Learning Nonlocal Constitutive Laws for Heterogeneous Material Modeling

Yue Yu
Department of Mathematics, Lehigh University

Present at CBMS Conference: Deep Learning and Numerical PDEs

June 20, 2023

Outline

- Goal: modeling material responses from data
- Part I: Learning a Linear & Homogenized Model
  - ✓ To Learn: a nonlocal kernel function
- Part II: Learning a Nonlinear & Heterogeneous Model
  - ✓ To Learn: a nonlocal neural constitutive law

Motivation and Background

Goal: prediction and monitoring of material responses

- Prediction and monitoring of material responses from experimental measurements are ubiquitous in applications from different fields, such as mechanical engineering, biomedical engineering, civil engineering, etc.

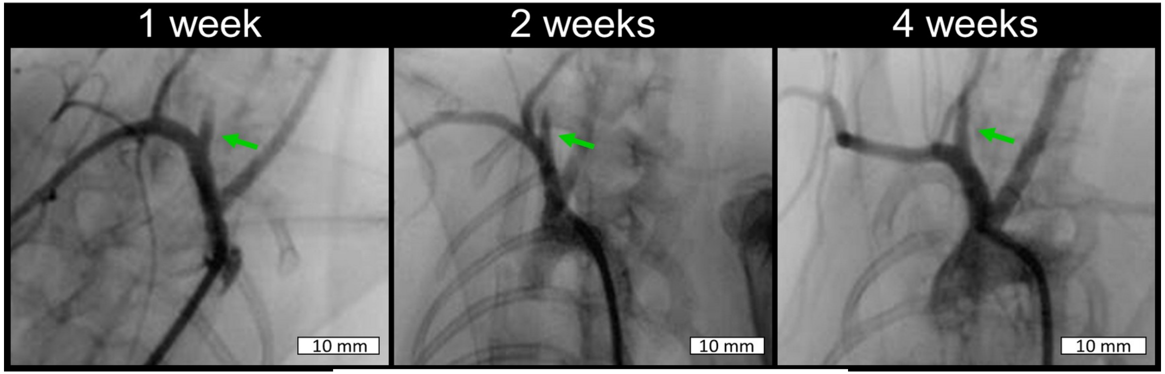
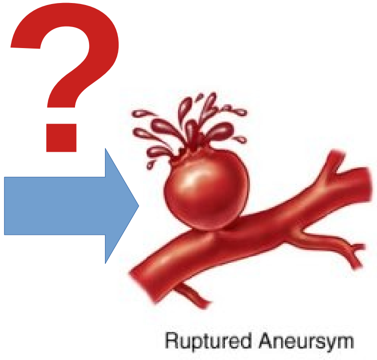
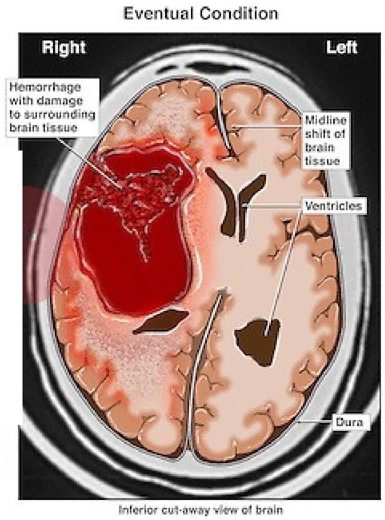


Image by Chung-Hao Lee group





Example 1: monitor aneurysm status and predict the possible hemorrhagic stroke.

Motivation and Background

Goal: prediction and monitoring of material responses

- Prediction and monitoring of material responses from experimental measurements are ubiquitous in applications from different fields, such as mechanical engineering, biomedical engineering, civil engineering, etc.

