Optimization with respect to order in a fractional diffusion model

Abner J. Salgado

Department of Mathematics
University of Tennessee

Nonlocal School on Fractional Equations
August 19, 2017

Joint work with Enrique Otárola (UTFSM) and Harbir Antil (GMU).

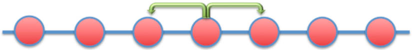
Supported by NSF grant DMS-1418784

Outline

Outline

Local jump random walk

- We consider a random walk of a particle along the real line.
- $h\mathbb{Z} = \{hz : z \in \mathbb{Z}\}$ — possible states of the particle.
- $u(x,t)$ — probability of the particle to be at $x \in h\mathbb{Z}$ at time $t \in \tau\mathbb{N}$.
- Local jump random walk: at each time step of size τ, the particle jumps to the left or right with probability $1/2$.



$$u(x, t+\tau) = \frac{1}{2}u(x+h, t) + \frac{1}{2}u(x-h, t)$$

If we consider $\tau = 2h^2$, then we obtain

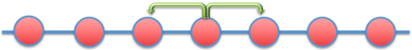
$$\frac{u(x,t+\tau) - u(x,t)}{\tau} = \frac{u(x+h,t) + u(x-h,t) - 2u(x,t)}{h^2}$$

Letting $h, \tau \downarrow 0$ yields

$$u_t - \Delta u = 0$$

Local jump random walk

- We consider a random walk of a particle along the real line.
- $h\mathbb{Z} = \{hz : z \in \mathbb{Z}\}$ — possible states of the particle.
- $u(x,t)$ — probability of the particle to be at $x \in h\mathbb{Z}$ at time $t \in \tau\mathbb{N}$.
- Local jump random walk: at each time step of size τ, the particle jumps to the left or right with probability $1/2$.



$$u(x,t+\tau) = \frac{1}{2}u(x+h,t) + \frac{1}{2}u(x-h,t)$$

If we consider $\tau = 2h^2$, then we obtain

$$\frac{u(x,t+\tau) - u(x,t)}{\tau} = \frac{u(x+h,t) + u(x-h,t) - 2u(x,t)}{h^2}$$

Letting $h, \tau \downarrow 0$ yields

$$u_t - \Delta u = 0$$

Long jump random walk

- The probability that the particle jumps from the point $hk \in h\mathbb{Z}$ to the point $hm \in h\mathbb{Z}$ is $\mathcal{K}(k-m) = \mathcal{K}(m-k)$.

$$u(x, t+\tau) = \sum_{k\in\mathbb{Z}} \mathcal{K}(k) u(x+hk, t),$$

since $\sum_{k\in\mathbb{Z}} \mathcal{K}(k) = 1$, this yields

$$u(x, t+\tau) - u(x,t) = \sum_{k\in\mathbb{Z}} \mathcal{K}(k) \left(u(x+hk, t) - u(x,t)\right)$$

- Let $\mathcal{K}(y) \sim |y|^{-(1+2s)}$ with $s \in (0,1)$.
- Choose $\tau = h^{2s}$, then $\frac{\mathcal{K}(k)}{\tau} = h\mathcal{K}(kh)$.

Let $h, \tau \downarrow 0$,

$$\partial_t u = \int_{\mathbb{R}} \frac{u(x+y,t) - u(x,t)}{|y|^{1+2s}} \, \mathrm{d}y \Leftrightarrow \partial_t u = -(-\Delta)^s u$$

Long jump random walk

Question: What were the fundamental ingredients that led to a fractional heat equation?

- $\mathcal{K}(y) \sim |y|^{-(1+2s)}$ with $s \in (0,1)$, but the construction would have worked with another kernel, thus obtaining another nonlocal operator.
- $\tau = h^{2s}$. The space and time must have a particular scaling!

Question: How do we find this scaling? How do we know the order of the fractional diffusion?

Long jump random walk

Question: What were the fundamental ingredients that led to a fractional heat equation?

- $\mathcal{K}(y) \sim |y|^{-(1+2s)}$ with $s \in (0,1)$, but the construction would have worked with another kernel, thus obtaining another nonlocal operator.
- $\tau = h^{2s}$. The space and time must have a particular scaling!

Question: How do we find this scaling? How do we know the order of the fractional diffusion?

Long jump random walk

Question: What were the fundamental ingredients that led to a fractional heat equation?

