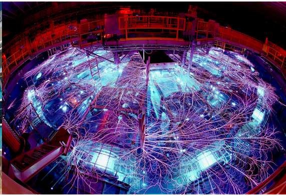
Exceptional service in the national interest

Sandia National Laboratories



Scalable methods for nonlocal models

Christian Glusa

Theoretical and Applied Aspects for Nonlocal Models
July 21, 2022

U.S. DEPARTMENT OF ENERGY
NNSA National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND NO. SAND2022-9177 C

Elliptic nonlocal operators

Let $\delta \in (0, \infty]$ be the *horizon*, $\Omega \subset \mathbb{R}^d$ a bounded open domain, define the *interaction domain*

$$\Omega_I := \{\mathbf{y} \in \mathbb{R}^d \setminus \Omega : |\mathbf{x} - \mathbf{y}| \leq \delta, \quad \text{for } \mathbf{x} \in \Omega\}.$$

We want to numerically solve equations involving the nonlocal operator

$$\mathcal{L}u(\mathbf{x}) = \text{p.v.} \int_{\Omega \cup \Omega_I} (u(\mathbf{y}) - u(\mathbf{x}))\gamma(\mathbf{x}, \mathbf{y}) d\mathbf{y}, \qquad \mathbf{x} \in \Omega,$$

with

$$\begin{aligned} &\gamma(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x}, \mathbf{y}) \, |\mathbf{x} - \mathbf{y}|^{-\beta(\mathbf{x}, \mathbf{y})} \, \mathcal{X}_{|\mathbf{x}-\mathbf{y}| \leq \delta}, \qquad && \mathbf{x}, \mathbf{y} \in \Omega \cup \Omega_I, \\ &\phi(\mathbf{x}, \mathbf{y}) > 0. \end{aligned}$$

- Examples:
  - Integral fractional Laplacian: $\phi \sim$ const, $\beta = d + 2s$, $s \in (0, 1)$, $\delta = \infty$
  - Tempered fractional Laplacian: $\phi(\mathbf{x}, \mathbf{y}) \sim \exp(-\lambda |\mathbf{x} - \mathbf{y}|)$
  - Truncated fractional Laplacian: δ finite
  - Variable order fractional Laplacians with varying coefficient: $\beta(\mathbf{x}, \mathbf{y}) = d + 2s(\mathbf{x}, \mathbf{y})$, $\phi(\mathbf{x}, \mathbf{y}) > 0$
  - Integrable kernels: constant kernel ($\beta = 0$), "peridynamic" kernel ($\beta = 1$)
- Assumptions (for now):
  - γ is symmetric.
  - Interaction domain is defined wrt ℓ_2-norm.

- Nonlocal Poisson's equation:

$$\begin{aligned} -\mathcal{L}u &= f \quad \text{in } \Omega, \\ u &= 0 \quad \text{in } \Omega_I. \end{aligned}$$

- Nonlocal heat equation:

$$\begin{aligned} u_t - \mathcal{L}u &= f \quad \text{in } (0,T) \times \Omega, \\ u &= 0 \quad \text{in } (0,T) \times \Omega_I, \\ u &= u_0 \quad \text{on } \{0\} \times \Omega. \end{aligned}$$

- Source control
- Parameter learning:

$$\min_{u,s,\delta,\dots} \frac{1}{2} \|u - u_d\|_{L^2}^2 + \mathcal{R}(s, \delta, \dots)$$

  subject to nonlocal equation.
- Remark: Homogeneous Dirichlet "boundary" condition for simplicity.

Goal

Assemble and solve nonlocal equations in similar complexity & memory as their local counterparts, i.e. $\mathcal{O}(n \log n)$.

Bilinear form

- We consider

$$a(u,v) = \frac{1}{2}\int_{\Omega} d\mathbf{x} \int_{\Omega} d\mathbf{y}\; \left[\left(u\left(\mathbf{x}\right) - u\left(\mathbf{y}\right)\right)\left(v\left(\mathbf{x}\right) - v\left(\mathbf{y}\right)\right)\right]\gamma(\mathbf{x},\mathbf{y})$$
$$+ \int_{\Omega} d\mathbf{x} \int_{\Omega_I} d\mathbf{y}\, u\left(\mathbf{x}\right) v\left(\mathbf{x}\right) \gamma(\mathbf{x},\mathbf{y}).$$

posed on $\widetilde{H}^s\left(\Omega\right)$ or $L^2\left(\Omega\right)$ respectively, where

$$H^s\left(\Omega\right) := \left\{u \in L^2\left(\Omega\right) \mid \|u\|_{H^s(\Omega)} < \infty\right\}, \quad \widetilde{H}^s\left(\Omega\right) := \left\{u \in H^s\left(\mathbb{R}^d\right) \mid u = 0 \text{ in } \Omega^c\right\},$$

and

$$\|u\|^2_{H^s(\Omega)} = \|u\|^2_{L^2(\Omega)} + \int_{\Omega} d\mathbf{x} \int_{\Omega} d\mathbf{y}\, \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{|\mathbf{x}-\mathbf{y}|^{d+2s}},$$
$$\|u\|^2_{\widetilde{H}^s(\Omega)} = \int_{\mathbb{R}^d} d\mathbf{x} \int_{\mathbb{R}^d} d\mathbf{y}\, \frac{(u(\mathbf{x}) - u(\mathbf{y}))^2}{|\mathbf{x}-\mathbf{y}|^{d+2s}}.$$