UAV

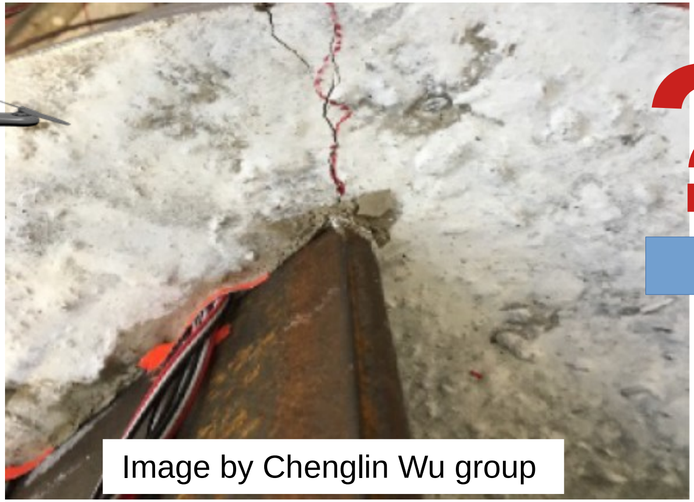


Image by Chenglin Wu group

Image by Francesco Pugliese

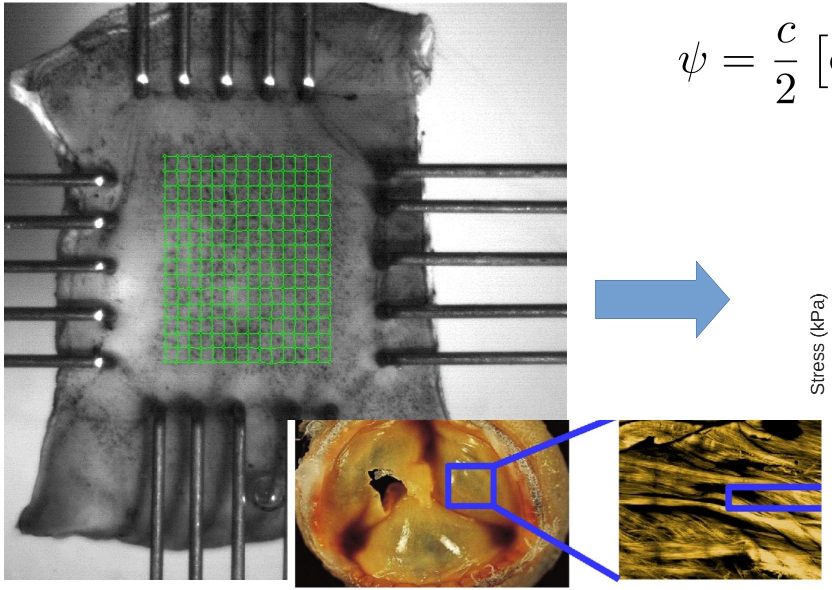
Example 2: monitor crack propagation and corrosion to predict the bridge serving life.

Motivation and Background

Goal: prediction and monitoring of material responses

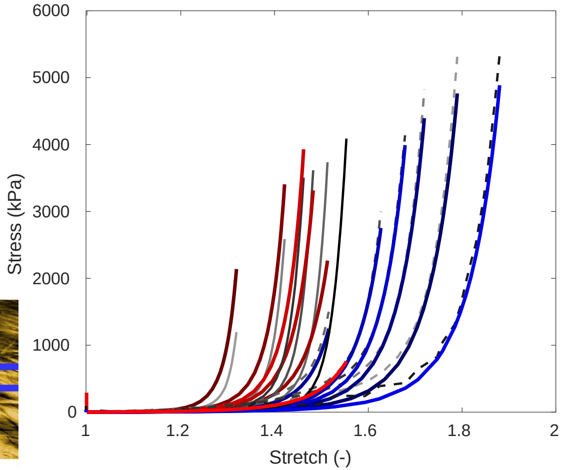
- In materials, small-scale dynamics and interactions affect the global behavior.
- The constitutive law is generally unknown, making the model calibration and validation challenging.

Step 1: Data collection (mechanical Testing of heart valve leaflet)

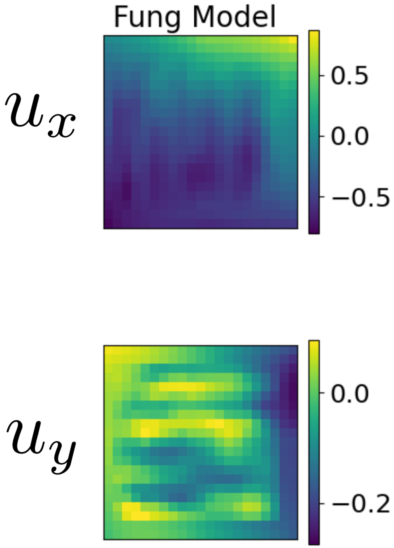


Step 2: Model selection and parameter fitting

$$\psi = \frac{c}{2}\left[\exp(a_1 E_{11}^2 + a_2 E_{22}^2 + 2a_3 E_{11} E_{22}) - 1\right]$$