- $\mathcal{K}(y) \sim |y|^{-(1+2s)}$ with $s \in (0,1)$, but the construction would have worked with another kernel, thus obtaining another nonlocal operator.
- $\tau = h^{2s}$. The space and time must have a particular scaling!

Question: How do we find this scaling? How do we know the order of the fractional diffusion?

Outline

The model problem

- The data:
  - $\Omega \subset \mathbb{R}^n$, open, convex and with Lipschitz boundary.
  - $\mathsf{f}, \mathsf{u}_d : \Omega \to \mathbb{R}$, "nice" enough.
- The problem: Find $(\bar{s}, \bar{\mathsf{u}})$ that minimize

$$J(s,u) = \frac{1}{2}\|u - \mathsf{u}_d\|^2_{L^2(\Omega)} + \varphi(s)$$

  subject to $(-\Delta)^s u = \mathsf{f}$.
- Where:
  - For $0 \leq \alpha < \beta \leq 1$, $\varphi \in C^2(\alpha, \beta)$ is nonnegative, convex and

$$\lim_{s \downarrow \alpha} \varphi(s) = \lim_{s \uparrow \beta} \varphi(s) = +\infty.$$

    For instance,

$$\varphi(s) = (s-\alpha)^{-1}(\beta - s)^{-1}, \qquad \varphi(s) = (s-\alpha)^{-1} e^{(\beta - s)^{-1}}.$$

  - $(-\Delta)^s$ denotes the fractional powers of the Dirichlet Laplacian.

The model problem

- The problem: Find $(\bar{s}, \bar{\mathsf{u}})$ that minimize

$$J(s,u) = \frac{1}{2}\|u - \mathsf{u}_d\|^2_{L^2(\Omega)} + \varphi(s)$$

  subject to

$$(-\Delta)^s u = \mathsf{f}.$$

Question: What are we trying to model here?

- Given some "observations/measurements" u_d, can we find the order of fractional diffusion s that best represents them?

Comment: This problem was originally considered by (Sprekels, Valdinoci 2016) for the fractional heat operator $\partial_t + (-\Delta)^s$, the authors show existence of solutions and optimality conditions.

Outline

Spectral theory 101

We consider the definition of $(-\Delta)^s$ based on spectral theory:

- $-\Delta : H^2(\Omega) \cap H_0^1(\Omega) \subset L^2(\Omega) \to L^2(\Omega)$ is symmetric, closed and unbounded and its inverse is compact.
- The eigenpairs $\{\lambda_k, \varphi_k\}$, i.e.

  $$-\Delta\varphi_k = \lambda_k\varphi_k, \qquad \varphi_k|_{\partial\Omega} = 0$$

  form an orthonormal basis of $L^2(\Omega)$.
- For u sufficiently smooth:

  $$u = \sum_{k=1}^{\infty} u_k\varphi_k \longmapsto (-\Delta)^s u := \sum_{k=1}^{\infty} u_k\lambda_k^s\varphi_k$$

- $(-\Delta)^s : \mathbb{H}^s(\Omega) \to \mathbb{H}^{-s}(\Omega)$, $\mathbb{H}^s(\Omega) = [H_0^1(\Omega), L^2(\Omega)]_{1-s}$.

The control to state map

- For $\mathsf{f} = \sum_k \mathsf{f}_k \varphi_k \in \mathbb{H}^{-s}(\Omega)$ the solution to the state equation is
$$\mathsf{u} = \sum_k \lambda_k^{-s} \mathsf{f}_k \varphi_k.$$
- This defines: $(0,1) \ni s \mapsto \mathcal{S}(s) = \sum_k \lambda_k^{-s} \mathsf{f}_k \varphi_k \in L^2(\Omega)$.

Theorem (properties of $\mathcal{S}$)

For $\mathsf{f} \in L^2(\Omega)$ the control to state map $\mathcal{S}$ is bounded
$\|\mathcal{S}(s)\|_{L^2(\Omega)} \lesssim 1$,
and three times Fréchet differentiable:

$$\mathcal{S}'(s) = -\sum_k \lambda_k^{-s} \ln(\lambda_k) \mathsf{f}_k \varphi_k =: \mathsf{u}'(s)$$
$$\mathcal{S}''(s) = \sum_k \lambda_k^{-s} \ln^2(\lambda_k) \mathsf{f}_k \varphi_k =: \mathsf{u}''(s)$$

with
$$\|\mathcal{S}^{(k)}(s)\|_{\mathbb{R} \to L^2(\Omega)} \lesssim s^{-k}, \ k = 1, 2, 3.$$

