Fast numerical methods and mathematical analysis of fractional PDEs

Hong Wang

Department of Mathematics, University of South Carolina

hwang@math.sc.edu

A Workshop on Future Directions in Fractional Calculus Research and Applications
Department of Statistics and Probability, Michigan State University
October 17-21, 2016

- Collaborators:
  T. Basu, Occidental College; H. Chen, L. Jia, W. Li, Shandong Normal U.; A. Cheng, N. Du, J. Gao, K. Wang, & X. Zhang, Shandong U.; H. Fu, China U. of Petroleum; X. Guo, C. Wang, Z. Wang & S. Yang, U. of South Carolina; J. Jia, Fudan U.; Z. Li, J. Pan, D. Yang & S. Zhu, East China Normal U.; M. Ng, Hong Kong Baptist U.; Y. Ren, Qilu U. of Technology; H. Tian, Ocean U. of China;
- Grant support
  The OSD/ARO MURI Grant W911NF-15-1-0562, and the National Science Foundation under Grants EAR-0934747, DMS-1216923 and DMS-1620194.

$$\partial_t u - d_+(x,t)\,{}_a^G D_x^\alpha u - d_-(x,t)\,{}_x^G D_b^\alpha u = f, \quad x \in (a,b),\ t \in (0,T], \qquad (1)$$
$$u(a,t) = u(b,t) = 0,\ t \in [0,T], \quad u(x,0) = u_0(x),\ x \in [a,b].$$

- d_+ and d_- are the left and right *variable* diffusivity coefficients (so analytical techniques do not apply, in general).
- The left- and right-sided Grünwald-Letnikov fractional derivatives of order $1 < \alpha < 2$ are defined by

$${}_a^G D_x^\alpha u(x,t) := \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon^\alpha} \sum_{k=0}^{\lfloor (x-a)/\varepsilon \rfloor} g_k^{(\alpha)} u(x - k\varepsilon, t),$$
$${}_x^G D_b^\alpha u(x,t) := \lim_{\varepsilon \to 0^+} \frac{1}{\varepsilon^\alpha} \sum_{k=0}^{\lfloor (b-x)/\varepsilon \rfloor} g_k^{(\alpha)} u(x + k\varepsilon, t) \qquad (2)$$

- $g_k^{(\alpha)} := (-1)^k \binom{\alpha}{k}$ with $\binom{\alpha}{k}$ being the fractional binomial coefficients.

A finite difference method (Lynch et al 2003, del-Castillo-Negrete et al 2004, Liu et al 2004, Meerschaert & Tadjeran 2004)

- FPDEs have significantly different features from integer-order PDEs.
- Let $x_i := a + ih$ and $t_m := m\Delta t$. The fully implicit finite difference scheme obtained by truncating (2) is *unconditionally unstable*!
- An *unconditionally stable* scheme is (Meerschaert & Tadjeran 2004) is

$$\frac{u_i^m - u_i^{m-1}}{\Delta t} - \frac{d_i^{+,m}}{h^\alpha} \sum_{k=0}^{i} g_k^{(\alpha)} u_{i-k+1}^m - \frac{d_i^{-,m}}{h^\alpha} \sum_{k=0}^{N-i+1} g_k^{(\alpha)} u_{i+k-1}^m = f_i^m \tag{3}$$

- The stiffness matrix $A^m = [a_{i,j}^m]_{i,j=1}^N$

$$a_{i,j}^m = \frac{1}{h^\alpha} \begin{cases} -\big(d_i^{+,m} + d_i^{-,m}\big) g_1^{(\alpha)} > 0, & j = i, \\ -\big(d_i^{+,m} g_2^{(\alpha)} + d_i^{-,m} g_0^{(\alpha)}\big) < 0, & j = i-1, \\ -\big(d_i^{+,m} g_0^{(\alpha)} + d_i^{-,m} g_2^{(\alpha)}\big) < 0, & j = i+1, \\ -d_i^{+,m} g_{i-j+1}^{(\alpha)} < 0, & j < i-1, \\ -d_i^{-,m} g_{j-i+1}^{(\alpha)} < 0, & j > i+1. \end{cases} \tag{4}$$

- In the matrix form of the finite difference scheme (3)

$$(I + \Delta t A^m) u^m = u^{m-1} + \Delta t f^m, \tag{5}$$

  - A^m is full and has to be assembled in any traditional scheme.
- We utilize the following properties of $g_k^{(\alpha)} := (-1)^k \binom{\alpha}{k}$ to conclude

$$\begin{aligned} & g_1^{(\alpha)} = -\alpha < 0, \quad 1 = g_0^{(\alpha)} > g_2^{(\alpha)} > g_3^{(\alpha)} > \cdots > 0, \\ & \sum_{k=0}^{\infty} g_k^{(\alpha)} = 0, \quad \sum_{k=0}^{m} g_k^{(\alpha)} < 0 \quad (m \geq 1), \\ & g_k^{(\alpha)} = \frac{\Gamma(k-\alpha)}{\Gamma(-\alpha)\Gamma(k+1)} = \frac{1}{\Gamma(-\alpha) k^{\alpha+1}} \Big(1 + O\Big(\frac{1}{k}\Big)\Big) \end{aligned} \tag{6}$$

  - $a_{i,i\pm k}/a_{i,i}$ decay at a rate of $1/k^{\alpha+1}$ as $k \to \infty$.

$$\begin{aligned} & a_{i,i}^m - \sum_{j=1, j\neq i}^{N} |a_{i,j}^m| \\ & = -(r_i^{+,m} + r_i^{-,m}) g_1^{(\alpha)} - r_i^{+,m} \sum_{k=0,k\neq 1}^{i} g_k^{(\alpha)} - r_i^{-,m} \sum_{k=0,k\neq 1}^{N-i} g_k^{(\alpha)} \\ & > -(r_i^{+,m} + r_i^{-,m}) g_1^{(\alpha)} - (r_i^{+,m} + r_i^{-,m}) \sum_{k=0,k\neq 1}^{\infty} g_k^{(\alpha)} = 0. \end{aligned} \tag{7}$$

  - A^m is a strictly diagonally dominant M-matrix, the scheme is monotone

Theorem

$$A^m = \left(\operatorname{diag}(d_i^{+,m})_{i=1}^N T^{\alpha,N} + \operatorname{diag}(d_i^{-,m})_{i=1}^N (T^{\alpha,N})^T\right)/h^\alpha, \tag{8}$$

$$T^{\alpha,N} := -\begin{bmatrix} g_1^{(\alpha)} & g_0^{(\alpha)} & 0 & \dots & 0 & 0 \\ g_2^{(\alpha)} & g_1^{(\alpha)} & g_0^{(\alpha)} & \ddots & \ddots & 0 \\ \vdots & g_2^{(\alpha)} & g_1^{(\alpha)} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ g_{N-1}^{(\alpha)} & \ddots & \ddots & \ddots & g_1^{(\alpha)} & g_0^{(\alpha)} \\ g_N^{(\alpha)} & g_{N-1}^{(\alpha)} & \dots & \dots & g_2^{(\alpha)} & g_1^{(\alpha)} \end{bmatrix}.$$

Theorem

$A^m v$ can be evaluated in $O(N \log N)$ operations for any vector v.