- For $\delta = \infty$, if $\gamma(\mathbf{x},\mathbf{y}) = \nabla_{\mathbf{y}} \cdot \mathbf{\Gamma}(\mathbf{x},\mathbf{y})$, can reduce integral from $\Omega \times \Omega^c$ to $\Omega \times \partial\Omega$. (E.g. $\Gamma(\mathbf{x},\mathbf{y}) \sim \frac{\mathbf{x}-\mathbf{y}}{|\mathbf{x}-\mathbf{y}|^{d+2s}}$ for the constant-order fractional kernel.)

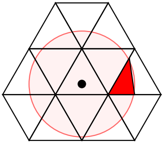
Finite element approximation

- Partition domain into shape-regular mesh $\mathcal{P}_h = \{K\}$ with edges e on the boundary $\partial\Omega$.
- Set $V_h \subset \widetilde{H}^s(\Omega)$ the space of continuous, piecewise linear functions.

$$
\begin{aligned}
a(u,v) =& \frac{1}{2} \sum_K \sum_{\tilde{K}} \int_K d\mathbf{x} \int_{\tilde{K}} d\mathbf{y}\, (u(\mathbf{x}) - u(\mathbf{y}))\, (v(\mathbf{x}) - v(\mathbf{y}))\, \gamma(\mathbf{x},\mathbf{y}) \\
&+ \sum_K \sum_e \int_K d\mathbf{x}\, u(\mathbf{x})\, v(\mathbf{x}) \int_e d\mathbf{y}\, \mathbf{n}_e \cdot \Gamma(\mathbf{x},\mathbf{y}).
\end{aligned}
$$

$\dim V_h =: n$

- Approximate cut elements with simplices, $\mathcal{O}(h_K^2)$ error[1]



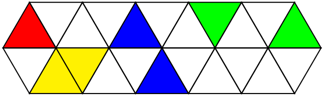
[1] Marta D'Elia, Max Gunzburger, and Christian Vollmann. "A cookbook for approximating Euclidean balls and for quadrature rules in finite element methods for nonlocal problems". In: *Mathematical Models and Methods in Applied Sciences* 31.08 (2021), pp. 1505–1567.

Quadrature

- In subassembly procedure, use quadrature to evaluate element pair contributions:

$$a^{K\times\tilde{K}}(\phi_i,\phi_j) = \frac{1}{2}\int_K d\mathbf{x}\int_{\tilde{K}} d\mathbf{y}\,(\phi_i(\mathbf{x})-\phi_i(\mathbf{y}))\,(\phi_j(\mathbf{x})-\phi_j(\mathbf{y}))\,\gamma(\mathbf{x},\mathbf{y})$$

- Treatment for element pairs $K\cap\tilde{K}\neq\emptyset$:



  - split $K\times\tilde{K}$ into sub-simplices,
  - Duffy transform onto a hypercube, with Jacobian canceling the singularity.
  - Choose quadrature order so that quadrature error $\leq$ discretization error[2]:
    - $|\log h_K|$ if the elements coincide (*red*),
    - $|\log h_K|^2$ if the elements share only an edge (*yellow*),
    - $|\log h_K|^3$ if the elements share only a vertex (*blue*),
    - $|\log h_K|^4$ if the elements are "near neighbours" (*green*), and
    - C if the elements are well separated.

[2] Mark Ainsworth and Christian Glusa. "Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver". In: *Computer Methods in Applied Mechanics and Engineering* (2017).

$\mathcal{O}(n \log n)$ approximations to the stiffness matrix

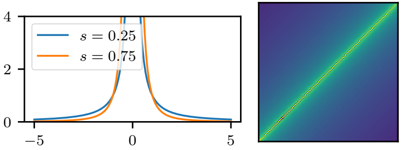


Figure: *Left:* Fractional kernels in $d = 1$ dimensions. *Right:* Magnitude of matrix entries.

Depending on δ and h:

- Straightforward discretization can lead to a fully dense matrix.
- Assembly and solve would have at least $\mathcal{O}(n^2)$ complexity and memory requirement.

Better approach

Panel clustering / Fast Multipole Method / hierarchical matrix approximation

- Find low-rank representations of off-diagonal matrix blocks.
- Lots of methods for computing a structurally sparse approximation, varying level of intrusiveness. I will show what I use: *panel clustering*.
- Important: we don't want to assemble a dense matrix and then compress it.
- Approximation incurs error. The game is to control it so that it is dominated by discretization error.

Cluster method: admissible clusters

First question: Which sub-blocks of the matrix do we want to compress?

Build tree of clusters of DoFs.

- root contains all unknowns
- subdivision based on coordinates
- distributed computations: first level given by partition of unknowns

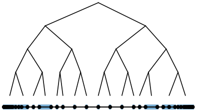


Figure: A cluster tree in $d = 1$ dimensions.

- Find cluster pairs (P, Q) that are *admissible* for approximation: sufficient separation compared to sizes.
- Matrix entries that are not part of any admissible cluster pair are assembled directly into the sparse near-field matrix A_{near}.

