Padé-parametric FEM approximation for fractional powers of elliptic operators on manifolds

Beiping Duan*

*Faculty of Computational Mathematics and Cybernetics, Shenzhen MSU-BIT University, Shenzhen 518172, P.R. China.
E-mail address: duanbeiping@smbu.edu.cn.

Aug. 8, 2023

Content

Background and existing methods

Definition of fractional powers of elliptic operators:

Suppose $\mathcal{L}$ is a positive definite operator, $s \in \mathbb{C}$, then the fractional power of $\mathcal{L}$ is defined by the Dunford-Taylor formula

$$\mathcal{L}^s = \frac{1}{2\pi i} \int_{\mathcal{C}} z^s (z - \mathcal{L})^{-1} dz,$$

with $\mathcal{C}$ a contour on complex plane surrounds the spectrum of $\mathcal{L}$.
We aim at solving the following problem

$$\mathcal{L}^\alpha u = f, \quad \alpha \in (0,1). \tag{1.1}$$

Existing approaches:

- quadrature schemes based on integral representations, e.g.,

$$\text{Bonito-Pasciak}^{1}: \mathcal{L}^{-\alpha} f = \frac{2\sin\pi\alpha}{\pi}\int_{-\infty}^{\infty} e^{2\alpha y}(\mathcal{I} + 2^{2y}\mathcal{L})^{-1} f\, dy$$

Quadrature error(trapezoidal rule): $\mathcal{O}(e^{-\pi\sqrt{\alpha(1-\alpha)N_s}})$

Aceto-Novati[2] :

$$\mathcal{L}^{-\alpha} f = \frac{2\sin\pi\alpha\tau^{-\alpha}}{\pi}\int_{-1}^{1} w^{-\alpha,\alpha-1}\left(\tau(1-t) + (1+t)\frac{\mathcal{L}}{\tau}\right)^{-1} f\, dt$$

Quadrature error(Gauss-Jacobi): $\mathcal{O}(\rho^{-N_s/\lambda_{h,\max}^{1/4}})$

[1] A. Bonito, and J. Pasciak. Numerical approximation of fractional powers of elliptic operators. Mathematics of Computation 84, no. 295 (2015): 2083-2110.

[2] L. Aceto, and P. Novati. Rational approximations to fractional powers of self-adjoint positive operators. Numerische Mathematik 143 (2019): 1-16.

- utilizing the best uniform rational approximations[3];

$$\text{BURA: } \mathcal{L}^{-\alpha} f \approx \lambda_1^{-\alpha} r_{k,\alpha}(\lambda_1 \mathcal{L}^{-1}) f$$

- solving Caffarelli-Silvestre extension problem (for $(-\Delta)^\alpha u = f$)

$$\begin{cases} \nabla \cdot (t^{1-2\alpha} \nabla U(x,t)) = 0, \\ -\lim\limits_{t\to 0^+} U_t(x,t) = 2^{1-2\alpha} \dfrac{\Gamma(1-\alpha)}{\Gamma(\alpha)} f, \end{cases}$$

  then $u(x) = U(x,0)$.
- solving Vabishchevich extension problem[4]

$$\begin{cases} U_t(x,t) + \alpha \mathcal{B}(\hat{\lambda}\mathcal{I} + t\mathcal{B})^{-1} U(x,t) = 0, \\ U(x,0) = \hat{\lambda}^{-\alpha} f, \end{cases}$$

  with $\mathcal{B} = \mathcal{L} - \hat{\lambda}\mathcal{I}$, then $u(x) = U(x,1)$.

[3] S. Harizanov, R. Lazarov, S. Margenov, P. Marinov, & J. Pasciak. (2020). Analysis of numerical methods for spectral fractional elliptic equations based on the best uniform rational approximation. Journal of Computational Physics, 408, 109285.

[4] P.N. Vabishchevich, Numerically solving an equation for fractional powers of elliptic operators, J. Comput. Phys. 282 (2015) 289–302.

Existing issues:

- Few are robust with respect to α, say, $\alpha \to 0^+, 1^-$;
- All focus on Euclidean cases, except [Bonito-Lei, 2021][5].

[5] A. Bonito, and W. Lei. Approximation of the Spectral Fractional Powers of the Laplace-Beltrami Operator. Numerical Mathematics: Theory, Methods & Applications 15.4, 2022.