Existence

Since the state equation always has a solution, we introduce the reduced cost

$$f(s) = J(s, \mathcal{S}(s)).$$

Theorem (existence)

There is an optimal pair $(\bar{s}, \bar{\mathrm{u}} = \mathcal{S}(\bar{s})) \in (\alpha, \beta) \times \mathbb{H}^{\bar{s}}(\Omega)$ *for which*

$$f(\bar{s}) \leq f(s), \quad \forall s \in (\alpha, \beta).$$

Proof.

- The function f in continuous on (α, β).
- Consider sequences $\alpha_k \downarrow \alpha$, $\beta_k \uparrow \beta$ and seek for

$$s_k = \underset{s \in [\alpha_k, \beta_k]}{\operatorname{argmin}} f(s).$$

- Any accumulation point of $\{s_k\}_{k \geq 1}$ is a minimizer.

Optimality conditions

Theorem (optimality conditions)

- *First order necessary condition: If $(\bar{s}, \bar{\mathsf{u}}(\bar{s}))$ is optimal, then*

$$(\bar{\mathsf{u}}(\bar{s}) - \mathsf{u}_d, \bar{\mathsf{u}}'(\bar{s}))_{L^2(\Omega)} + \varphi'(\bar{s}) = 0 \qquad (f'(\bar{s}) = 0).$$

- *second order sufficient condition: If $(\bar{s}, \bar{\mathsf{u}}(\bar{s}))$ satisfies the first order condition and, in addition,*

$$\|\bar{\mathsf{u}}'(\bar{s})\|^2_{L^2(\Omega)} + (\bar{\mathsf{u}}(\bar{s}) - \mathsf{u}_d, \bar{\mathsf{u}}''(\bar{s}))_{L^2(\Omega)} + \varphi''(\bar{s}) > 0, \quad (f''(\bar{s}) > 0)$$

  then the pair is optimal.

- In essence, we are dealing with the unconstrained minimization of a twice differentiable function over an open set.

What about (local) uniqueness?

Assume that φ is strongly convex, i.e., there is $\xi > 0$

$$(\varphi'(s_1) - \varphi'(s_2)) \cdot (s_1 - s_2) \geq \xi |s_1 - s_2|^2, \forall s_1, s_2 \in (\alpha, \beta)$$

then we have:

Theorem (local uniqueness)

Assume that φ is strongly convex, $\|\mathsf{f}\|_{L^2(\Omega)}$ and $\|\mathsf{u}_d\|_{L^2(\Omega)}$ are small enough. If $\bar{s}$ is optimal, there is $\delta > 0$ and $\eta > 0$ such that

$$f(s) \geq f(\bar{s}) + \eta |s - \bar{s}|^2, \quad \forall s \in (\alpha, \beta) \cap (\bar{s} - \delta, \bar{s} + \delta).$$

- This implies local uniqueness.

Outline

Disclaimer

Up to now we could have

$$\alpha = 0 \qquad \beta = 1$$

from now on we require

$$\alpha > 0 \qquad \beta < 1$$

Discretization in s

- For $\sigma > 0$ introduce the centered difference operator

$$d_\sigma \psi(s) = \frac{1}{2\sigma} \left(\psi(s+\sigma) - \psi(s-\sigma) \right).$$

- Recall that, for $\psi \in C^3$

$$|\psi'(s) - d_\sigma \psi(s)| = \mathcal{O}(\sigma^2).$$

- We will discretize the first order optimality condition and seek for $s_\sigma \in (\alpha, \beta)$ such that

$$j_\sigma(s_\sigma) = (\mathsf{u}(s_\sigma) - \mathsf{u}_d, d_\sigma \mathsf{u}(s_\sigma))_{L^2(\Omega)} + \varphi'(s_\sigma) = 0.$$

How do we find s_σ?

```
0 < σ ≪ 1 and set s_l, s_r ∈ (α, β), with s_l < s_r;                ▷ Initialization
                                         ▷ We take care of possible degenerate cases
if j_σ(s_l) = 0 then
    s_σ = s_l;
end if
if j_σ(s_r) = 0 then
    s_σ = s_r;
end if
                                                                   ▷ Root isolation
while j_σ(s_r) < 0 do
    s_r := s_r + σ;
end while
while j_σ(s_l) > 0 do
    s_l := s_l − σ;
end while
                                                                        ▷ Bisection
k = 1;
repeat
    s_k = (1/2)(s_l + s_r);
    if j_σ(s_k) = 0 then
        s_σ = s_k;
        break;                                       ▷ The solution has been found
    end if
    if j_σ(s_l) j_σ(s_k) > 0 then                                      ▷ Sign check
        s_l = s_k;
    else
        s_r = s_k;
    end if
    k = k + 1;
until forever
```

Bisection method

Does the root isolation step finish?