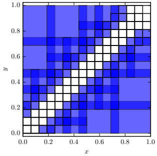


Figure: Elements of admissible cluster pairs in blue. Overlaps in dark blue.

Cluster method – $\mathcal{H}$-matrices

Let $P, Q \subset \Omega$, P and Q admissible.
Let ϕ, ψ be FE basis functions with $\operatorname{supp}\phi \subset P$, $\operatorname{supp}\psi \subset Q$.

$$a(\phi, \psi) = -\int_\Omega \int_\Omega \gamma(\mathbf{x}, \mathbf{y})\, \phi(\mathbf{x})\, \psi(\mathbf{y})\,.$$

Let $\boldsymbol{\xi}_\alpha^P$ be Chebyshev nodes in P and L_α^P the associated Lagrange polynomials. Then

$$\gamma(\mathbf{x}, \mathbf{y}) \approx \sum_{\alpha,\beta=1}^{m^d} \gamma\left(\boldsymbol{\xi}_\alpha^P, \boldsymbol{\xi}_\beta^Q\right) L_\alpha^P(\mathbf{x})\, L_\beta^Q(\mathbf{y})\,, \quad \mathbf{x} \in P,\ \mathbf{y} \in Q.$$

and

$$a(\phi, \psi) \approx -\sum_{\alpha,\beta=1}^{m^d} \gamma\left(\boldsymbol{\xi}_\alpha^P, \boldsymbol{\xi}_\beta^Q\right) \int_P \phi(\mathbf{x})\, L_\alpha^P(\mathbf{x})\ d\mathbf{x} \int_Q \psi(\mathbf{y})\, L_\beta^Q(\mathbf{y})\ d\mathbf{y}.$$

- Decouples ϕ and ψ, "sparsifies" off-diagonal matrix blocks.
- Replaces subblock of $a(\cdot,\cdot)$ with a low rank approximation $U_P \Sigma_{(P,Q)} U_Q^T$ with tall and skinny U_P, U_Q.
- If we stop now, we have constructed a so-called $\mathcal{H}$-matrix approximation:

$$A \approx A_{\text{near}} + A_{\text{far}} = A_{\text{near}} + \sum_{(P,Q)\text{ admissible}} U_P \Sigma_{(P,Q)} U_Q^T.$$

Cluster method – $\mathcal{H}^2$-matrices

For $\boldsymbol{x}$ in a sub-cluster P of Q, i.e. $P \subset Q$,

$$L_\alpha^Q(\boldsymbol{x}) = \sum_{\beta=1}^{m^d} L_\alpha^Q\left(\boldsymbol{\xi}_\beta^P\right) L_\beta^P(\boldsymbol{x}).$$

Need to compute

- Far-field coefficients $\int_P \phi(\boldsymbol{x}) L_\alpha^P(\boldsymbol{x}) \, d\boldsymbol{x}$ only for leaves of the cluster tree,
- shift coefficients $L_\alpha^Q\left(\boldsymbol{\xi}_\beta^P\right)$,
- kernel approximations $\gamma\left(\boldsymbol{\xi}_\alpha^P, \boldsymbol{\xi}_\beta^Q\right)$,
- near-field entries.

$\mathcal{H}^2$-matrix approximation[3][4]

FE assembly and matrix-vector product in $\mathcal{O}\left(n \log^{2d} n\right)$ operations.

- Finite δ: need to be able to form clusters that fit within the horizon.
- Less intrusive but more costly way of computing far-field interactions via entry sampling: Adaptive Cross Approximation (ACA)

[3] Mark Ainsworth and Christian Glusa. "Towards an efficient finite element method for the integral fractional Laplacian on polygonal domains". In: *Contemporary Computational Mathematics-A Celebration of the 80th Birthday of Ian Sloan*. Springer, 2018, pp. 17–57.

[4] Mark Ainsworth and Christian Glusa. "Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver". In: *Computer Methods in Applied Mechanics and Engineering* (2017).

Operator interpolation[5,6]

Parameter learning problem requires operators for different values of s and δ.

- Piecewise Chebyshev interpolation in s:

Lemma

Let $s \in [s_{\min}, s_{\max}] \subset (0,1)$, $\delta \in (0,\infty)$, and let $\eta > 0$. Assume that $u(s) \in H_{\Omega}^{s+1/2-}(\mathbb{R}^n)$, $v \in H_{\Omega}^{s}(\mathbb{R}^n)$. There exists a partition of $[s_{\min}, s_{\max}]$ into sub-intervals $\mathcal{S}_k$ and interpolation orders M_k such that the piecewise Chebyshev interpolant $\tilde{a}(\cdot,\cdot;s,\delta)$ satisfies:

$$\left|a(u(s),v;s,\delta) - \tilde{a}(u(s),v;s,\delta)\right| \leq \eta \left\|u(s)\right\|_{H_{\Omega}^{\bar{s}_2(s)}(\mathbb{R}^n)} \left\|v\right\|_{H_{\Omega}^{s}(\mathbb{R}^n)},$$

and the total number of interpolation nodes satisfies

$$\sum_{k=1}^{K} (M_k + 1) \leq C \left|\log \eta\right|.$$

The constant C depends on δ and $s_{\max}$.