Content

Notations

We consider $\mathcal{L}$ as an operator defined on a C^3-smooth manifold. Let

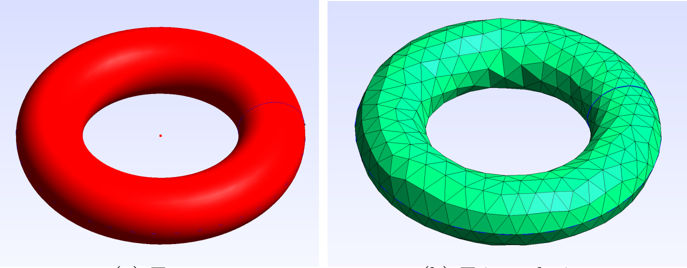
$$\begin{aligned} \mathcal{M}: &\quad C^3 - \text{smooth manifold with boundary } \Gamma(\text{may be empty}); \\ a(\vec{x}): &\quad C^\kappa(\mathcal{M}), \text{real, scalar;} \\ b(\vec{x}): &\quad C^\kappa(\mathcal{M}), \text{real, scalar;} \end{aligned}$$

and $\mathcal{L} := \mathcal{L}_\mathcal{M}$, given by for $\forall w, v \in \mathbb{H}^1(\mathcal{M})$

$$\int_\mathcal{M} \mathcal{L}_\mathcal{M} w \, v d\mathcal{M} = \int_\mathcal{M} a(\vec{x}) \nabla_\mathcal{M} w (\nabla_\mathcal{M} v)^T + b(\vec{x}) w v \, d\mathcal{M}$$

Then **our target** is solving

$$\mathcal{L}_\mathcal{M}^\alpha u = f \quad \text{or equivalently } u = \mathcal{L}_\mathcal{M}^{-\alpha} f.$$



(a) Torus (b) Triangulation

Figure 1: Triangulation for a torus

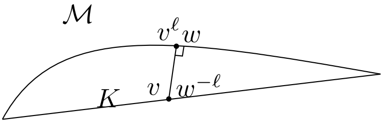


Figure 2: Lift operation

The scheme can be splitted into two steps:

- Approximate $\mathcal{M}$ by union of finitely many non-degenerate 2-dimensional simplices, then approximate $\mathcal{L}_{\mathcal{M}}^{-\alpha}$ by $\mathcal{L}_{\mathcal{M}_h}^{-\alpha}$ with

$$\begin{aligned}&\int_{\mathcal{M}_h} \mathcal{L}_{\mathcal{M}_h} w_h \, v_h d\mathcal{M}_h \\ &= \int_{\mathcal{M}_h} a_h(\vec{x}) \nabla_{\mathcal{M}_h} w_h (\nabla_{\mathcal{M}_h} v_h)^T + b_h(\vec{x}) w_h v_h d\mathcal{M}_h,\end{aligned} \tag{2.1}$$

where $a_h = I_h(a^{-\ell})$, $b_h = I_h(b^{-\ell})$ with I_h the piecewise linear interpolation operator on $\mathcal{M}_h$.

- Reformulate $\mathcal{L}_{\mathcal{M}_h}^{-\alpha}$ into

$$\mathcal{L}_{\mathcal{M}_h}^{-\alpha} = \frac{1}{\hat{\lambda}^\alpha} \prod_{k=0}^{L} \left[(\hat{\lambda}\mathcal{I} + t_{k+1}\mathcal{B}_h)(\hat{\lambda}\mathcal{I} + t_k\mathcal{B}_h)^{-1} \right]^{-\alpha}$$
$$= \frac{1}{\hat{\lambda}^\alpha} \prod_{k=0}^{L} \left[1 + \tau_k \mathcal{B}_h (\hat{\lambda}\mathcal{I} + t_k\mathcal{B}_h)^{-1} \right]^{-\alpha}$$

where $0 = t_0 < t_1 < \cdots < t_L < t_{L+1} = 1$, $\tau_k = t_{k+1} - t_k$, $\hat{\lambda} \in (0, \lambda_{1,h}]$ and $\mathcal{B}_h = \mathcal{L}_{\mathcal{M}_h} - \hat{\lambda}\mathcal{I}$.
Then approximate each term $(1+t)^{-\alpha}$ by diagonal Padé approximation $r_m(t)$ which gives

$$\left[1 + \tau_k \mathcal{B}_h (\hat{\lambda}\mathcal{I} + t_k\mathcal{B}_h)^{-1} \right]^{-\alpha} \approx r_m(\tau_k \mathcal{B}_h (\hat{\lambda}\mathcal{I} + t_k\mathcal{B}_h)^{-1}).$$

Why rational approximation?

For scalar case,

$$(1+t)^{-\alpha} \approx c_0 t + c_1 t + c_2 t^2 + \cdots, c_n t^n.$$

For operator function, polynomial cannot do this job:

$$(1+(-\Delta))^{-\alpha} f \approx [c_0 + c_1(-\Delta) + c_2(-\Delta)^2 + \cdots, c_n(-\Delta)^n] f.$$

We hope to find rational approximation with the following form:

$$(1+(-\Delta))^{-\alpha} f \approx \frac{a_1 + b_1(-\Delta)}{1 + c_1(-\Delta)} \cdot \frac{a_2 + b_2(-\Delta)}{1 + c_2(-\Delta)} \cdots \frac{a_n + b_n(-\Delta)}{1 + c_n(-\Delta)} f$$

$$\frac{a_i + b_i(-\Delta)}{1 + c_i(-\Delta)} f = \kappa + \frac{\kappa'}{1 + c_i(-\Delta)} f$$

Solve $(1 - c_i\Delta) w = f$.