The matrix $T^{\alpha,N}$ is embedded into a $2N \times 2N$ circulant matrix $C^{\alpha,2N}$

$$C^{\alpha,2N} := \begin{bmatrix} T^{\alpha,N} & S^{\alpha,N} \\ S^{\alpha,N} & T^{\alpha,N} \end{bmatrix}, \quad S^{\alpha,N} := \begin{bmatrix} 0 & g_N^{(\alpha)} & \dots & \dots & g_3^{(\alpha)} & g_2^{(\alpha)} \\ 0 & 0 & g_N^{(\alpha)} & \dots & \ddots & g_3^{(\alpha)} \\ 0 & 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \ddots & 0 & g_N^{(\alpha)} \\ g_0^{(\alpha)} & 0 & \dots & 0 & 0 & 0 \end{bmatrix}.$$

- Let $c^{\alpha,2N}$ be the first column of $C^{\alpha,2N}$. Then $C^{\alpha,2N}$ can be decomposed as

$$C^{\alpha,2N} = F_{2N}^{-1} \, \mathrm{diag}(F_{2N} c^{\alpha,2N}) \, F_{2N} \tag{9}$$

- A fast matrix-vector multiplication $A^m \mathsf{v}$ is formulated as follows
  - For any $\mathsf{v} \in \mathbb{R}^N$, define v_{2N} by

$$\mathsf{v}_{2N} = \begin{bmatrix} \mathsf{v} \\ 0 \end{bmatrix}, \quad C^{\alpha,2N}\mathsf{v}_{2N} = \begin{bmatrix} T^{\alpha,N} & S^{\alpha,N} \\ S^{\alpha,N} & T^{\alpha,N} \end{bmatrix} \begin{bmatrix} \mathsf{v} \\ 0 \end{bmatrix} = \begin{bmatrix} T^{\alpha,N}\mathsf{v} \\ S^{\alpha,N}\mathsf{v} \end{bmatrix}. \tag{10}$$

  - $F_{2N}\mathsf{v}_{2N}$ can be carried out in $O(N \log N)$ operations via FFT, so $C^{\alpha,2N}\mathsf{v}_{2N}$ can be evaluated in $O(N \log N)$ operations.
  - The first N entries of $C^{\alpha,2N}\mathsf{v}_{2N}$ yields $T^{\alpha,N}\mathsf{v}$.
  - Similarly, $(T^{\alpha,N})^T\mathsf{v}$ can be evaluated in $O(N \log N)$ operations.
  - $A^m\mathsf{v}$ can be evaluated in $O(N \log N)$ operations.

- The fast algorithm
  - is not lossy, since no compression used in evaluating $A^m v$;
  - retains the conservation, stability, and convergence of the underlying scheme;
  - is nonintrusive, only the matrix-vector multiplication module needs to be modified.

- By (8)–(10), the fast algorithm is *matrix-free*.
  - The evaluatation of $A^m v$ requires only formulating the vectors $\{d_i^{\pm,m}\}_{i=1}^N$ and $c^{\alpha,2N}$;
  - The storage of A^m requires only storing the $(3N+1)$ parameters $\{d_i^{\pm,m}\}_{i=1}^N$ and $\{g_i^{(\alpha)})_{i=0}^N$.
  - In contrast, any traditional method requires the assembly of the full stiffness matrix A^m.

$$\begin{aligned}
&\partial_t u(x,y,t) - d_+(x,y,t)\,{}_a^G D_x^\alpha u(x,y,t) - d_-(x,y,t)\,{}_x^G D_b^\alpha u(x,y,t)\\
&\quad -e_+(x,y,t)\,{}_c^G D_y^\beta u(x,y,t) - e_-(x,y,t)\,{}_y^G D_d^\beta u(x,y,t) = f(x,y,t),\\
&\quad (x,y) \in \Omega := (a,b)\times(c,d), \quad t \in (0,T], \quad 1 < \alpha, \beta < 2\\
&\quad u(x,y,t) = 0, \qquad (x,y) \in \partial\Omega, \quad t \in [0,T],\\
&\quad u(x,y,0) = u_o(x,y), \qquad (x,y) \in \overline{\Omega}.
\end{aligned} \tag{11}$$

- The fractional spatial derivatives are only in the coordinate directions.

- A two-dimensional shifted finite difference scheme is

$$
\begin{aligned}
&\frac{u_{i,j}^m - u_{i,j}^{m-1}}{\Delta t} - \frac{d_{i,j}^{+,m}}{h_1^\alpha} \sum_{k=0}^{i} g_k^{(\alpha)} u_{i-k+1,j}^m - \frac{d_{i,j}^{-,m}}{h_1^\alpha} \sum_{k=0}^{N_1-i+1} g_k^{(\alpha)} u_{i+k-1,j}^m \\
&- \frac{e_{i,j}^{+,m}}{h_2^\beta} \sum_{l=0}^{j} g_j^{(\beta)} u_{i,j-l+1}^m - \frac{e_{i,j}^{-,m}}{h_2^\beta} \sum_{k=0}^{N_2-i+1} g_l^{(\beta)} u_{i,j+l-1}^m = f_{i,j}^m, \\
&1 \le i \le N_1, \quad 1 \le j \le N_2, \quad m = 1, 2, \ldots, M.
\end{aligned}
\tag{12}
$$

Let $N = N_1 N_2$. Introduce N-dimensional vectors u^m and f^m defined by

$$
\begin{aligned}
u^m &:= \left[u_{1,1}^m, \cdots, u_{N_1,1}^m, u_{1,2}^m, \cdots, u_{N_1,2}^m, \cdots, u_{1,N_2}^m, \cdots, u_{N_1,N_2}^m\right]^T, \\
f^m &:= \left[f_{1,1}^m, \cdots, f_{N_1,1}^m, f_{1,2}^m, \cdots, f_{N_1,2}^m, \cdots, f_{1,N_2}^m, \cdots, f_{N_1,N_2}^m\right]^T.
\end{aligned}
\tag{13}
$$

The finite difference scheme (12) can be expressed in the matrix form

$$
(I + \Delta t A^m) u^m = u^{m-1} + \Delta t f^m. \tag{14}
$$

- $A^{m,x}$ accounts for the coupling of all the nodes in the x direction
  - $A^{m,x}$ is block-diagonal with full diagonal blocks.
  - Each diagonal block $A_j^{m,x}$ is identical to that for a 1D problem

$$A_j^{m,x} = -\mathrm{diag}(r_j^{+,m})T^{\alpha,N_1} - \mathrm{diag}(r_j^{-,m})(T^{\alpha,N_1})^T. \qquad (15)$$

  - $A^{m,x}v$ can be evaluated in $N_2 O(N_1 \log N_1) = O(N \log N)$ operations.
  - $A^{m,x}$ can be stored in $N_2 O(N_1) = O(N)$ memory.
- $A^{m,y}$ accounts for the coupling of all the nodes in the y direction.
  - As the labelling runs x first, $A^{m,y}$ is a full block matrix but with sparse matrix blocks.
  - We prove that $A^{m,y}$ is block-Toeplitz-circulant-block

$$A^{m,y} = -\mathrm{diag}\big(s_j^{+,m}\big)_{j=1}^{N_2}\big(T^{\beta,N_2} \otimes I_{N_1}\big) - \mathrm{diag}\big(s_j^{-,m}\big)_{j=1}^{N_2}\big((T^{\beta,N_2})^T \otimes I_{N_1}\big). \qquad (16)$$