- Combined with hierarchical matrix approach: $\mathcal{O}(n \log^{2d+1} n)$ complexity & memory.
- Also allows to evaluate derivatives wrt s.
- Assembly for different values of δ is achieved by splitting the kernel into infinite horizon, singular part, and δ-dependent regular part.

[5] Olena Burkovska and Max Gunzburger. "Affine approximation of parametrized kernels and model order reduction for nonlocal and fractional Laplace models". In: *SIAM Journal on Numerical Analysis* 58.3 (2020), pp. 1469–1494.

[6] Olena Burkovska, Christian Glusa, and Marta D'Elia. "An optimization-based approach to parameter learning for fractional type nonlocal models". In: *Computers & Mathematics with Applications* (2021).

Conditioning and scalable solvers

- $\mathcal{O}(n \log n)$ matrix-vector product in all cases $\rightarrow$ can explore iterative solvers
- Steady-state:
  - Fractional kernel, $\delta = \infty$[7]: $\kappa(\mathbf{A}) \sim h^{-2s} \sim n^{2s/d}$
  - Fractional kernel, $\delta \leq \delta_0$[8]: $\kappa(\mathbf{A}) \sim \delta^{2s-2} h^{-2s} \sim \delta^{2s-2} n^{2s/d}$
  - Constant kernel, δ finite[8]: $\kappa(\mathbf{A}) \sim \delta^{-2}$
- Time-dependent:
  - $\kappa(\mathbf{M} + \Delta t \mathbf{A}) \sim 1 + \Delta t\, \kappa(\mathbf{A})$
  - Depending on time-stepper and CFL condition, this is well-conditioned for small s, large δ.

- Scalable solver options:
  - Multigrid
    - Geometric (GMG)
    - Algebraic (AMG)
  - Domain decomposition
    - Substructuring
    - Schwarz methods
  - Krylov methods
    The matrix is well-conditioned in the certain parameter regimes, e.g.
    - constant kernel, δ large, or
    - or fractional kernel, s small, δ large.

---

[7] Mark Ainsworth, William McLean, and Thanh Tran. "The conditioning of boundary element equations on locally refined meshes and preconditioning by diagonal scaling". In: *SIAM Journal on Numerical Analysis* 36.6 (1999), pp. 1901–1932.

[8] Burak Aksoylu and Zuhal Unlu. "Conditioning analysis of nonlocal integral operators in fractional Sobolev spaces". In: *SIAM Journal on Numerical Analysis* 52.2 (2014), pp. 653–677.

Geometric multigrid (GMG)

- Hierarchy of meshes from uniform or adaptive refinement
- Restriction / prolongation given by nesting of FE spaces
- Assembly into hierarchical or CSR matrix format on every level
- Smoothers:
  - Jacobi,
  - Chebyshev,
  - Gauss-Seidel when CSR matrix format is used.
- Coarse solve: convert to dense or CSR matrix

Numerical Examples in 2D – Timings

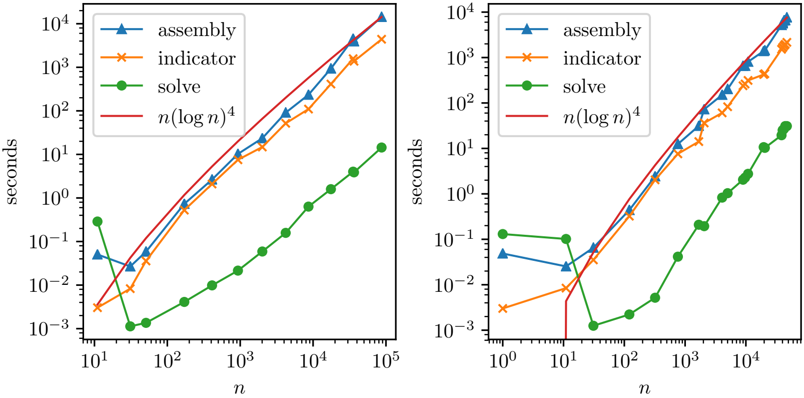
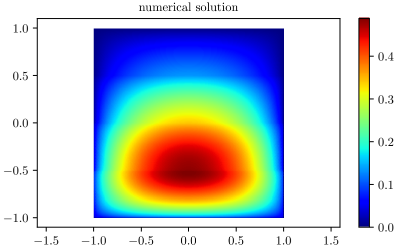
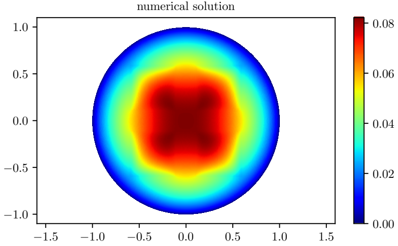


Figure: Timings for assembly of the stiffness matrix for fractional kernels, $\delta = \infty$, solution of linear system using GMG and computation of the error indicators for the two-dimensional problem. $s = 0.25$ on the left, $s = 0.75$ on the right.