Lemma (root isolation)

If σ is sufficiently small there are $s_l, s_r \in (\alpha, \beta)$ for which

$$j_\sigma(s) < 0 \; s \in (\alpha, s_l), \qquad j_\sigma(s) > 0 \; s \in (s_r, \beta).$$

A standard argument then yields

Lemma (convergence of bisection)

The bisection method generates a sequence $\{s_k\}_{k\geq 1}$ that satisfies

$$|s_\sigma - s_k| \lesssim 2^{-k}.$$

With $j_\sigma(s_\sigma) = 0$.

What about convergence of s_σ?

- The optimal $\bar{s}$ need not be unique. Thus, we do not expect convergence of the whole family s_σ to s.
- The following statement is the best we can hope for.

Lemma (convergence of s_σ)

The family $\{s_\sigma\}_{\sigma>0}$ has a convergent subsequence and any accumulation point satisfies the first order condition.

- If we focus on one of these subsequences we can establish a rate.

Theorem (rate in σ)

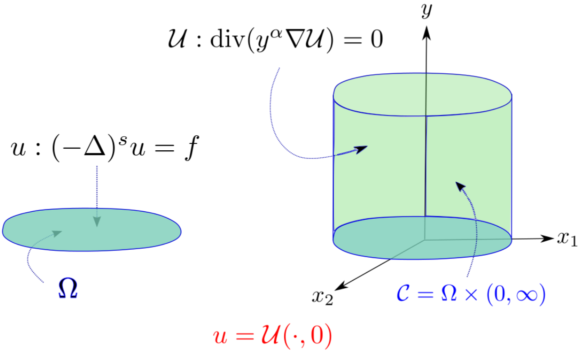
If σ is sufficiently small we have

$$|\bar{s} - s_\sigma| \lesssim \sigma^2.$$

Outline

The α-harmonic extension

Molčanov, Ostrovskiĭ (1969), Caffarelli, Silvestre (2007), Cabré, Tan (2010), Capella et al. (2011), Stinga Torrea (2010–2012).



- $s \in (0,1)$ and $\alpha = 1 - 2s \in (-1,1)$.
- $\partial_{\nu^\alpha} \mathcal{U} = -\lim_{y \downarrow 0} y^\alpha \partial_y \mathcal{U}$ on $\Omega \times \{0\}$.
- $d_s = 2^\alpha \Gamma(1-s)/\Gamma(s)$.

The α-harmonic extension

- Recall that $\alpha = 1 - 2s \in (-1, 1)$, y^α is degenerate $(\alpha > 0)$ or singular $(\alpha < 0)$!
- But y^α is a Muckenhoupt weight.
- The domain $\mathcal{C} = \Omega \times (0, \infty)$ is infinite!
- We can consider a truncated version and incur in an exponentially small error:

$$\|\mathcal{U} - \mathcal{V}\|_{\mathring{H}^1_L(y^\alpha, \mathcal{C}_\Upsilon)} \lesssim e^{-\sqrt{\lambda_1}\Upsilon/4}.$$

- The solution has a rather bad behavior $\mathcal{U}_{yy} \approx y^{-\alpha-1}$ as $y \approx 0^+$.
- We use anisotropic meshes.

Discretization

- Denote: $\mathscr{T}_\mathcal{Y}$ the mesh and $\mathbb{V}(\mathscr{T}_\mathcal{Y})$ the discrete space. Then

$$\|\mathcal{V} - V_{\mathscr{T}_\mathcal{Y}}\|_{\mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathcal{Y})} = \inf_{W \in \mathbb{V}(\mathscr{T}_\mathcal{Y})} \|\mathcal{V} - W\|_{\mathring{H}^1_L(y^\alpha, \mathcal{C}_\mathcal{Y})},$$

  and set $W = \Pi\mathcal{V} \in \mathbb{V}(\mathscr{T}_\mathcal{Y})$. We need to construct a suitable interpolation operator.