Main results

Denote by

$$u = \mathcal{L}_{\mathcal{M}}^{-\alpha} f$$

$$u_h = \mathcal{L}_{\mathcal{M}_h}^{-\alpha} f_h$$

$$U_h^{L+1} = \frac{1}{\hat{\lambda}^{\alpha}} \prod_{k=0}^{L} r_m(\tau_k \mathcal{B}_h (\hat{\lambda}\mathcal{I} + t_k \mathcal{B}_h)^{-1}) f_h.$$

Assumption 2.1

For given $f \in \dot{\mathbb{H}}^{\delta}(\mathcal{M})$, we assume that $f_h \in V_h$ satisfies the stability condition $\|f_h\|_h \le c\|f\|_{\dot{H}^{\delta}(\mathcal{M})}$ and processes the approximation property

$$\|f_h^{\ell} - f\|_{H^{-1}(\mathcal{M})} \le ch^{1+\delta}\|f\|_{\dot{H}^{\delta}(\mathcal{M})}. \tag{2.2}$$

Theorem 2.1 (Parametric FEM error)

Suppose $f \in \dot{\mathbb{H}}^\delta(\mathcal{M})$ and f_h satisfies Assumption 2.1. Then for h small enough

$$\|u - u_h^\ell\|_{L^2(\mathcal{M})} \le c \begin{cases} |2\alpha + \delta - 2|^{-1} h^{\min(\delta+2\alpha,2)} \|f\|_{\dot{H}^\delta(\mathcal{M})}, & \delta + 2\alpha \neq 2 \\ |\ln h| h^2 \|f\|_{\dot{H}^\delta(\mathcal{M})}, & \delta + 2\alpha = 2 \end{cases}$$

with c independent of α and h.

Theorem 2.2 (Padé approximation error)

Let $\hat{\lambda} \in (0, \lambda_{1,h}]$ then

$$\left\| u_h - U_h^{L+1} \right\|_{L^2(\mathcal{M}_h)} \le \hat{c} \hat{\lambda}^{-\alpha}\, 32^{-\frac{N_s}{\lceil \log_2(\lambda_{h,max}/\hat{\lambda}) \rceil}} \|f_h\|_{L^2(\mathcal{M}_h)}$$

where N_s is the number of total solves, $\hat{c} \approx \frac{(\alpha+2)2^{\alpha-1}\pi}{\Gamma(1-\alpha)\Gamma(1+\alpha)}$.

Note $\hat{c}$ is bounded with respect to α, which implies our scheme is robust for $\alpha \in (0,1)$.

Content

Padé approximation

We denote the L, M Padé approximation to $A(t)$ by

$$[L, M] = \frac{P_L}{Q_M}$$

where P_L is a polynomial of degree at most L, and Q_M is a polynomial of degree at most M. The formal power series

$$A(t) = \sum_{n=0}^{\infty} c_n t^n$$

determines the coefficients of P_L and Q_M by the equation

$$A(t) - \frac{P_L}{Q_M} = \mathcal{O}(t^{L+M+1}).$$

Padé approximation of $(1+t)^{-\alpha}$

Lemma 3.1 (Duan-IMANA)

Suppose $\{t_j(\beta,\gamma)\}_{j=1}^m$ are the roots of $J_m^{\beta,\gamma}(t), t \in [0,1]$ enumerated in increasing order, then

$$r_m(t) = \prod_{i=1}^{m} \frac{1+t_i(\alpha,-\alpha)t}{1+t_i(-\alpha,\alpha)t}. \tag{3.1}$$

To ensure the stability of our scheme, say,

$$\mathcal{L}_{\mathcal{M}_h}^{-\alpha} f_h = \prod_{k=0}^{L} r_m(\tau_k \mathcal{B}_h(\delta\mathcal{I} + t_k\mathcal{B}_h)^{-1}) f_h \tag{3.2}$$

the following interlacing property is needed.

Proposition 3.1 (Duan-IMANA)

For $\alpha \in (0,1)$ it holds

$$\begin{aligned} 0 &< t_1(\alpha,-\alpha) < t_1(-\alpha,\alpha) < t_2(\alpha,-\alpha) < t_2(-\alpha,\alpha) \\ &< \cdots < t_j(\alpha,-\alpha) < t_j(-\alpha,\alpha) < \cdots < t_m(\alpha,-\alpha) < t_m(-\alpha,\alpha) < 1. \end{aligned}$$

Theorem 3.2 (Duan-IMANA)

For $\alpha \in [0,1]$,

$$0 < r_m(t) - (1+t)^{-\alpha} \leq c'_\alpha \frac{2^{-4m} t^{2m+1}}{2^{mt}}, \quad t \in [0,1]$$

with r_m *the* (m,m)*-type Padé approximant and* $c'_\alpha \approx \frac{\alpha\pi}{2\Gamma(1-\alpha)\Gamma(1+\alpha)}$.