- In the numerical experiments the data are given as follows
  - $a_+(x,y,z,t) = a_-(x,y,z,t) = b_+(x,y,z,t) = b_-(x,y,z,t) = c_+(x,y,z,t) = c_-(x,y,z,t) = D = 0.005$
  - $f = 0$, $\alpha = \beta = \gamma = 1.8$, $\Omega = (-1,1)^3$, $[0,T] = [0,1]$.
  - The true solution is expressed via the inverse Fourier transform

$$\begin{aligned} u(x,y,z,t) &= \frac{1}{\pi}\int_0^\infty e^{-2D|\cos(\frac{\pi\alpha}{2})|(t+0.5)\xi^\alpha}\cos(\xi x)d\xi \\ &\quad \times \frac{1}{\pi}\int_0^\infty e^{-2D|\cos(\frac{\pi\beta}{2})|(t+0.5)\eta^\beta}\cos(\eta y)d\eta \\ &\quad \times \frac{1}{\pi}\int_0^\infty e^{-2D|\cos(\frac{\pi\gamma}{2})|(t+0.5)\zeta^\gamma}\cos(\zeta z)d\zeta. \end{aligned} \tag{17}$$

  - The initial condition $u_o(x,y,z)$ is chosen to be $u(x,y,z,0)$.
- The Meerschaert & Tadjeran FDM and the fast FDM implemented in Fortran 90 on a workstation of 120 GB of memory.

Table: The CPU of the FDM and fast FDM

$h = \Delta t$	# of nodes	The FDM	The fast FDM
2^{-3}	4,096	1h 4m 26s	0.58s
2^{-4}	32,768	2 months 25d 9h 12m	5.74s
2^{-5}	262,144	N/A	1m 6s
2^{-6}	2,097,152	N/A	14m 22s
2^{-7}	16,777,216	N/A	3h 49m 56s
2^{-8}	134,217,728	N/A	3days 3h 18m 52s

- It would take the regular FDM about 1,000 years of CPU times on state of the art supercomputers (10 petaflops, Nov 2011) to finish the simulation, provided that the computer has enough memory.
- Parallelization was used in measuring the peak performance of supercomputers. The nonlocal nature of FPDEs makes the communications in the numerical simulations global, which further increases the CPU times of the FDM simulations.

Conservative FDEs (del-Castillo-Negrete et al. 2004; Ervin & Roop 2005; Wheatcraft & Meerschaert 2008; Zhang et al. 2007)

$$
\begin{aligned}
&-D\Big(K(x)\big(\theta\ {}_0^{C,l}D_x^{1-\beta}u - (1-\theta)\ {}_x^{C,r}D_1^{1-\beta}u\big)\Big) = f(x), \quad x \in (0,1), \\
&u(0) = u_l, \quad u(1) = u_r, \qquad 0 < \beta < 1, \quad 0 \le \theta \le 1.
\end{aligned}
\tag{18}
$$

- derived from a local mass balance + a fractional Fick's law.
- θ is the weight of forward versus backward transition probability.
- The left- and right-fractional integrals, Caputo and Riemann-Liouville fractional derivatives are defined by

$$
\begin{aligned}
{}_0I_x^{\beta}u(x) = {}_0D_x^{-\beta}u(x) &:= \int_0^x \frac{(x-s)^{\beta-1}u(s)}{\Gamma(\beta)}ds, \\
{}_xI_1^{\beta}u(x) = {}_xD_1^{-\beta}u(x) &:= \int_x^1 \frac{(s-x)^{\beta-1}u(s)}{\Gamma(\beta)}ds, \\
{}_0^{C,l}D_x^{1-\beta}u := {}_0I_x^{\beta}Du, &\qquad {}_x^{C,r}D_1^{1-\beta}u := -{}_xI_1^{\beta}Du, \\
{}_0^{R,l}D_x^{1-\beta}u := D\ {}_0I_x^{\beta}u, &\qquad {}_x^{C,r}D_1^{1-\beta}u := -D\ {}_xI_1^{\beta}u.
\end{aligned}
\tag{19}
$$

- Galerkin formulation: given $f \in H^{-(1-\frac{\beta}{2})}(0,1)$, seek $u \in H_0^{1-\frac{\beta}{2}}(0,1)$

$$B(u,v) = \langle f, v\rangle, \qquad \forall\, v \in H_0^{1-\frac{\beta}{2}}(0,1). \tag{20}$$

Here $B : H_0^{1-\frac{\beta}{2}}(0,1) \times H_0^{1-\frac{\beta}{2}}(0,1) \to \mathbb{R}$ is defined to be

$$\begin{aligned} B(u,v) :=\ & \theta\big\langle K\ {}_0D_x^{-\beta} Du, Dv\big\rangle + (1-\theta)\big\langle K\ {}_xD_1^{-\beta} Du, Dv\big\rangle \\ =\ & \theta\big(K\ {}_0D_x^{-\beta/2} Du, {}_xD_1^{-\beta/2} Dv\big)_{L^2(0,1)} \\ & +(1-\theta)\big(K\ {}_xD_1^{-\beta/2} Du, {}_0D_x^{-\beta/2} Dv\big)_{L^2(0,1)} \end{aligned}$$

$\langle\cdot,\cdot\rangle$ is the duality pair between $H^{-(1-\frac{\beta}{2})}(0,1)$ and $H_0^{1-\frac{\beta}{2}}(0,1)$.

- The coercivity of $B(\cdot,\cdot)$ is derived as follows

$$\begin{aligned} B(u,u) &= K\big({}_0D_x^{-\beta/2}Du, {}_xD_1^{-\beta/2}Du\big)_{L^2(0,1)} \\ &= -\cos\big((1-\beta/2)\pi\big)K|u|^2_{H^{1-\beta/2}(0,1)} \\ &= \cos\big(\beta\pi/2\big)K|u|^2_{H^{1-\beta/2}(0,1)}. \end{aligned}$$

Theorem

$B(\cdot,\cdot)$ is coercive and continuous on $H_0^{1-\frac{\beta}{2}}(0,1) \times H_0^{1-\frac{\beta}{2}}(0,1)$. Hence, the Galerkin weak formulation (20) *has a unique solution. Moreover,*

$$\|u\|_{H^{1-\frac{\beta}{2}}(0,1)} \leq C\|f\|_{H^{-(1-\frac{\beta}{2})}(0,1)}.$$

- Let $S_h(0,1) \subset H_0^{1-\frac{\beta}{2}}(0,1)$ be the finite element space of piecewise polynomials of degree $m-1$. Find $u_h \in S_h(0,1)$ such that

$$B(u_h, v_h) = \langle f, v_h \rangle, \qquad \forall v_h \in S_h(0,1).$$

- **Assume** that the true solution $u \in H^m(0,1) \cap H_0^{1-\frac{\beta}{2}}(0,1)$. Then the optimal-order error estimate in the energy norm holds

$$\|u_h - u\|_{H^{1-\frac{\beta}{2}}(0,1)} \leq Ch^{m-1+\beta/2} \|u\|_{H^m(0,1)}.$$

- **Assume** that the dual problem has the full regularity for **each** $g \in L^2$. Then the optimal-order error estimate in the L^2 norm holds **for** $u \in H^m(0,1) \cap H_0^{1-\frac{\beta}{2}}(0,1)$
- Extensions to spectral Galerkin methods and other methods were proved under the same assumptions.