Step 3: Prediction by solving PDEs



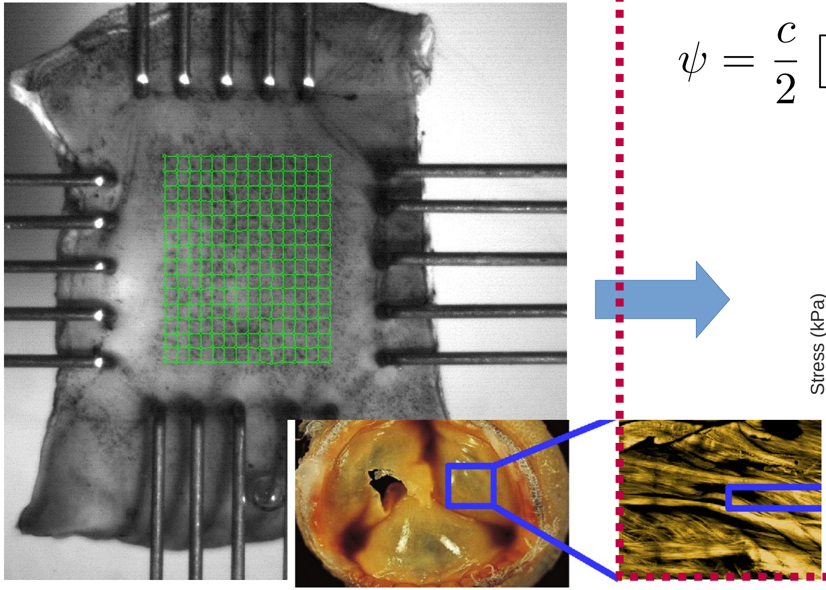
Motivation and

To learn:
a nonlocal constitutive law from experimental measurements

Goal: prediction and monitoring of materia... nses

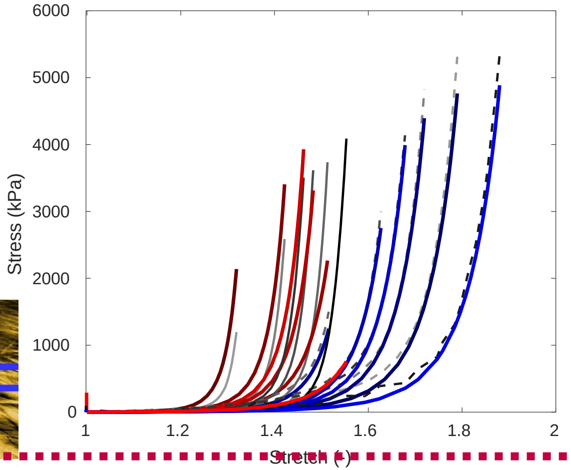
- In materials, small-scale dynamics and interactions affect th... l behavior.
- The constitutive law is generally unknown, making the model ... bration and validation challenging.

Step 1: Data collection (mechanical Testing of heart valve leaflet)

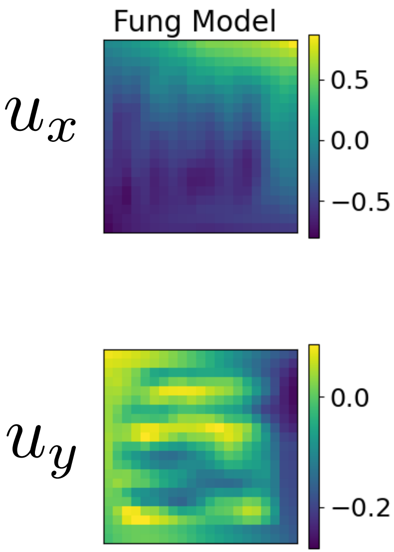


Step 2: Model selection and parameter fitting

$$\psi = \frac{c}{2}\left[\exp(a_1 E_{11}^2 + a_2 E_{22}^2 + 2a_3 E_{11}E_{22}) - 1\right]$$

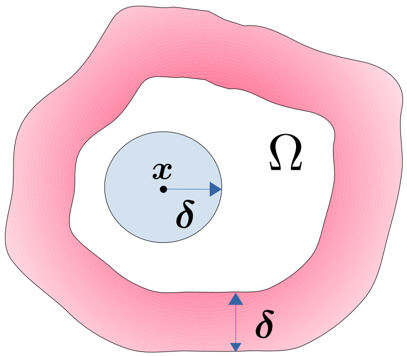


Step 3: Prediction by solving PDEs



What is nonlocal model?