Proof. Set $z_n = r_n - r_{n+1}$, then

$$z_n(t) = \frac{P_n(t)Q_{n+1}(t) - P_{n+1}(t)Q_n(t)}{Q_n(t)Q_{n+1}(t)},$$

so

$$Q_n(t)Q_{n+1}(t)z_n(t) = P_n(t)Q_{n+1}(t) - P_{n+1}(t)Q_n(t) \in \mathcal{P}^{2n+1}.$$

This implies

$$P_n(t)Q_{n+1}(t) - P_{n+1}(t)Q_n(t) = \eta_n t^{2n+1}. \tag{3.3}$$

Note

$$P_n(t) = 1 + \sum_{j=1}^{n} a_n^j b_n^j(-\alpha)t^j, \quad Q_n(t) = 1 + \sum_{j=1}^{n} a_n^j b_n^j(\alpha)t^j \tag{3.4}$$

where

$$b_n^j(\alpha) = (n+\alpha)((n-1)+\alpha)\cdots((n+1-j)+\alpha) \tag{3.5}$$

and

$$a_n^j = \frac{n(n-1)\cdots(n+1-j)}{j!2n(2n-1)\cdots(2n+1-j)} \qquad \text{for } j = 1, 2, \ldots, n. \tag{3.6}$$

Manipulations lead to

$$\eta_n = c_\alpha \frac{\Gamma(n+1-\alpha)\Gamma(n+1+\alpha)\Gamma(n+1)\Gamma(n+2)}{\Gamma(2n+1)\Gamma(2n+3)}$$

where $c_\alpha = \frac{2\alpha}{\Gamma(1-\alpha)\Gamma(1+\alpha)}$. Thus, we get $\eta_n > 0$ which implies that $r_n(t) > r_{n+1}(t)$.

Next we shall show the error. Appealing to Stirling's formula it follows

$$\frac{\Gamma(n+1-\alpha)\Gamma(n+1+\alpha)\Gamma(n+1)\Gamma(n+2)}{\Gamma(2n+1)\Gamma(2n+3)} \approx 2\pi 2^{-4n-3} = \frac{\pi}{4}2^{-4n}.$$

Thus we get $\eta_n \leq c'_\alpha 2^{-4n}$ with $c'_\alpha \approx \frac{\pi}{4}c_\alpha$. Telescoping sums,

$$r_m(t) - (1+t)^{-\alpha} = \sum_{n=m}^{\infty}(r_n - r_{n+1}) \leq c'_\alpha \sum_{n=m}^{\infty} \frac{2^{-4n}t^{2n+1}}{Q_n(t)Q_{n+1}(t)}. \tag{3.7}$$

It remains to provide a lower bound for the denominator.

Set ξ_j, for $j = 1, \ldots, n$, to be the roots of the n'th Legendre polynomial $J_n^{0,0}(t)$. Since $Q_n(0) = 1$,

$$Q_n(t) = \prod_{j=1}^{n} (1 + t_j(-\alpha, \alpha)t).$$

Note $t_j(-\alpha, \alpha) > t_j(0, 0)$ so that

$$Q_n(t) \geq \prod_{i=1}^{n} (1 + \xi_i t) := Q_n^0(t), \quad \text{for all} \quad t \geq 0.$$

Now,

$$\log_2 Q_n^0(t) = \sum_{i=1}^{n} \log_2(1 + \xi_i t).$$

Applying the inequalities for the roots of Legendre polynomials,

$$\cos\left(\frac{2i}{2n+1}\pi\right) < 2\xi_i - 1 < \cos\left(\frac{2i-1}{2n+1}\pi\right), \quad i = 1, 2, \cdots, n,$$

we get

$$\log_2 Q_n(t) > \sum_{i=1}^{n} \log_2\left(1 + \frac{t}{2}\left(1 + \cos\frac{2i\pi}{2n+1}\right)\right).$$

Note that $\frac{t}{2}(1+\cos\frac{2i\pi}{2n+1}) \in [0,1]$ and $\log_2(1+\eta) \geq \eta$ for $\eta \in [0,1]$ so

$$\log_2 Q_n(t) > \sum_{i=1}^{n} \frac{t}{2}\left(1+\cos\frac{2i\pi}{2n+1}\right) = \frac{nt}{2} + \frac{t}{2}\sum_{i=1}^{n}\cos\frac{2i\pi}{2n+1}.$$

The well known identity $\sum_{i=1}^{n}\cos\frac{2i\pi}{2n+1} = -\frac{1}{2}$ implies $\log_2 Q_n(t) > \frac{nt}{2} - \frac{t}{4}$ so that

$$Q_n(t)Q_{n+1}(t) \geq 2^{\frac{nt}{2}-\frac{t}{4}}2^{\frac{(n+1)t}{2}-\frac{t}{4}} = 2^{nt}.$$