- Conservative and non-conservative FDEs are not equivalent.
- Finite element/volume methods are suited for conservative FDEs.
- Finite difference methods are suited for nonconservative FDEs.
- In many applications, local mass conservation is crucial.
- A finite-volume scheme naturally has second-order accuracy in space, without a Richardson extrapolation as in finite difference methods.
- Let $u = \sum_{j=1}^N u_j \phi_j$, $\boldsymbol{u} := [u_1, u_2, \dots, u_N]^T$, $\boldsymbol{f} := [f_1, f_2, \dots, f_N]^T$, $\boldsymbol{A} := [A_{i,j}]_{i,j=1}^N$. Integrating (18) over $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ yields

$$
\begin{aligned}
&\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f}, \quad f_i := \int_{x_{i-1/2}}^{x_{i+1/2}} f(x)dx, \quad 1 \le i, j \le N. \\
&A_{i,j} := \Big[K(x) \big(\theta \, {}_0^{C,l} D_x^{1-\beta} u - (1-\theta) \, {}_x^{C,r} D_1^{1-\beta} u \big) \Big]_{x=x_{i+1/2}}^{x=x_{i-1/2}}.
\end{aligned}
\tag{21}
$$

Theorem

$$A = \gamma(\beta)\big(K_-\ T_L^{\beta,N} + K_+\ T_R^{\beta,N}\big), \quad K_\pm := \mathrm{diag}\big(\{K(x_{i\pm\frac{1}{2}})\}_{i=1}^N\big) \tag{22}$$

where $T_L^{\beta,N}$ and $T_R^{\beta,N}$ are full *Toeplitz matrices. So A can be stored in $O(N)$ memory and Av can be evaluated in $O(N\log N)$ operations for any $v \in \mathbb{R}^N$.*

+ A fast Krylov subspace iterative method reduces the computational complexity of each iteration from $O(N^2)$ to $O(N\log N)$.
− For problem (18), the condition number $\kappa(A) = O(h^{-(2-\beta)})$.
− The number of Krylov iterations is $O(h^{-(1-\beta/2)}) = O(N^{1-\beta/2})$, leading to an overall computational complexity of $O(N^{2-\beta/2}\log N)$.

- This calls for an effective and efficient preconditioner.

Theorem

$M := T_L^{\beta,N} + T_R^{\beta,N}$ *is a full symmetric and positive-definite, Toeplitz matrix.*

- Outline of (a perburbation-based) proof: Let $K_0 := \mathrm{diag}\big(\{K(x_i)\}_{i=1}^N\big)$.

$$\begin{aligned} &\gamma(\beta)^{-1}K_0^{-1}A \\ &\quad = K_0^{-1}K_-\ T_L^{\beta,N} + K_0^{-1}K_+\ T_R^{\beta,N} \\ &\quad = K_0^{-1}\big[K_0 + (K_- - K_0)\big]T_L^{\beta,N} + K_0^{-1}\big[K_0 + (K_+ - K_0)\big]T_L^{\beta,N} \\ &\quad = M + K_0^{-1}\Big[(K_- - K_0)T_L^{\beta,N} + (K_+ - K_0)T_R^{\beta,N}\Big] \\ &\quad = M + O(h). \end{aligned} \tag{23}$$

- M is a good preconditioner for the finite volume scheme (21).

$$\big(K_0^{-1}K_-\ T_L^{\beta,N} + K_0^{-1}K_+\ T_R^{\beta,N}\big)u = \gamma(\beta)^{-1}K_0^{-1}Au = \gamma(\beta)^{-1}K_0^{-1}f. \tag{24}$$

- M can be inverted via the superfast algorithm (Ammar & Gragg, 1988) in $O(N\log^2 N)$ operations.

- The data in (18):$\beta = 0.2$, $\theta = 0.5$, $K(x) = \Gamma(1.2)(1+x)$, $u_l = u_r = 0$.
- The true solution $u(x) = x^2(1-x)^2$, f is computed accordingly

	Gauss			CGS		
N	$\lVert u-u_G\rVert_{L^\infty}$	CPU(s)		$\lVert u-u_C\rVert_{L^\infty}$	CPU(s)	Itr. #
2^5	2.018×10^{-4}	0.000		2.018×10^{-4}	0.000	32
2^6	5.157×10^{-5}	0.000		5.157×10^{-5}	0.000	65
2^7	1.294×10^{-5}	0.000		1.294×10^{-5}	0.016	128
2^8	3.214×10^{-6}	0.047		3.214×10^{-6}	0.141	217
2^9	7.893×10^{-7}	0.500		7.893×10^{-7}	3.359	599
2^{10}	1.887×10^{-7}	7.797		1.886×10^{-7}	2 m 2 s	1,110
2^{11}	4.030×10^{-8}	2 m 38 s		4.047×10^{-8}	21 m 13 s	2,624
2^{12}	6.227×10^{-9}	24 m 29 s		7.468×10^{-8}	4 h 19 m	7,576
2^{13}	5.783×10^{-9}	3 h 27 m		N/A	$>$ 2 days	$>$ 20,000
	FCGS			PFCGS		
	$\lVert u-u_F\rVert_{L^\infty}$	CPU(s)	Itr. #	$\lVert u-u_S\rVert_{L^\infty}$	CPU(s)	Itr. #
2^5	2.018×10^{-4}	0.000	32	2.018×10^{-4}	0.000	6
2^6	5.157×10^{-5}	0.016	63	5.157×10^{-5}	0.000	5
2^7	1.294×10^{-5}	0.031	128	1.294×10^{-5}	0.000	5
2^8	3.214×10^{-6}	0.125	248	3.214×10^{-6}	0.006	5
2^9	7.893×10^{-7}	0.578	576	7.893×10^{-7}	0.016	5
2^{10}	1.886×10^{-7}	2.281	1,078	1.887×10^{-7}	0.047	5
2^{11}	4.037×10^{-8}	9.953	1,997	4.038×10^{-8}	0.078	5
2^{12}	1.587×10^{-8}	57.27	5,130	6.194×10^{-9}	0.188	5
2^{13}	2.372×10^{-8}	2 m 52 s	7,410	4.345×10^{-9}	0.391	5

- Use the numerical solutions by Gaussian elimination as a benchmark:
  - The conjugate gradient squared (CGS) method diverges, due to significant amount of round-off errors.
  - The fast CGS (FCGS) reduced the CPU time significantly, as the operations for each iteration is reduced from $O(N^2)$ to $O(N \log N)$.
    - The number of iterations is still $O(N^{1-\beta/2})$,
    - It is less accurate than Gaussian at fine meshes due to round-off errors.
  - The preconditioner M is optimal, so the preconditioned FCGS (PFCGS) has an overall computational cost of $O(N \log^2 N)$.
    - It significantly reduces round-off errors.
    - It generates more accurate solutions than Gaussian elimination.
    - It further reduces CPU time.