Basic concepts:

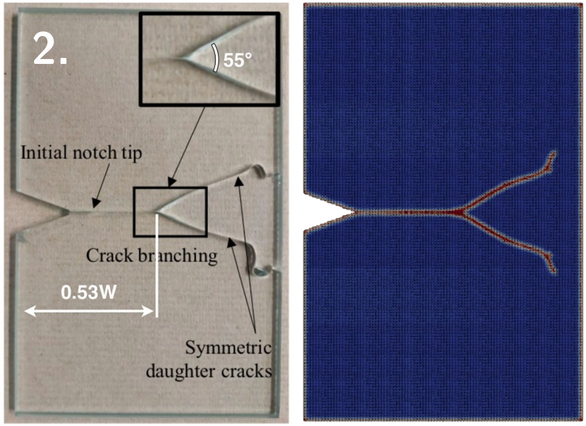


- The state of a system at any point depends on the state in a **neighborhood** of points
- Interactions can occur **at distance, without contact**
- Solutions can be irregular: non-differentiable, singular, discontinuous

Facts:

These models can capture effects that traditional PDEs **hard to capture**

1) Multiscale behavior (*nonlocal as an upscaled/homogenized model*)
2) Discontinuities such as cracks and fractures
3) Anomalous behavior such as superdiffusion and subdiffusion *(fractional operators)*

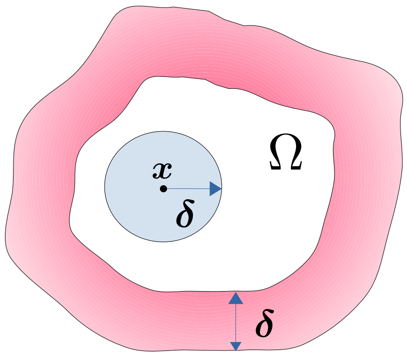


Glass fracture simulation, Yu et al. [2021]

What is nonlocal model?

Basic concepts:

- The state of a system at any point depends on the state in a **neighborhood** of points
- Interactions can occur **at distance, without contact**
- Solutions can be irregular: non-differentiable, singular, discontinuous

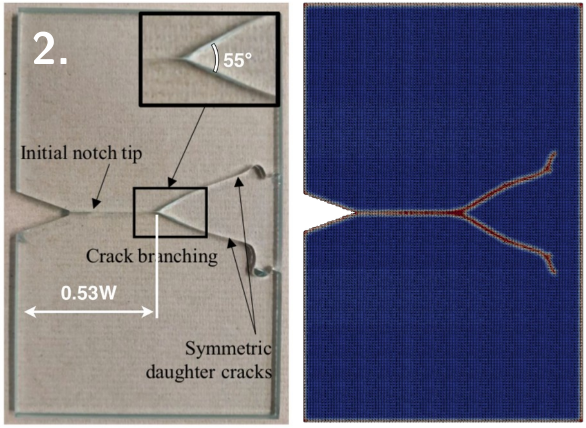


A general nonlocal mechanical (peridynamics) model:

$$\rho(\mathbf{x})\ddot{\mathbf{u}}(\mathbf{x},t) = \int_{B_\delta(\mathbf{x})} \boxed{\mathbf{g}(\mathbf{y},\mathbf{x},\mathbf{u},t)\; dy} + \mathbf{f}(\mathbf{x},t)$$

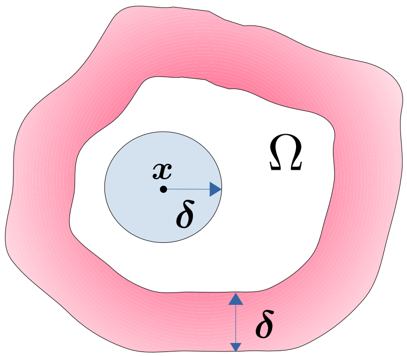
The integrants depend on material properties, microstructure, etc

Glass fracture simulation, Yu et al. [2021]



S. A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces, Journal of the Mechanics and Physics of Solids 48 (1) (2000) 175–209.