Fractional kernel, variable order[9]





$f \equiv 1, \delta = 0.5$

$$s(\mathbf{x}, \mathbf{y}) = \frac{1}{2}(\sigma(\mathbf{x}_1) + \sigma(\mathbf{y}_1))$$

$$\sigma(z) = \begin{cases} 1/5 & \text{if } z < -1/2, \\ 2/5 & \text{if } -1/2 \le z < 0, \\ 3/5 & \text{if } 0 \le z < 1/2, \\ 4/5 & \text{if } 1/2 \le z. \end{cases}$$

$f \equiv 1, \delta = \infty$

$$s(\mathbf{x}, \mathbf{y}) = \begin{cases} 0.25 & \text{if } \mathbf{x}, \mathbf{y} \in \text{islands}, \\ 0.75 & \text{if } \mathbf{x}, \mathbf{y} \notin \text{islands}, \\ 0.75 & \text{else}. \end{cases}$$

[9] Marta D'Elia and Christian A. Glusa. *A fractional model for anomalous diffusion with increased variability. Analysis, algorithms and applications to interface problems*. (Accepted in Numerical Methods for Partial Differential Equations). 2021.

FEM convergence for variable s

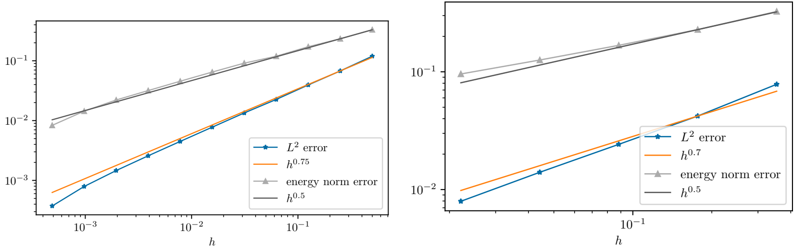


Figure: Convergence in L^2 and energy norm for a 1D example (*left*) and a 2D example with four material layers (*right*).

Rate of convergence, fractional kernels

	$\|\|e\|\|$	$\|\|e\|\|_{L^2}$
constant kernels (literature)	$h^{1/2-\varepsilon}$	$h^{\min\{1,1/2+s\}-\varepsilon}$
variable kernels (observed)	$h^{1/2-\varepsilon}$	$h^{\min\{1,1/2+\underline{s}\}-\varepsilon}$

$\underline{s} = \min s(\boldsymbol{x}, \boldsymbol{y})$

$\Rightarrow$ Possibly straightforward extension of regularity theory?

Solvers for Time-Dependent Problems: CG and GMG

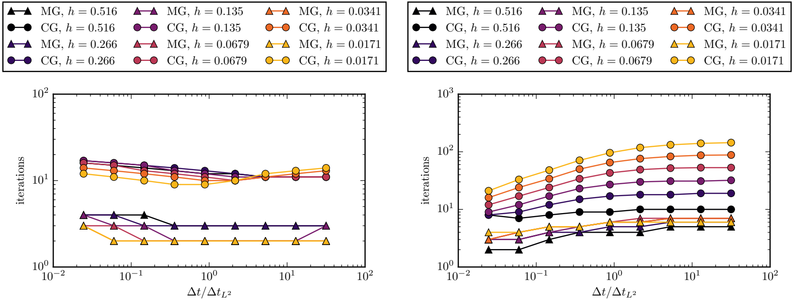


Figure: Fractional kernel. Number of iterations for CG and GMG depending on Δt for $s = 0.25$ (*left*) and $s = 0.75$ (*right*). Δt_{L^2} is the time-step that balances discretisation errors in time and space with respect to the L^2-norm.

Conjugate gradient is a competitive solver when the fractional order s is small and the time step Δt is not too large.

Algebraic multigrid (WIP)

Motivation:

- Adaptively refined / graded meshes can make geometric multigrid painful.
- Use of established algebraic multigrid framework: Trilinos/MueLu
  - Lots of features (more smoothers, coarse solvers, multigrid cycles, etc)
  - Able to handle coefficient and mesh variations
  - Runs on lots of different computing architectures (CPU, threads, GPUs, etc)

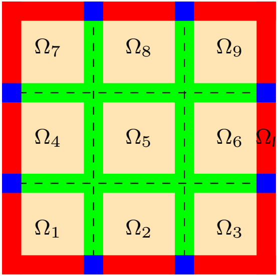
Approach:

- Algebraic multigrid constructs coarse problems using sparsity patterns and matrix entries
  $\rightarrow$ Cannot directly use matrix $\boldsymbol{A}$ when $\delta \gg h$ and hierarchical matrix format is used.
- Construct hierarchy for an auxiliary operator:
  - PDE operators, e.g. $(\nabla u, \nabla v)$,
  - (distance) Graph Laplacian wrt mesh,
  - near field part of hierarchical matrix after some filtering.
- Triple matrix products $\boldsymbol{A}_c = \boldsymbol{RAP}$ where $\boldsymbol{R}$ and $\boldsymbol{P}$ are sparse and $\boldsymbol{A}$ an $\mathcal{H}$- or $\mathcal{H}^2$-matrix
- Recompression of coarse matrix $\boldsymbol{A}_c$

		memory (finest level)		iterations (time)
unknowns	# MPI ranks	dense	$\mathcal{H}^2$	CG+AMG
11,193	4	0.93 GB	0.18 GB	7 (0.22s)
45,169	18	15.2 GB	0.89 GB	9 (0.82s)
181,473	72	245 GB	5.1 GB	15 (2.1s)
727,489	288	3,943 GB	17.8 GB	9 (3.75s)
n	$\sim n$	$\sim n^2$	$\sim n \log^4 n$	constant # iterations?