Regularity of the boundary-value problem of FDEs (Jin et al 2015; W. et al 2014, 2016; W. & Zhang 2015)

- Error estimates were proved for numerical methods for FDEs, under the **assumption** that the true solution is smooth.
- For integer-order elliptic or parabolic PDEs, smooth data (and domain for multi-D problem) $\Longrightarrow$ smooth solution.
  - $u(x) = (x^{2-\beta} - x^{1-\beta})/\Gamma(3-\beta) \notin W^{1,1/\beta}(0,1)$ is the solution of
$$D\big(\,{}_0D_x^{-\beta} Du\big) = 1, \quad x \in (0,1), \qquad u(0) = u(1) = 0 \tag{25}$$
  - In particular, $u \notin H^1(0,1)$ for $1/2 \le \beta \le 1$.
- For FDEs smooth data does not ensure smooth solutions
  - No conditions in the literature to ensure smooth solutions to FDEs.
  - The Nitsche-lifting based proof of optimal-order L^2 error estimates in the literature does not hold even for constant $K > 0$.
  - What conditions ensures that high-order methods $\Longrightarrow$ high-order convergence rates?
  - Solutions may have boundary layers and other singularity, which need to be resolved numerically.

- Solutions to FDEs with smooth data and domain may have boundary layers, a uniform mesh is not effective.
  - Finite-difference methods out of the question, as Grünwald-Letnikov derivatives are inherently defined on uniform meshes.
  - Riemann-Liouville and Caputo derivatives offer such flexibilities.
- Bebause of the nonlocal nature of FDEs, a numerical scheme discretized on an arbitrarily adaptively refined mesh
  - offers great flexbility and effective approximation property
  - offers possible advantage on its theoretical analysis
  - destroys the structure of its stiffness matrix and so efficiency.
- Motivation: balancing flexibility and efficiency.

- We assume a geometrically refined mesh towards the left endpoint.

Theorem

The matrix A can be decomposed as

$$\begin{aligned} A &= \frac{1}{\Gamma(\beta+1)}\Big[\mathrm{diag}(K^-)\big(\gamma Q_l + (1-\gamma)Q_r\big) \\ &\qquad -\mathrm{diag}(K^+)\big(\gamma P_l + (1-\gamma)P_r\big)\Big]\mathrm{diag}\big(\{h_i^{\beta-1}\}_{i=1}^m\big). \end{aligned}$$

- P_l, P_r, Q_l and Q_r are Toeplitz.
- A has an additional diagonal matrix (reflecting the impact of the mesh sizes) multiplier to that on the uniform mesh.

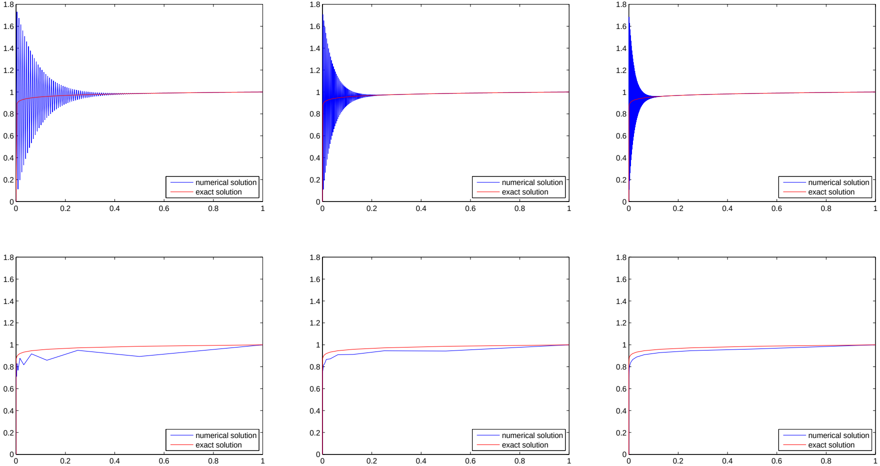
- Consider (18) with $K = 1$, $f = 0$, $\beta = 0.98$, $\theta = 1$, $u_l = 0$, $u_r = 1$, i.e.,

$$
\begin{aligned}
D\big({}_0D_x^{-\beta} Du\big) &= 0, \quad x \in (0,1), \\
u(0) = 0, \quad u(1) &= 1
\end{aligned}
$$

Its solution $u(x) = x^{1-\beta}$ for $x \in (0,1)$.

	N	CPU	#of iterations
Gauss	256	0.640s	
	512	5.567s	
	1024	59s	
CGS	256	2.978s	256
	512	29s	512
	1024	403s	1024
FCGS	256	0.073s	256
	512	0.139s	512
	1024	0.391s	1024

Figure: First row: numerical solutions on a uniform mesh of $n = 256, 512, 1024$; Second row: numerical solutions on a geometrically refined mesh $n = 48, 64, 96$.



- Solutions to FDEs with smooth data and domain may have boundary layers. Numerical solution of FDEs
  - with a uniform mesh is not effective.
  - with a gridded mesh may resolve the boundary layers, but does not necessarily provide an accurate global approximation.
- We propose to use a composite mesh that consists of
  - a uniform mesh in most of the domain,
  - a gridded mesh in the cells near the (left) boundary.
- The key issue is the structure of the stiffness matrix:

$$A = \begin{bmatrix} A_{l,l} & A_{l,r} \\ A_{r,l} & A_{r,r} \end{bmatrix}. \tag{26}$$

  - $A_{r,r}$, corresponding to the uniform mesh, has a Toeplitz-like structure.
  - $A_{l,l}$, corresponding to the gridded mesh, has a Toeplitz-like structure with an extra right diagonal multiplier.

- The off-diagonal submatrices $A_{l,r}$ and $A_{r,l}$
  - are full due to the nonlocal nature of FDEs,
  - are not Toeplitz-like.

Theorem

$$
\begin{aligned}
A_{l,r} &= \frac{(1-\gamma)h^{\beta-1}}{\Gamma(\beta+1)}\big(\mathrm{diag}(K_l^-)E - \mathrm{diag}(K_l^+)D\big),\\
A_{r,l} &= \frac{\gamma}{\Gamma(\beta+1)}(\mathrm{diag}(K_r^-)H - \mathrm{diag}(K_r^+)G)\mathrm{diag}(\{h_i^{\beta-1}\}_{i=1}^m).
\end{aligned}
$$

- The typical entries of D and E are of the form

$$
d_{i,j} = 2(j+1-3\cdot 2^{i-m-1})^\beta - (j-3\cdot 2^{i-m-1})^\beta - (j+2-3\cdot 2^{i-m-1})^\beta,
$$

$$
\begin{aligned}
g_{i,j} = {}& \left[2^{m-j+1}\left(i+\frac{3}{2}\right)-1\right]^\beta - \frac{3}{2}\left[2^{m-j+1}\left(i+\frac{3}{2}\right)-2\right]^\beta\\
&+\frac{1}{2}\left[2^{m-j+1}\left(i+\frac{3}{2}\right)-4\right]^\beta.
\end{aligned}
$$

- Use a fractional binomial expansion, we have

$$
\begin{aligned}
D \approx & -2\binom{\beta}{2}[1,1,\ldots,1]^T\left[\frac{1}{2^{2-\beta}},\frac{1}{3^{2-\beta}},\ldots,\frac{1}{(n-1)^{2-\beta}}\right] \\
& -2\binom{\beta}{4}[1,1,\ldots,1]^T\left[\frac{1}{2^{4-\beta}},\frac{1}{3^{4-\beta}},\ldots,\frac{1}{(n-1)^{4-\beta}}\right] \\
& +18\binom{\beta}{3}[2^{-m},2^{-m+1},\ldots,2^{-1}]^T\left[\frac{1}{2^{3-\beta}},\frac{1}{3^{3-\beta}},\ldots,\frac{1}{(n-1)^{3-\beta}}\right] \\
& -108\binom{\beta}{4}[2^{-2m},2^{-2m+2},\ldots,2^{-2}]^T\left[\frac{1}{2^{4-\beta}},\frac{1}{3^{4-\beta}},\ldots,\frac{1}{(n-1)^{4-\beta}}\right].
\end{aligned}
$$

  - The matrices can be approximated by a finite sum of low-rank matrices.
  - The matrix-vector multiplication can be performed in $O(N)$ operations.