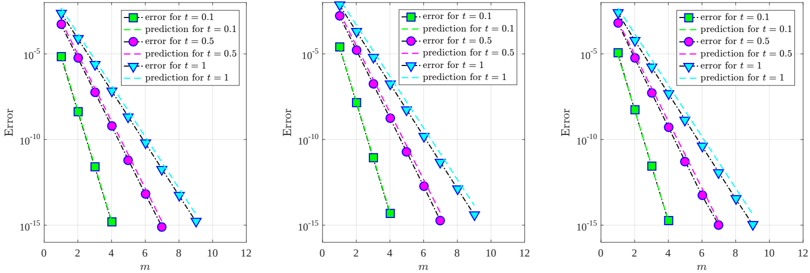
Combining this with (3.7) gives

$$r_m(t) - (1+t)^{-\alpha} \leq \frac{32}{31}c'_\alpha \frac{2^{-4m}t^{2m+1}}{2^{mt}} \quad \text{for } t \in [0,1]$$

with

$$c'_\alpha = \frac{\pi}{4}c_\alpha \approx \frac{\alpha\pi}{2\Gamma(1-\alpha)\Gamma(1+\alpha)}.$$

How sharp is the bound?



(a) $\alpha = 0.1$ (b) $\alpha = 0.5$ (c) $\alpha = 0.9$

Figure 3: $r_m(t) - (1+t)^{-\alpha}$ for different α and t and theoretical predictions.

Recalling

$$u_h = \mathcal{L}_{\mathcal{M}_h}^{-\alpha} f_h = \frac{1}{\hat{\lambda}^{\alpha}} \prod_{k=0}^{L} \left[1 + \tau_k \mathcal{B}_h (\delta \mathcal{I} + t_k \mathcal{B}_h)^{-1}\right]^{-\alpha} f_h$$

$$U_{L+1} = \frac{1}{\hat{\lambda}^{\alpha}} \prod_{k=0}^{L} r_m \left(\tau_k \mathcal{B}_h (\delta \mathcal{I} + t_k \mathcal{B}_h)^{-1}\right) f_h$$

Theorem 3.3

Denote $\theta_n(\lambda) = \frac{\tau_n(\lambda-\delta)}{\delta + t_n(\lambda-\delta)}$ *for* $n = 0, 1, \cdots, L$ *and suppose the mesh* $\{t_l\}_{l=0}^{L+1}$ *satisfies* $\theta_n(\lambda_{h,max}) \leq 1$*. Then for positive integer* m *we have*

$$\|\mathcal{L}_{\mathcal{M}_h}^{-\alpha} f_h - U_{L+1}\|_{L^2(\mathcal{M}_h)} \leq \widetilde{c} \delta^{-\alpha} 2^{-5m} \|f_h\|_{L^2(\mathcal{M}_h)}, \quad \forall f_h \in \mathbb{L}^2(\mathcal{M}_h),$$

where $\widetilde{c} = 2^{\alpha+1} c'_{\alpha}$.

Proof. Let $v_n(\lambda) = (\hat{\lambda} + t_n(\lambda - \hat{\lambda}))^{-\alpha}$, $\mu_0(\lambda) = \hat{\lambda}^{-\alpha}$ and for $n = 1, 2, \ldots, L+1$,

$$\mu_n(\lambda) = r_m(\tau_{n-1}(\lambda - \hat{\lambda})(\hat{\lambda} + t_{n-1}(\lambda - \hat{\lambda}))^{-1})\mu_{n-1}(\lambda),$$

and use the fact

$$U_{L+1} = \sum_{j=1}^{D} \mu_{L+1}(\lambda_{j,h})(f_h, \psi_{j,h})_{\mathcal{M}_h} \psi_{j,h},$$

$$u_h = \sum_{j=1}^{D} v_{L+1}(\lambda_{j,h})(f_h, \psi_{j,h})_{\mathcal{M}_h} \psi_{j,h}.$$

We note that, $v_{n+1} = (1 + \theta_n)^{-\alpha} v_n$ with $\theta_n(\lambda) = \frac{\tau_n(\lambda - \hat{\lambda})}{\hat{\lambda} + t_n(\lambda - \hat{\lambda})}$ for $n = 0, \ldots, L$. Thus,

$$\mu_{n+1} - v_{n+1} := e_{n+1} = (1 + \theta_n)^{-\alpha} e_n + [r_m(\theta_n) - (1 + \theta_n)^{-\alpha}]\mu_n.$$

Then

$$|e_{n+1}| \leq (1 + \theta_n)^{-\alpha}|e_n| + c'_\alpha \hat{\lambda}^{-\alpha} \frac{2^{-4m} {\theta_n}^{2m+1}}{2^{m\theta_n}}.$$