  N. O. S. *Piecewise polynomial interpolation in Muckenhoupt weighted Sobolev spaces*. Numer. Math 2016.

- If the mesh is suitably graded:

$$\begin{aligned}\|u - V_{\mathscr{T}_\mathcal{Y}}(\cdot, 0))\|_{\mathbb{H}^s(\Omega)} &\le \|\nabla(\mathcal{U} - V_{\mathscr{T}_\mathcal{Y}})\|_{L^2(y^\alpha, \mathcal{C})} \\ &\lesssim |\log \#\mathscr{T}_\mathcal{Y}|^s \#\mathscr{T}_\mathcal{Y}^{-\frac{1}{n+1}}.\end{aligned}$$

  which is near optimal estimate in terms of degrees of freedom.

  N. O. S. *A PDE approach to fractional diffusion*. Found. Comp. Math. 2015.

- This formulation allows us to devise multigrid methods

  L. Chen, N. O. S. *Multilevel methods for nonuniformly elliptic equations*. Math. Comp. 2016.

The discrete control to state map

- All the hidden constants in the previous discussion depend on s, but since $s \in (\alpha, \beta) \Subset (0,1)$ they are uniformly controlled.
- Define $S_{\mathscr{T}} : (\alpha, \beta) \to \mathbb{U}(\mathscr{T}_\Omega)$ by $s \mapsto U_{\mathscr{T}_\Omega} = V_{\mathscr{T}_\mathcal{Y}}(\cdot, 0)$

Lemma (continuity of $S_{\mathscr{T}}$)

For every $\mathscr{T}_\mathcal{Y}$ the map $S_{\mathscr{T}}$ is continuous on (α, β).

- All norms in finite dimensions are equivalent.

Fully discrete scheme

Define

$$j_{\sigma,\mathscr{T}}(s) = (U_{\mathscr{T}_\Omega}(s) - \mathsf{u}_d, d_\sigma U_{\mathscr{T}_\Omega}(s))_{L^2(\Omega)} + \varphi'(s).$$

We seek for $s_{\sigma,\mathscr{T}}$ such that

$$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}}) = 0.$$

- The continuity of $S_{\mathscr{T}}$ implies that we can find it by using bisection as before.

Error estimates

- As before, we can only expect that a subsequence of $\{s_{\sigma,\mathscr{T}}\}_{\mathscr{T}}$ converges to a s_σ.
- If we extract this subsequence then we have.

Theorem (rate of convergence)

If $\mathsf{f} \in \mathbb{H}^{1-\epsilon}(\Omega)$ *for all* $\epsilon > 0$ *we have*

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim \sigma^{-1} |\log(\#\mathscr{T}_{\mathcal{Y}})|^2 \, (\#\mathscr{T}_{\mathcal{Y}})^{-1/(n+1)} + \sigma^2.$$

Corollary (explicit rate)

Choose $\sigma \approx |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2/3} \, (\#\mathscr{T}_{\mathcal{Y}})^{-\frac{1}{3(n+1)}}$ *then*

$$|\bar{s} - s_{\sigma,\mathscr{T}}| \lesssim |\log(\#\mathscr{T}_{\mathcal{Y}})|^{2/3} \, (\#\mathscr{T}_{\mathcal{Y}})^{-\frac{2}{3(n+1)}}$$

Outline

Generalities

- $\Omega = (0,1)^2$
- $\mathscr{Y} = 1 + \frac{1}{3}(\#\mathscr{T}_\Omega)$
- $\sigma = \frac{1}{2.5}\left(\#\mathscr{T}_{\mathscr{Y}}\right)^{-1/9}$
- The initial bounds are $s_l = 0.3$ and $s_r = 0.9$

In this geometry we have

$$\lambda_{k,l} = \pi^2(k^2 + l^2), \qquad \varphi_{k,l}(x,y) = \sin(k\pi x)\sin(l\pi y).$$

So if, for $s \in (0,1)$

$$\mathsf{f} = \lambda_{2,2}^s \varphi_{2,2} \implies \mathsf{u} = \varphi_{2,2}$$