- A preconditioner based on T. Chan's circulant preconditioner C_n, which minimizes $\|A - C_n\|_F$ over all circulant matrices.
- We define a block-diagonal-circulant-block preconditioner M for A

$$M := \begin{bmatrix} M_1 & 0 \\ 0 & M_2 \end{bmatrix} \tag{27}$$

  - M_1 is a preconditioner for $A_{l,l}$
  - M_2 is a preconditioner for $A_{r,r}$

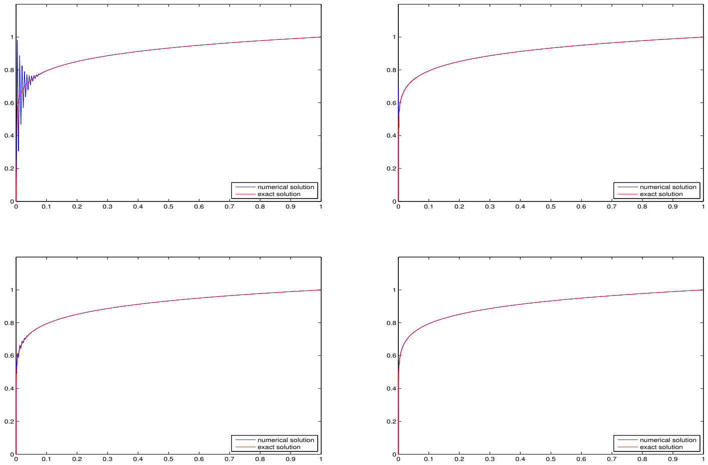
- Consider (18) with $K = 1$, $f = 0$, $\theta = 1$, $\beta = 0.9$, $u_l = 0$, $u_r = 1$, i.e.,

$$\begin{aligned} D\big({}_0D_x^{-\beta} Du\big) &= 0, \quad x \in (0,1), \\ u(0) = 0, \quad u(1) &= 1 \end{aligned}$$

Its solution $u(x) = x^{1-\beta}$ for $x \in (0,1)$.

n	$\|u_n - u\|$	$\|u_{n,m} - u\|$	$\|u_{n,m} - u\|$
128	4.3546×10^{-1}	2.6805×10^{-1}, $m = 7$	2.0315×10^{-1}, $m = 11$
256	4.0630×10^{-1}	2.3336×10^{-1}, $m = 8$	1.3403×10^{-1}, $m = 16$
512	3.7909×10^{-1}	2.0315×10^{-1}, $m = 9$	8.2504×10^{-2}, $m = 22$
1024	3.5370×10^{-1}	1.7685×10^{-1}, $m = 10$	3.8488×10^{-2}, $m = 32$
8192	2.8730×10^{-1}	1.6668×10^{-1}, $m = 13$	N/A

Figure: First row: numerical solutions on a uniform mesh of n=256, 8192;
Second row: numer. solns. on a composite mesh with $n = 256$ and m = 8, 16.



- Consider (18) with $K = 1$, $\theta = 0.5$, $\beta = 0.95$, $u_l = 0$, $u_r = 1$,

$$f(x) = \frac{(1-\gamma)(1-\beta)}{\Gamma(\beta)x(1-x)^{1-\beta}}, \qquad u(x) = x^{1-\beta}, \quad x \in (0,1).$$

	m	n	Error	Iterations
Gauss	2^3	2^8	1.4379×10^{-1}	
	2^4	2^9	1.0491×10^{-1}	
	2^5	2^{10}	5.8194×10^{-2}	
CGS	2^3	2^8	1.4379×10^{-2}	48
	2^4	2^9	1.0491×10^{-1}	77
	2^5	2^{10}	5.8194×10^{-2}	142
FCGS	2^3	2^8	1.4379×10^{-1}	48
	2^4	2^9	1.0491×10^{-1}	78
	2^5	2^{10}	5.8194×10^{-2}	150
PFCGS	2^3	2^8	1.4379×10^{-1}	9
	2^4	2^9	1.0491×10^{-1}	13
	2^5	2^{10}	5.8194×10^{-2}	16

Table: Numerical results on a uniform mesh

	n	Error	Iterations	CPUs
Gauss	2^8	1.8827×10^{-1}		0.01s
	2^9	1.8206×10^{-1}		0.01s
	2^{10}	1.7596×10^{-1}		0.05s
	2^{11}	1.7002×10^{-1}		0.25s
	2^{12}	1.6425×10^{-1}		1.25s
	2^{13}	1.5867×10^{-1}		9.76s
	2^{14}	1.5327×10^{-1}		97s
CGS	2^8	1.8827×10^{-1}	46	0.01s
	2^9	1.8206×10^{-1}	66	0.01s
	2^{10}	1.7596×10^{-1}	94	0.18s
	2^{11}	1.7002×10^{-1}	133	0.86s
	2^{12}	1.6425×10^{-1}	188	4.94s
	2^{13}	1.5867×10^{-1}	266	30.78s
	2^{14}	1.5327×10^{-1}	379	187s
FCGS	2^8	1.8827×10^{-1}	46	0.05s
	2^9	1.8206×10^{-1}	66	0.16s
	2^{10}	1.7596×10^{-1}	94	0.29s
	2^{11}	1.7002×10^{-1}	133	1.16s
	2^{12}	1.6425×10^{-1}	188	2.00s
	2^{13}	1.5867×10^{-1}	266	12s
	2^{14}	1.5327×10^{-1}	379	27s
PFCGS	2^8	1.8827×10^{-1}	8	0.02s
	2^9	1.8206×10^{-1}	8	0.02s
	2^{10}	1.7596×10^{-1}	9	0.05s
	2^{11}	1.7002×10^{-1}	10	0.09s
	2^{12}	1.6425×10^{-1}	10	0.14s
	2^{13}	1.5867×10^{-1}	10	0.66s
	2^{14}	1.5327×10^{-1}	11	1.00s

- FPDEs have significantly different mathematical and numerical properties from their integer-order analogues. For instance,
- For $u_l = u_r = 0$, (18) and its Riemann-Liouville analogue coincide. They are well posed if K is a positive constant.
- When u_l, u_r do not vanish, then (18) is well posed for a postive constant K. But its Riemann-Liouville analogue does not admit a solution.