Table: 2d fractional Poisson problem, $s = 0.75$, $\delta = \infty$, smoothed aggregation

Substructuring[10,11,12]



- Assume $\delta = \mathcal{O}(h)$.
- Cover with overlapping subdomains $\Omega \cup \Omega_I = \bigcup \Omega_i$, diam $(\Omega_i \cap \Omega_j) \sim \delta$ for adjacent subdomains.
- Duplicate unknowns in overlaps:

$$\boldsymbol{A}\boldsymbol{u} = \boldsymbol{f} \Leftrightarrow \begin{pmatrix} \boldsymbol{A}_{\epsilon\epsilon} & \boldsymbol{M}^T \\ \boldsymbol{M} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_\epsilon \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\epsilon \\ 0 \end{pmatrix}$$

  - $\boldsymbol{A}_{\epsilon\epsilon}$ is block diagonal by subdomain, partition-of-unity type scaling included.
  - For floating subdomains, local matrix $\boldsymbol{A}_p$ is singular.
  - $\boldsymbol{M}$ has entries $\{\pm 1, 0\}$, encodes the identity constraints on the overlaps (non-redundant).

---

[10] Giacomo Capodaglio, Marta D'Elia, Pavel Bochev, and Max Gunzburger. "An energy-based coupling approach to nonlocal interface problems". In: *Computers & Fluids* 207 (2020), p. 104593.

[11] Xiao Xu, Christian Glusa, Marta D'Elia, and John T. Foster. "A FETI approach to domain decomposition for meshfree discretizations of nonlocal problems". In: *Computer Methods in Applied Mechanics and Engineering* 387 (2021), p. 114148.

[12] WIP with Bochev, Capodaglio, D'Elia, Gunzburger, Klar, Vollmann

Reduced system and Dirichlet preconditioner

- Let nullspace of $\boldsymbol{A}_{\epsilon\epsilon}$ be given by $\boldsymbol{Z}$.
- Eliminate primal variables from

$$\begin{pmatrix} \boldsymbol{A}_{\epsilon\epsilon} & \boldsymbol{M}^T \\ \boldsymbol{M} & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{u}_\epsilon \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \boldsymbol{f}_\epsilon \\ 0 \end{pmatrix}$$

  and obtain

$$\begin{aligned} \boldsymbol{P}_0 K \lambda &= \boldsymbol{P}_0 (\boldsymbol{M}\boldsymbol{A}_{\epsilon\epsilon}^\dagger \boldsymbol{f}_\epsilon) \\ \boldsymbol{G}^T \boldsymbol{\lambda} &= \boldsymbol{Z}^T \boldsymbol{f}_\epsilon, \end{aligned}$$

  where $\boldsymbol{K} = \boldsymbol{M}\boldsymbol{A}_{\epsilon\epsilon}^\dagger \boldsymbol{M}^T$, $\boldsymbol{G} = \boldsymbol{M}\boldsymbol{Z}$, $\boldsymbol{P}_0 = \boldsymbol{I} - \boldsymbol{G}(\boldsymbol{G}^T\boldsymbol{G})^\dagger \boldsymbol{G}^T$.
- Use projected CG to solve system.
- $\boldsymbol{P}_0$ acts as a "coarse grid".
- Preconditioner for $\boldsymbol{K}$:
  - Let $\boldsymbol{A}_p$, $\boldsymbol{M}_p$ be local parts of $\boldsymbol{A}_{\epsilon\epsilon}$ and $\boldsymbol{M}$.
  - Write $\boldsymbol{K} = \sum_{p=1}^{P} \boldsymbol{M}_p \boldsymbol{A}_p^\dagger \boldsymbol{M}_p^T = \sum_{p=1}^{P} \widetilde{\boldsymbol{M}}_p \boldsymbol{S}_p^\dagger \widetilde{\boldsymbol{M}}_p^T$.
  - Dirichlet preconditioner: $\boldsymbol{Q} = \sum_{p=1}^{P} \widetilde{\boldsymbol{M}}_p \boldsymbol{S}_p \widetilde{\boldsymbol{M}}_p^T$.
- Results shown use Manuel Klar's (U of Trier) assembly code
  `https://gitlab.uni-trier.de/klar/nonlocal-assembly`