Implementation

In fact, to implement the algorithm more efficiently, we rewrite the rational function $r_m(t)$ as a sum of partial fractions:

$$r_m(t) = \beta_0 + \sum_{i=1}^{m} \frac{\beta_i}{1 + t_i(-\alpha, \alpha)t} \tag{3.8}$$

where

$$\beta_0 = \prod_{i=1}^{m} \frac{t_i(\alpha, -\alpha)}{t_i(-\alpha, \alpha)} > 0 \quad \text{and} \quad \beta_i = \frac{\prod_{j=1}^{m}(1 - t_j(\alpha, -\alpha)/t_i(-\alpha, \alpha))}{\prod_{j \neq i}(1 - t_j(-\alpha, \alpha)/t_i(-\alpha, \alpha))} > 0.$$

To reach $t_{L+1} = 1$ with fewer steps, say, to get smaller L, the best choice would be $\theta_n(\lambda_{h,max}) = 1$ for $n = 0, 1, \cdots, L$. Thus we obtain $t_0 = 0$,

$$t_1 = \frac{\hat{\lambda}}{\lambda_{h,max} - \hat{\lambda}}, \;\; t_{n+1} = \min\{2^{n+1} t_1 - t_1, 1\}, \quad n = 1, 2, \cdots, L, \tag{3.9}$$

thus

$$L + 1 = \lceil \log_2(\lambda_{h,max}/\hat{\lambda}) \rceil. \tag{3.10}$$

Matrix form

Denote by $\mathbf{R}_l^i = \beta_i[\hat{\lambda}\mathbf{M}+(t_l+t_i(-\alpha,\alpha)\tau_l)(\mathbf{S}-\hat{\lambda}\mathbf{M})]^{-1}[\hat{\lambda}\mathbf{M}+t_l(\mathbf{S}-\hat{\lambda}\mathbf{M})]$

Algorithm 1

(a) Set $\vec{U}_0 = \hat{\lambda}^{-\alpha}\vec{F}$;

(b) For $l = 0, 1, \ldots, L$:

 (i) For $i = 1, \cdots, m$, *solve for* $\vec{U}_{l+1}^i$: $\vec{U}_{l+1}^i = \mathbf{R}_l^i \vec{U}_l$;

 (ii) Set $\vec{U}_{l+1} = \beta_0\vec{U}_l + \sum_{i=1}^m \beta_i\vec{U}_{l+1}^i$;

(c) end.

Content

Appealing to Balakrishnan formula we know

$$u = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \mu^{-\alpha} (\mu\mathcal{I} + \mathcal{L}_{\mathcal{M}})^{-1} f \, d\mu, \tag{4.1}$$

and

$$u_h = \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \mu^{-\alpha} (\mu\mathcal{I} + \mathcal{L}_{\mathcal{M}_h})^{-1} f_h \, d\mu. \tag{4.2}$$

So to compare the error between u and u_h, we introduce the auxiliary problem: for $f \in \dot{\mathbb{H}}^\delta(\mathcal{M})$, find $w_\mu \in \mathbb{H}^1(\mathcal{M})$ such that for $\forall v \in \mathbb{H}^1(\mathcal{M})$

$$a_\mu(\cdot,\cdot) : \ \langle \mathcal{L}_{\mathcal{M}} w_\mu + \mu w_\mu, v\rangle_{\mathcal{M}} = \langle f, v\rangle_{\mathcal{M}} := F(v). \tag{4.3}$$

Correspondingly we denote $w_{h,\mu} \in V_h$ as the solution of

$$a_{h,\mu}(\cdot,\cdot) : \ \mathcal{L}_{\mathcal{M}_h} w_{h,\mu} + \mu w_{h,\mu} = f_h. \tag{4.4}$$

Theorem 4.1

Suppose $a(\vec{x}) \in H^1(\mathcal{M}) \cap C^\kappa(\mathcal{M})$, $b(\vec{x}) \in C^\kappa(\mathcal{M})$, $f \in \dot{\mathbb{H}}^\delta(\mathcal{M})$ *and* f_h *satisfies Assumption 2.1, where* $\delta \in [0,2]$ *and* $\kappa = \min\{2\alpha+\delta, 2\}$. *Let* $u = \mathcal{L}_{\mathcal{M}}^{-\alpha} f$ *and* $u_h = \mathcal{L}_{\mathcal{M}_h}^{-\alpha} f_h$, *then for* h *small enough*

$$\|u - u_h^\ell\| \leq c \begin{cases} |2\alpha+\delta-2|^{-1} h^{\min(\delta+2\alpha,2)} \|f\|_{\dot{H}^\delta(\mathcal{M})} & \alpha+\delta/2 \neq 1, \\ h^2 |\ln h| \|f\|_{\dot{H}^\delta(\mathcal{M})} & \alpha+\delta/2 = 1, \end{cases}$$

with c *independent of* α *and* h.

