularity of solutions and finite element approximations, SIAM J. Numer. Anal. 55 (2) (2017) 472–495, http://dx.doi.org/10.1137/15M1033952.

[7] G. Acosta, F.M. Bersetche, J.P. Borthagaray, A short FE implementation for a 2d homogeneous Dirichlet problem of a fractional Laplacian, Comput. Math. Appl. 74 (4) (2017) 784–816, http://dx.doi.org/10.1016/j.camwa.2017.05.026.

[8] V.J. Ervin, N. Heuer, J.P. Roop, Regularity of the solution to 1-D fractional order diffusion equations, Math. Comp. 87 (313) (2018) 2273–2294, http://dx.doi.org/10.1090/mcom/3295.

[9] F. Izsák, B.J. Szekeres, Models of space-fractional diffusion: A critical review, Appl. Math. Lett. 71 (2017) 38–43, http://dx.doi.org/10.1016/j.aml.2017.03.006.

[10] A. Bonito, J.P. Borthagaray, R.H. Nochetto, E. Otárola, A.J. Salgado, Numerical methods for fractional diffusion, Comput. Vis. Sci. 19 (5–6) (2018) 19–46, http://dx.doi.org/10.1007/s00791-018-0289-y.

[11] A. Lischke, G. Pang, M. Gulian, et al., What is the fractional Laplacian? A comparative review with new results, J. Comput. Phys. 404 (2020) 109009, 62, http://dx.doi.org/10.1016/j.jcp.2019.109009.

[12] M. Ilić, F. Liu, I. Turner, V. Anh, Numerical approximation of a fractional-in-space diffusion equation. I, Fract. Calc. Appl. Anal. 8 (3) (2005) 323–341.

[13] B.J. Szekeres, F. Izsák, Finite element approximation of fractional order elliptic boundary value problems, J. Comput. Appl. Math. 292 (2016) 553–561, http://dx.doi.org/10.1016/j.cam.2015.07.026.

[14] A. Bonito, J.E. Pasciak, Numerical approximation of fractional powers of elliptic operators, Math. Comp. 84 (295) (2015) 2083–2110, http://dx.doi.org/10.1090/S0025-5718-2015-02937-8.

[15] A. Bonito, W. Lei, J.E. Pasciak, The approximation of parabolic equations involving fractional powers of elliptic operators, J. Comput. Appl. Math. 315 (2017) 32–48, http://dx.doi.org/10.1016/j.cam.2016.10.016.

[16] P.N. Vabishchevich, Numerical solution of nonstationary problems for a space-fractional diffusion equation, Fract. Calc. Appl. Anal. 19 (1) (2016) 116–139, http://dx.doi.org/10.1515/fca-2016-0007.

[17] P.N. Vabishchevich, Numerical solution of time-dependent problems with fractional power elliptic operator, Comput. Methods Appl. Math. 18 (1) (2018) 111–128, http://dx.doi.org/10.1515/cmam-2017-0028.

[18] S. Harizanov, R. Lazarov, S. Margenov, P. Marinov, Numerical solution of fractional diffusion–reaction problems based on BURA, Comput. Math. Appl. (2019) http://dx.doi.org/10.1016/j.camwa.2019.07.002.

[19] M. Karkulik, J.M. Melenk, $\mathcal{H}$-Matrix approximability of inverses of discretizations of the fractional Laplacian, Adv. Comput. Math. 45 (5–6) (2019) 2893–2919, http://dx.doi.org/10.1007/s10444-019-09718-5.

[20] C. Vollmann, V. Schulz, Exploiting multilevel toeplitz structures in high dimensional nonlocal diffusion, Comput. Vis. Sci. 20 (1–2) (2019) 29–46, http://dx.doi.org/10.1007/s00791-018-00306-6.

[21] F. Izsák, B.J. Szekeres, Efficient computation of matrix power-vector products: Application for space-fractional diffusion problems, Appl. Math. Lett. 86 (2018) 70–76, http://dx.doi.org/10.1016/j.aml.2018.06.019.

[22] R.H. Nochetto, E. Otárola, A.J. Salgado, A PDE approach to fractional diffusion in general domains: a priori error analysis, Found. Comput. Math. 15 (2015) 733–791, http://dx.doi.org/10.1007/s10208-014-9208-x.

[23] A. Ern, J.-L. Guermond, Theory and practice of finite elements, Applied Mathematical Sciences, vol. 159, Springer-Verlag, New York, 2004, p. xiv+524.

[24] B.J. Szekeres, F. Izsák, Finite difference approximation of space-fractional diffusion problems: the matrix transformation method, Comput. Math. Appl. 73 (2) (2017) 261–269, http://dx.doi.org/10.1016/j.camwa.2016.11.021.

[25] A.V. Balakrishnan, Fractional powers of closed operators and the semigroups generated by them, Pacific J. Math. 10 (1960) 419–437.

[26] L.A. Caffarelli, P.R. Stinga, Fractional elliptic equations, Caccioppoli estimates and regularity, Ann. Inst. H. Poincaré Anal. Non Linéaire 33 (3) (2016) 767–807, http://dx.doi.org/10.1016/j.anihpc.2015.01.004.