Lemma

$B(w,w) < 0$ for some $K(x)$ of two positive constants and $w \in H_0^{1-\frac{\beta}{2}}(0,1)$

Let $K(x)$ and $w \in H_0^1(0,1) \subset H_0^{1-\frac{\beta}{2}}(0,1)$ be defined by

$$K(x) := \begin{cases} K_l, & x \in (0,1/2), \\ 1, & x \in (1/2,1). \end{cases} \qquad w(x) := \begin{cases} 2x, & x \in (0,1/2], \\ 2(1-x), & x \in [1/2,1). \end{cases}$$

$${}_0^{C,l}D_x^{1-\beta} w(x) = \begin{cases} 2x^\beta/\Gamma(\beta+1), & x \in (0,1/2), \\ 2\big(x^\beta - 2(x-1/2)^\beta\big)/\Gamma(\beta+1), & x \in (1/2,1). \end{cases}$$

$$B(w,w) = 2^{1-\beta}\big(K_l - \big(2^{\beta+1}-3\big)\big)/\Gamma(\beta+2).$$

As $0 < \log_2 3 - 1 < 1$, choose $\log_2 3 - 1 < \beta < 1$ so that $2^{\beta+1} - 3 > 0$. Select $K_l > 0$ such that $K_l - \big(2^{\beta+1}-3\big) < 0$. For such K and w, $B(w,w) < 0$.

- Consider the one-sided problem ((18) with $\theta = 1$)

$$-D\big(K\ {}_0D_x^{-\beta} Du\big) = f(x), \quad x \in (0,1), \quad u(0) = u(1) = 0. \tag{28}$$

- For a variable K

$$\begin{aligned} B(u,v) \quad &= \theta\big\langle K\ {}_0I_x^{\beta} Du, Dv\big\rangle + (1-\theta)\big\langle K\ {}_xI_1^{\beta} Du, Dv\big\rangle \\ &\neq \theta\big\langle K Du, {}_xI_1^{\beta} Dv\big\rangle + (1-\theta)\big\langle K Du, {}_0I_x^{\beta} Dv\big\rangle \\ &\neq \big(K\ {}_0I_x^{\beta/2} Du, {}_xI_1^{\beta/2} Dv\big)_{L^2(0,1)} \end{aligned}$$

- Even the best possible (last) form cannot guarantee the coercivity of B

$$\begin{aligned} &\big(K\ {}_0I_x^{\beta/2} Du, {}_xI_1^{\beta/2} Du\big)_{L^2(0,1)} \\ &\quad \not\geq K_{min}\big({}_0I_x^{\beta/2} Du, {}_xI_1^{\beta/2} Du\big)_{L^2(0,1)} \\ &\quad = \cos\big(\beta\pi/2\big) K_{min} |u|^2_{H^{1-\beta/2}(0,1)}. \end{aligned}$$

  - ${}_0I_x^{\beta/2} Du$ and ${}_xI_1^{\beta/2} Du$ do not always have the same sign on (0,1).
  - One can choose a smooth K such that the left-hand side negative.

- For a variable K, the Galerkin formulation is *not* coercive on any product space $H \times H$ so $H_0^{1-\frac{\beta}{2}}(0,1) \times H_0^{1-\frac{\beta}{2}}(0,1)$ is not a feasible choice.
- That the FDE is a local mass balance incorporated with a fractional Fick's law motivates a Petrov-Galerkin formulation: Seek $u \in H_0^{1-\beta}(0,1)$ such that

$$A(u,v) := \int_0^1 K(x)\ \big({}_0D_x^{-\beta} Du\big) Dv dx = \langle f, v\rangle, \quad \forall v \in H_0^1(0,1) \tag{29}$$

Theorem

Assume $0 < \beta < 1/2$ *and* $0 < K_{min} \le K \le K_{max} < \infty$. *Then*

$$\begin{aligned} \inf_{w \in H_0^{1-\beta}(0,1)} \sup_{v \in H_0^1(0,1)} \frac{A(w,v)}{\|w\|_{H^{1-\beta}(0,1)} \|v\|_{H^1(0,1)}} \ge \gamma(\beta) > 0, \\ \sup_{w \in H_0^{1-\beta}(0,1)} A(w,v) > 0 \quad \forall\, v \in H_0^1(0,1) \setminus \{0\}. \end{aligned} \tag{30}$$

Hence, (29) has a unique solution $u \in H_0^{1-\beta}(0,1)$ *with the estimate*

$$\|u\|_{H^{1-\beta}(0,1)} \le (K_{max}/\gamma)\|f\|_{H^{-1}(0,1)}. \tag{31}$$

- $P_N[-1,1]$: the space of polynomials of degree $\le N$ on $[-1,1]$
- $L_n(x)$: the nth degree Legendre polynomial on $[-1,1]$

$$L_0(x) = 1, \quad L_1(x) = x, \quad L_{n+1}(x) = \frac{2n+1}{n+1} x L_n(x) - \frac{n}{n+1} L_{n-1}(x), \ n \ge 1,$$

$$\int_{-1}^{1} L_n(x) L_m(x) dx = \frac{2}{2n+1} \delta_{m,n}, \qquad L_n(\pm 1) = (\pm 1)^n$$

- $\phi_n(x) := L_n(x) - L_{n+2}(x)$ are linearly independent with $\phi(\pm 1) = 0$.

$$S_N[-1,1] := \{v \in P_N[-1,1] : v(-1) = v(1) = 0\} = \mathrm{span}\{\phi_n\}_{n=0}^{N-2}.$$

- A spectral-Galerkin method: Seek $u_N \in S_N[-1,1]$ such that

$$B(u_N, v_N) = \langle f, v_N \rangle, \qquad \forall v_N \in S_N[-1,1].$$

Theorem

(Huang et al. 2013; Zeng et al. 2014) If $u \in H^r \cap H_0^{1-\beta/2}$ and $1-\beta/2 \le s \le r$, then

$$\|u_N - u\|_{H^s} \le CN^{-(r-s)}\|u\|_{H^r}, \quad 1-\beta/2 \le s \le r. \tag{32}$$

Assume full regularity of the dual problem for each right-hand side, then the estimate holds for $0 \le s \le r$.

Theorem

(W. & Zhang 2015) The solution u to problem (18) with $0 < \beta < 1/2$, constant K and f and $u_l = u_r = 0$ is not in $H^{\frac{3}{2}-\beta}$ but in $B_\infty^{\frac{3}{2}-\beta}(L^2)$.
The best provable convergence rate in (32) is $r = \frac{3}{2}-\beta$.

- In particular, the convergence rate in $\|\cdot\|_{H^{1-\beta/2}}$ is $O(N^{-(1-\beta)/2})$.