Weak scaling – 2D, constant kernel

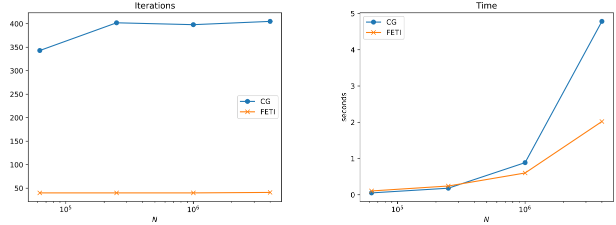


Figure: $\delta = 8e - 3 \rightarrow \kappa \sim$ const

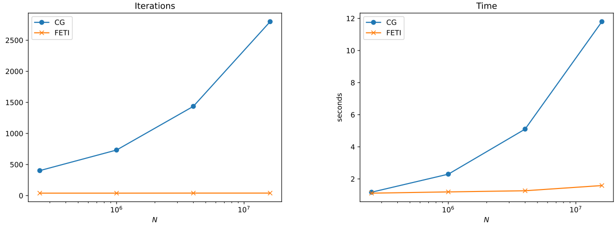


Figure: $\delta = 4h \rightarrow \kappa \sim N$

Weak scaling – 2D, fractional kernel, $s = 0.4$

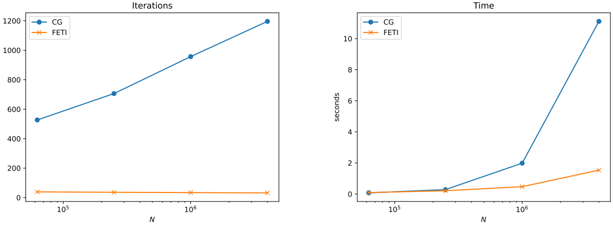


Figure: $\delta = 8e - 3 \rightarrow \kappa \sim N^s$

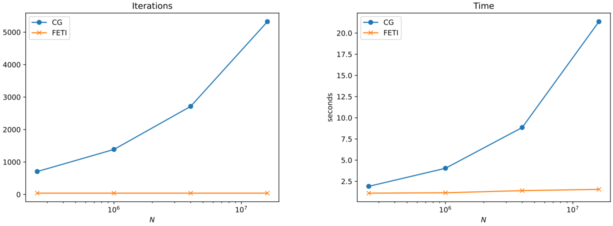


Figure: $\delta = 4h \rightarrow \kappa \sim N$

Strong scaling, 2D

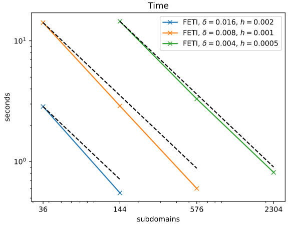


Figure: constant kernel, $\delta = 8h$.

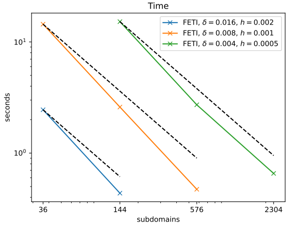


Figure: fractional kernel, $s = 0.4$, $\delta = 8h$.

Schwarz methods (WIP, with Pierre Marchand (INRIA))

- Drawback of substructuring: cannot handle $\delta \gg h$.
- Schwarz method
  - overlapping subdomain restrictions $\{\boldsymbol{R}_p\}$, local matrices $\boldsymbol{A}_p = \boldsymbol{R}_p \boldsymbol{A} \boldsymbol{R}_p^T$
  - partition of unity $\sum_{p=1}^{P} \boldsymbol{R}_p^T \boldsymbol{D}_p \boldsymbol{R}_p = \boldsymbol{I}$, with $\{\boldsymbol{D}_p\}$ diagonal
  - additive Schwarz preconditioner: $\boldsymbol{Q}_1 := \sum_{p=1}^{P} \boldsymbol{R}_p^T \boldsymbol{A}_p^{-1} \boldsymbol{R}_p$, or restricted additive Schwarz
- No global information exchange $\rightarrow$ need a coarse grid
- GenEO approach:
  Span coarse space using solutions of subdomain eigenvalue problems
  $\boldsymbol{D}_p \boldsymbol{A}_p \boldsymbol{D}_p \boldsymbol{v}_{p,k} = \lambda_{p,k} \boldsymbol{B}_p \boldsymbol{v}_{p,k}$, where $\boldsymbol{B}_p$ is similar to $\boldsymbol{A}_p$, but assembled over a modified local mesh.
- Distributed $\mathcal{H}$-matrix is built using Pierre Marchand's Htool library
  `https://github.com/htool-ddm/htool`
- HPDDM library for DD and GenEO `https://github.com/hpddm/hpddm`
- 2D fractional Poisson problem, $s = 0.75$, $\delta = \infty$

		memory (finest level)		iterations (time)
unknowns	# MPI ranks	dense	$\mathcal{H}$	GMRES+DD
65,025	72	31.5 GB	5.4 GB	21 (1.34s)
261,121	288	508 GB	12.6 GB	23 (0.96s)
1,046,529	1152	8,160 GB	86 GB	24 (2.4s)

- Caveats:
  - solver setup needs improvement, working on alternative low-rank approximations
  - direct solves (subdomain, coarse) and eigenvalue problems in dense format