Proof. Use the following integral then analyze term by term:

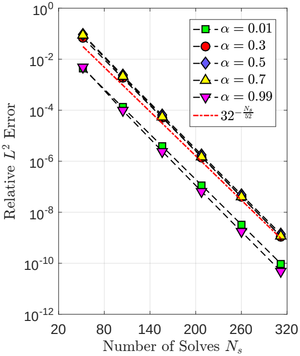
$$\|u - u_h^\ell\| \leq \frac{\sin(\pi\alpha)}{\pi} \int_0^\infty \mu^{-\alpha} \|e_\mu\| d\mu.$$

Content

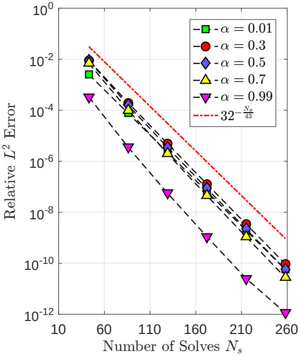
Example 1. To verify the robustness of our algorithm we take the checkerboard problem on domain $[-1,1]\times[-1,1]$ with homogeneous boundary for the first test, say, $\mathcal{M}$ is a plane square. We apply a standard finite difference scheme with more than 10^6 degree of freedom to discrete $\mathcal{L}_\mathcal{M}$ with $a(\vec{x})=1, b(\vec{x})=0$ and

$$f=\begin{cases} 1, & \text{for} \quad x_1x_2 \geq 0; \\ -1 & \text{for} \quad x_1x_2 < 0. \end{cases}$$

The mesh we use is geometrical refined around the boundary and along $x_1, x_2 = 0$: we first divide each direction into $N_1 = 2N_0 = 1000$ intervals $\{I_k\}_{k=1}^{N_1}$, then refine I_k with $k = 1, N_0, N_0+1, N_1$ by adding $p = 12$ nodes in each of them, denoting by $\{x_k^n\}_{n=1}^{p}$ which are exponentially clustered correspondingly at $-1^+, 0^-, 0^+$ and 1^- with a speed of 2^{-n}. One can obtain $\Lambda \approx 3.355\times 10^{13}$ and the total degree of freedom is $1047^2 = 1096209 > 10^6$ for such $\mathcal{L}_{\mathcal{M}_h}$. It is also worth to point out that in each direction the ratio of the mesh size is $h_{max}/h_{min} = 2^p = 4096$.

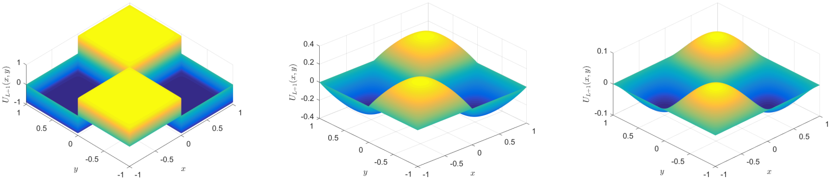


(a) $\hat{\lambda} = 0.01$



(b) $\hat{\lambda} = 4$

Figure 4: $\|u_h - U_{L+1}\|_h / \|u_h\|_h$ under different $\hat{\lambda}$



(a) $\alpha = 0.01$ (b) $\alpha = 0.5$ (c) $\alpha = 0.99$

Figure 5: Numerical solutions under different α

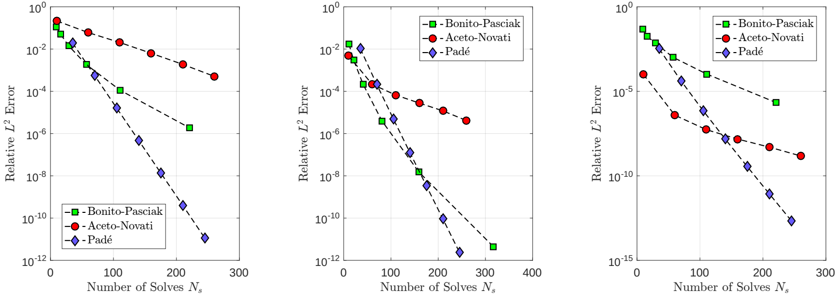
We would also like to compare our scheme with Bonito-Pasciak scheme and Aceto-Novati scheme. We employ **Example 1** with $N_0 = 25, p = 12$ to do the test. Running few iterations of power method gives $\Lambda = 7.5 \times 10^{10}$. We still set $\hat{\lambda} = 4$ in our scheme. The results are presented in Fig.6. One can observe that our scheme gives better results for N_s large enough. Recall that as a function of N_s,

$$\text{Bonito-Pasciak: } \mathcal{O}(e^{-\pi\sqrt{\alpha(1-\alpha)N_s}})$$

which indicates that the scheme will degenerate for α close to 1 or 0.

$$\text{Aceto-Novati: } C\sin(\alpha\pi)\lambda_{h,max}^{-\alpha/2}\exp\left(-4N_s\left(\frac{\lambda_{h,min}}{\lambda_{h,max}}\right)^{1/4}\right),$$

which implies slow convergence when the condition number of $\mathcal{L}_{\mathcal{M}_h}$ is large, especially in the case of $\alpha < 0.5$.