Theorem

For $0 < \beta < 1/2$ the true solution u to a one-dimensional, one-sided FDE can be decomposed as

$$u = u_l + \big(u_r - u_l - {}_{-1}^{C}D_1^{\beta} w_f\big)\big({}_{-1}^{C}D_1^{\beta} w_b\big)^{-1}\, {}_{-1}^{C}D_x^{\beta} w_b + {}_{-1}^{C}D_x^{\beta} w_f. \tag{33}$$

$$\begin{aligned} -D\big(K(x)Dw_f\big) &= f, \quad x \in (-1,1); \qquad w_f(-1) = w_f(1) = 0, \\ -D\big(K(x)Dw_b\big) &= 0, \quad x \in (-1,1); \qquad w_b(-1) = 0,\ w_b(1) = 1. \end{aligned} \tag{34}$$

- Use SPG to solve the second-order DE (34) (Canuto et al 2006, Shen et al 2011): Find $w_N \in S_N[-1,1]$ such that

$$\big(K(x)Dw_N, Dv_N\big)_{L^2(-1,1)} = \big(f, v_N\big)_{L^2(-1,1)}, \quad \forall v_N \in S_N[-1,1].$$

- Use (33) to postprocess w_N to obtain u_N

- Properties of $\left({}_{-1}^{C}D_{1}^{\beta}w_b\right)^{-1}$ and ${}_{-1}^{C}D_{x}^{\beta}w_b$
  - ${}_{-1}^{C}D_{x}^{\beta}w_b$ can be evaluated as follows

$$\begin{aligned}
{}_{-1}^{C}D_{x}^{\beta}w_b &= {}_{-1}D_{x}^{-(1-\beta)}Dw_b \\
&= \left(\int_{-1}^{1}\frac{1}{K(s)}ds\right)^{-1}{}_{-1}D_{x}^{-(1-\beta)}\frac{1}{K(x)} \\
&= \left(\int_{-1}^{1}\frac{1}{K(s)}ds\right)^{-1}\frac{1}{\Gamma(1-\beta)}\int_{-1}^{x}\frac{1}{K(s)(x-s)^{\beta}}ds > 0.
\end{aligned}$$

  - $\left({}_{-1}^{C}D_{1}^{\beta}w_b\right)^{-1}$ is well defined and ${}_{-1}^{C}D_{x}^{\beta}w_b$ is bounded in $L^{\infty}(-1,1)$.

- Spetral method offers additional computational benefit
  - Evaluating ${}_{-1}^{C}D_{x}^{\beta}w_N$ requires numerical integration of a weakly singular integral of Dw_N.
  - Spectral method can carry out the calculation analytically in a systematic manner.

- $J_n^{\mu,\nu}(x)$ – the nth order Jacobi polynomials that are orthogonal with respect to the Jacobi weight function $\omega^{\mu,\nu} := (1-x)^\mu(1+x)^\nu$

$$
\begin{aligned}
J_0^{\mu,\nu} &= 1, \quad J_1^{\mu,\nu} = \frac{1}{2}(\mu+\nu+2)x + \frac{1}{2}(\mu-\nu), \\
J_{n+1}^{\mu,\nu} &= \left(a_n^{\mu,\nu}x - b_n^{\mu,\nu}\right)J_n^{\mu,\nu} - c_n^{\mu,\nu}J_{n-1}^{\mu,\nu} \\
&= \frac{n+\mu+1}{n!\Gamma(n+\mu+\nu+1)}\sum_{k=0}^{n}\binom{n}{k}\frac{\Gamma(n+k+\mu+\nu+1)}{\Gamma(k+\mu+1)}\left(\frac{x-1}{2}\right)^k, \\
&\quad n \geq 1
\end{aligned}
$$

where $a_n^{\mu,\nu}$, $b_n^{\mu,\nu}$, and $c_n^{\mu,\nu}$ are constants having explicit expressions.

Theorem

(Huang et al 2011; Shen et al 2011) For $\mu > 0$,

$$
\begin{aligned}
{}_{-1}^{R}D_x^{\mu} L_n(x) &= \frac{\Gamma(n+1)}{\Gamma(n-\mu+1)}(1+x)^{-\mu} J_n^{\mu,-\mu}(x), \quad x \in [-1,1], \\
{}_{x}^{R}D_1^{\mu} L_n(x) &= \frac{\Gamma(n+1)}{\Gamma(n-\mu+1)}(1-x)^{-\mu} J_n^{-\mu,\mu}(x), \quad x \in [-1,1].
\end{aligned}
$$

- The SPG solution $w_N \in S_N[-1,1]$ can be expressed as

$$
w_N(x) = \sum_{n=0}^{N-2} d_n \phi_n(x) = \sum_{n=0}^{N-2} d_n (L_n(x) - L_{n+2}(x)).
$$

$$
\begin{aligned}
{}_{-1}^{C}D_x^{\beta} w_N = {}_{-1}^{R}D_x^{\beta} w_N \quad &= \sum_{n=0}^{N-2} d_n (1+x)^{-\beta} \Big(\frac{\Gamma(n+1)}{\Gamma(n+1-\beta)} J_n^{\beta,-\beta}(x) \\
&\quad - \frac{\Gamma(n+3)}{\Gamma(n+3-\beta)} J_{n+2}^{\beta,-\beta}(x) \Big).
\end{aligned}
$$

Theorem

(W. & Zhang 2015) Let $0 < \beta < 1/2$, $K \in C^m[-1,1]$, *and* $f \in H^{m-1}(-1,1)$ *for any* $m \geq 1$. *Then,*

$$\|u_N - u\|_{L^2(-1,1)} \leq CN^{-m}.$$

where $C = C\big(\beta, m, \|K\|_{C^m[-1,1]}, \|f\|_{H^{m-1}(-1,1)}\big)$.

- $K = 1$, $u_l = 0$, $u_r = 2$, and

$$f(x) \quad = -\frac{\Gamma(7)}{2^{2-\beta}\Gamma(5+\beta)}\Big(\frac{x+1}{2}\Big)^{4+\beta}.$$

- This gives the true solution $u(x) = \Big(\frac{x+1}{2}\Big)^{1-\beta} + \Big(\frac{x+1}{2}\Big)^{6}$.
- For SPG, $\|u_N - u\|_{L^2(-1,1)} \leq C_\kappa N^{-\kappa}$.
- For our improvements, $\|u_N - u\|_{L^2(-1,1)} \leq C_\kappa e^{-\kappa N}$.

Table: The comparison of the SPG and ISPG methods (W. & Zhang 2015)

	$\|u_{SPG,N} - u\|_{L^2(0,1)}$			$\|u_{ISPG,N} - u\|_{L^2(0,1)}$		
N	$\beta = 0.1$	$\beta = 0.5$	$\beta = 0.9$	$\beta = 0.1$	$\beta = 0.5$	$\beta = 0.9$
4	2.139e-03	5.104e-02	1.677	9.377e-03	2.319e-02	7.737e-02
5	1.334e-03	4.195e-02	0.472	8.451e-04	2.823e-03	1.283e-02
6	9.014e-04	3.431e-02	1.331	6.482e-06	1.087e-04	9.541e-04
7	6.738e-04	2.676e-02	0.439	4.185e-07	3.892e-06	7.135e-06
8	5.204e-04	2.308e-02	1.119	5.348e-08	3.943e-07	5.563e-07
9	4.126e-04	1.913e-02	0.415	9.807e-09	6.239e-08	7.625e-08
10	3.342e-04	1.691e-02	0.986	2.280e-09	1.307e-08	1.468e-08
11	2.755e-04	1.454e-02	0.395	6.296e-10	3.324e-09	3.481e-09
12	2.306e-04	1.309e-02	0.893	1.984e-10	9.811e-10	9.807e-10
13	1.955e-04	1.154e-02	0.380	6.952e-11	3.248e-10	3.105e-10
14	1.676e-04	1.052e-02	0.824	2.656e-11	1.183e-10	1.097e-10
15	1.450e-04	9.439e-03	0.366	1.091e-11	4.659e-11	4.182e-11
C_κ	0.034	0.342	2.479	0.675	4.343	35.521
κ	2.016	1.315	0.600	1.800	1.817	1.985

Thank You

for Your Attention!