Advertisement: PyNucleus, a FEM code for nonlocal problems

- Written in Python, lots of optimized kernels compiled to C via Cython.
- Compatible with NumPy/SciPy
- Simplical meshes in 1D, 2D, (3D); uniform refinement with boundary snapping options
- Mesh (re)partitioning using (PAR)METIS
- Finite Element discretizations: discontinuous P_0, continuous P_1, P_2, P_3
- Assembly of local differential operators
- Lots of solvers (direct, Krylov, simple preconditioners),
  and in particular geometric multigrid
  WIP: AMG (Trilinos/MueLu), DD (Htool&HPDDM)
- MPI distributed computations via mpi4py
- Assembly of the nonlocal operators in weak form:

$$a(u,v) = \frac{1}{2} \iint_{(\Omega\cup\Omega_I)^2} (u(\mathbf{x}) - u(\mathbf{y}))(v(\mathbf{x}) - v(\mathbf{y}))\gamma(\mathbf{x},\mathbf{y})d\mathbf{y}d\mathbf{x}$$

  into
  - CSR sparse matrix ($\delta \sim h$),
  - dense matrix ($\delta \gg h$),
  - $\mathcal{H}^2$ hierarchical matrix ($\delta \gg h$; only tested for fractional kernels)
- For fractional kernels: quadrature orders are tuned for optimal convergence.
- Code: `https://github.com/sandialabs/PyNucleus`
- Documentation and examples: `https://sandialabs.github.io/PyNucleus`

Code example


```python
from PyNucleus import (kernelFactory, nonlocalMeshFactory, dofmapFactory,
                       functionFactory, HOMOGENEOUS_DIRICHLET, solverFactory)

# Infinite horizon fractional kernel
kernel = kernelFactory('fractional', dim=2, s=0.75, horizon=inf)

# Mesh for unit disc, no interaction domain for homogeneous Dirichlet
mesh, _ = nonlocalMeshFactory('disc', kernel=kernel,
                              boundaryCondition=HOMOGENEOUS_DIRICHLET,
                              hTarget=0.15)

dm = dofmapFactory('P1', mesh)                                # P1 finite elements
f = functionFactory('constant', 1.)                           # constant forcing
b = dm.assembleRHS(f)                                         # $\int_\Omega f\phi_i$
A = dm.assembleNonlocal(kernel, matrixFormat='h2')            # $a(\phi_i, \phi_j)$, hierarchical
u = dm.zeros()                                                # solution vector

# solve with diagonally preconditioned CG
solver = solverFactory('cg-jacobi', A=A, setup=True)
solver(b, u)
u.plot()
```

- The documentation contains two examples of how to setup and solve local and nonlocal problems with a lot more explanations.
- The repository contains several drivers that demonstrate some of the code capabilities.

Conclusion

- Discretized fractional equations are dense, but not structurally dense.
  $\rightarrow$ approximation of off-diagonal matrix blocks
- Multigrid and domain decomposition solvers are optimal for nonlocal problems.
- Resulting approaches have essentially the same complexity as PDE case, allow for complex domains.

Thanks for listening!

MATNIP

Funding:
The MATNIP LDRD project (PI: Marta D'Elia) develops for the first time a rigorous nonlocal interface theory based on physical principles that is consistent with the classical theory of partial differential equations when the nonlocality vanishes and is mathematically well-posed. This will improve the predictive capability of nonlocal models and increase their usability at Sandia and, more in general, in the computational-science and engineering community. Furthermore, this theory will provide the groundwork for the development of nonlocal solvers, reducing the burden of prohibitively expensive computations.

References I

[1] Mark Ainsworth and Christian Glusa. “Aspects of an adaptive finite element method for the fractional Laplacian: A priori and a posteriori error estimates, efficient implementation and multigrid solver”. In: *Computer Methods in Applied Mechanics and Engineering* (2017).

[2] Mark Ainsworth and Christian Glusa. “Towards an efficient finite element method for the integral fractional Laplacian on polygonal domains”. In: *Contemporary Computational Mathematics-A Celebration of the 80th Birthday of Ian Sloan*. Springer, 2018.

[3] Mark Ainsworth, William McLean, and Thanh Tran. “The conditioning of boundary element equations on locally refined meshes and preconditioning by diagonal scaling”. In: *SIAM Journal on Numerical Analysis* 36.6 (1999).

[4] Burak Aksoylu and Zuhal Unlu. “Conditioning analysis of nonlocal integral operators in fractional Sobolev spaces”. In: *SIAM Journal on Numerical Analysis* 52.2 (2014).

[5] Olena Burkovska, Christian Glusa, and Marta D’Elia. “An optimization-based approach to parameter learning for fractional type nonlocal models”. In: *Computers & Mathematics with Applications* (2021).

[6] Olena Burkovska and Max Gunzburger. “Affine approximation of parametrized kernels and model order reduction for nonlocal and fractional Laplace models”. In: *SIAM Journal on Numerical Analysis* 58.3 (2020).

[7] Giacomo Capodaglio et al. “An energy-based coupling approach to nonlocal interface problems”. In: *Computers & Fluids* 207 (2020).

References II

[8] Marta D'Elia and Christian A. Glusa. *A fractional model for anomalous diffusion with increased variability. Analysis, algorithms and applications to interface problems*. 2021.

[9] Marta D'Elia, Max Gunzburger, and Christian Vollmann. "A cookbook for approximating Euclidean balls and for quadrature rules in finite element methods for nonlocal problems". In: *Mathematical Models and Methods in Applied Sciences* 31.08 (2021).

[10] Xiao Xu et al. "A FETI approach to domain decomposition for meshfree discretizations of nonlocal problems". In: *Computer Methods in Applied Mechanics and Engineering* 387 (2021).