(a) $\alpha = 0.1$ (b) $\alpha = 0.5$ (c) $\alpha = 0.9$

Figure 6: Error plots for different schemes under different α

Example 2. To verify Theorem 4.1, we consider

$$\mathcal{M} = \left\{(x_1, x_2, x_3) | x_1^2 + x_2^2 + x_3^2 = 1\right\},$$

say, $\mathcal{M}$ is a spherical surface in $\mathbb{R}^3$ with radius $r = 1$. We take $\mathcal{L}_{\mathcal{M}} = -\Delta_{\mathcal{M}}$ and

$$f = \begin{cases} 1, & \text{for} \quad x_3 \geq 0; \\ -1 & \text{for} \quad x_3 < 0. \end{cases}$$

Table 1: $\|u^{-\ell} - u_h\|_{L^2(\mathcal{M}_h)}$ under different mesh

$\alpha \backslash Dof$	153	606	2418	9666
$\alpha = 0.01$	3.8192e-01	2.6089e-01	1.8143e-01	1.2642e-01
	(0.5)	0.55	0.53	0.52
$\alpha = 0.3$	8.1961e-02	3.9167e-02	1.8433e-02	8.6204e-03
	(1.1)	1.07	1.09	1.10
$\alpha = 0.5$	2.5687e-02	1.0153e-02	3.7672e-03	1.3619e-03
	(1.5)	1.35	1.43	1.47
$\alpha = 0.7$	1.1322e-02	3.3840e-03	9.5737e-04	2.6565e-04
	(1.9)	1.76	1.83	1.85
$\alpha = 0.99$	1.8734e-02	4.8750e-03	1.2329e-03	3.0923e-04
	(2.0)	1.96	1.99	2.00

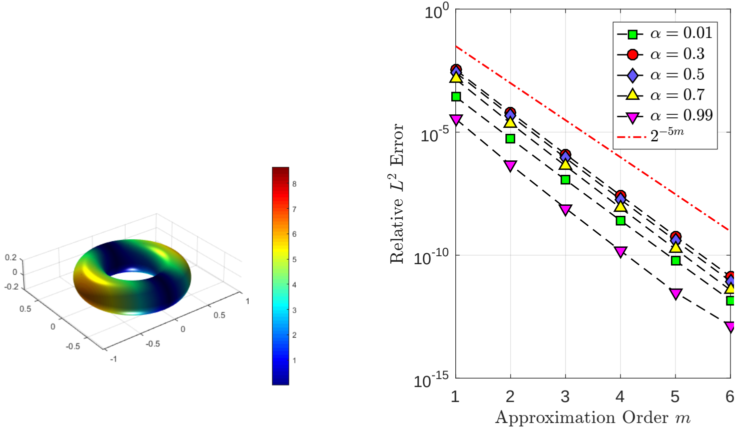
Example 3. In this example we take a torus as $\mathcal{M}$, which is given parametrically by

$$\vec{x} = [(R+r\cos\varphi_1)\cos\varphi_2, (R+r\cos\varphi_1)\sin\varphi_2, r\sin\varphi_1], \quad \varphi_1, \varphi_2 \in [0, 2\pi)$$

with $R = 0.5, r = 0.2$. We set $\mathcal{L}_{\mathcal{M}} = -\Delta_{\mathcal{M}} + \mathcal{I}$ and

$$f = H\cos\left(\arctan\left(\frac{x_2}{x_1}\right)\right)$$

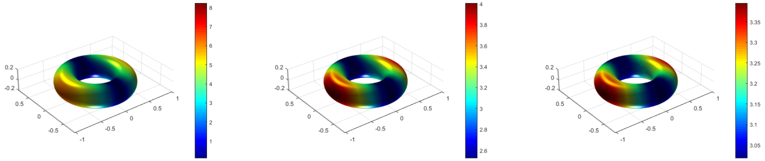
with H the mean curvature of the torus, see Fig. 7(a) for it. The surface is triangulated by 237,568 simplices with 118,784 vertexes.



(a) Source term f

(b) Relative L^2-errors against m

Figure 7: Source term and relative L^2-errors



(a) $\alpha = 0.01$ (b) $\alpha = 0.5$ (c) $\alpha = 0.99$

Figure 8: Numerical solutions for $\alpha = 0.01, 0.5$ and 0.99, respectively

In this case

$$\hat{\lambda} = \lambda_{min} = 1, \quad \lambda_{h,max} \leq \Lambda = 1.78 \times 10^6, \quad L + 1 = 21.$$

References

- B. Duan. Padé-parametric FEM approximation for fractional powers of elliptic operators on manifolds, IMA J. Numerical Analysis, 2022

Thank you for your attention!