Example 1: $\bar{s} = 1/2$

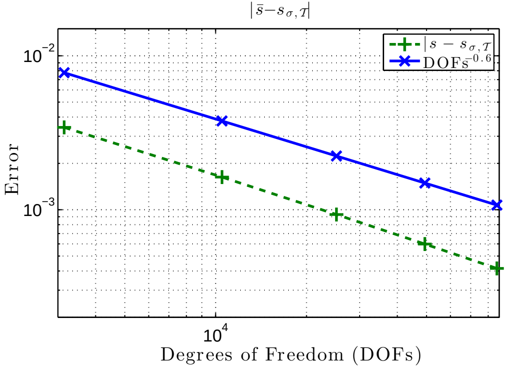
Set

$$\varphi(s) = \frac{1}{s(1-s)}$$

The following table shows the computed value of $s_{\sigma,\mathscr{T}}$ and the number of bisection iterations.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	4.96572e-01	-8.89011e-14	53
10496	4.98371e-01	-8.38218e-14	53
25137	4.99069e-01	3.49235e-14	53
49348	4.99402e-01	1.52327e-12	53
85529	4.99585e-01	6.28221e-12	53

Example 1: $\bar{s} = 1/2$. Convergence rate



- The rate of convergence is $\mathcal{O}(\#\mathscr{T}_{\mathcal{Y}}^{-0.6})$ which is better than the predicted rate of -0.22!

Example 2: $\bar{s} = (3 - \sqrt{5})/2$

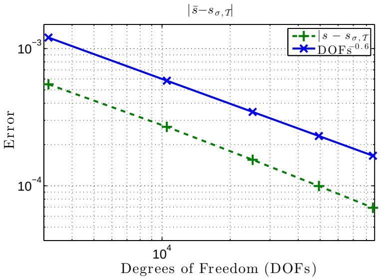
Set

$$\varphi(s) = \frac{1}{s} e^{\frac{1}{(1-s)}}$$

The following table shows the computed value of $s_{\sigma,\mathscr{T}}$ and the number of bisection iterations.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	3.81417e-01	9.99201e-16	46
10496	3.81697e-01	-2.52812e-13	53
25137	3.81811e-01	1.36418e-12	53
49348	3.81866e-01	2.66251e-12	53
85529	3.81897e-01	3.53083e-12	53

Example 2: $\bar{s} = (3 - \sqrt{5})/2$. Convergence rate



- The rate of convergence is $\mathcal{O}(\#\mathscr{T}_{\mathcal{Y}}^{-0.6})$ which is better than the predicted rate of -0.22!

Example 3. Unknown solution

$$
\begin{aligned}
\varphi(s) &= \frac{1}{s} e^{\frac{1}{(1-s)}}, \\
\mathsf{u}_d &= \max\left\{0, \frac{1}{2} - \sqrt{|x - \frac{1}{2}|^2 + |y - \frac{1}{2}|^2}\right\}, \\
\mathsf{f} &= 10 \notin \mathbb{H}^{\mu}(\Omega),\ \mu \geq \frac{1}{2}
\end{aligned}
$$

We do not know the solution, but we can still compute.
The following table shows the computed value of $s_{\sigma,\mathscr{T}}$ and the number of bisection iterations.

$\#\mathscr{T}_{\mathcal{Y}}$	$s_{\sigma,\mathscr{T}}$	$j_{\sigma,\mathscr{T}}(s_{\sigma,\mathscr{T}})$	N
3146	4.44005e-01	4.22951e-12	53
10496	4.47239e-01	2.97451e-11	53
25137	4.48182e-01	-3.20792e-11	53
49348	4.48544e-01	4.83542e-11	53
85529	4.48690e-01	2.68390e-10	53

Outline

Recap

- Parameter identification problem: The parameter is the order of the fractional elliptic operator.
- Existence and optimality conditions: Local uniqueness under smallness assumptions.
- Semidiscrete scheme: Convergence up to subsequences. Rate of convergence for subsequences.
- Fully discrete scheme: ídem.

Open questions

- Can we let $\alpha = 0$ and $\beta = 1$? The numerics seem to indicate that this is not an issue.
- Modulo technicalities we can also handle the time dependent problem, where the state equation is $\partial_t u + (-\Delta)^s u = \mathsf{f}$.
- Completely open: Space time fractional $\partial_t^\gamma u + (-\Delta)^s u = \mathsf{f}$ and optimize in s and γ.
- Ongoing: Consider the integral version of $(-\Delta)^s$ (with M. D'Elia).