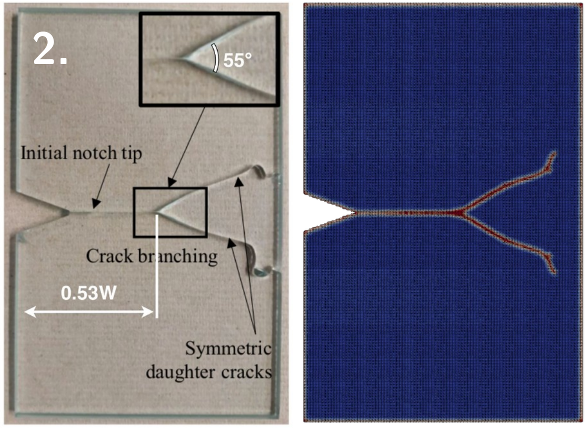
What is nonlocal model?



Basic concepts:

- The state of a system at any point depends on the state in a **neighborhood** of points
- Interactions can occur **at distance, without contact**
- Solutions can be irregular: non-differentiable, singular, discontinuous

Glass fracture simulation, Yu et al. [2021]



A general nonlocal mechanical (peridynamics) model:

$$\rho(\mathbf{x})\ddot{\mathbf{u}}(\mathbf{x},t) = \int_{B_\delta(\mathbf{x})} \boxed{\mathbf{g}(\mathbf{y},\mathbf{x},\mathbf{u},t)\; d\mathbf{y}} + \mathbf{f}(\mathbf{x},t)$$

Learn the integrants from data pairs

$$\{\mathbf{u}_i(\mathbf{x},t), \mathbf{f}_i(\mathbf{x},t)\}_{i=1}^N$$

S. A. Silling, Reformulation of elasticity theory for discontinuities and long-range forces, Journal of the Mechanics and Physics of Solids 48 (1) (2000) 175–209.

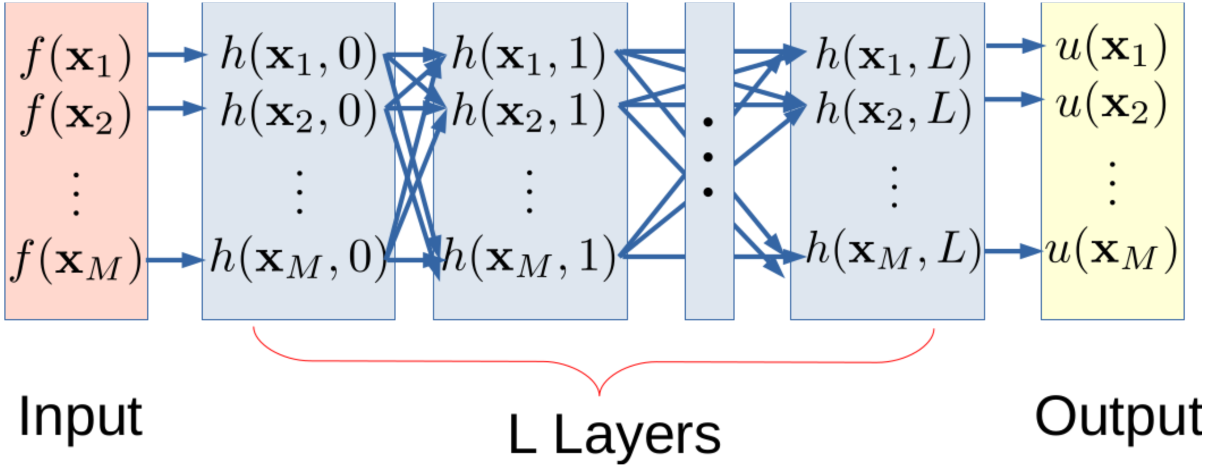
What is nonlocal model?

Goal: learn nonlocal constitutive laws for material modeling

- Desired properties: 1. the learnt model should be **generalizable to future prediction tasks**.
  2. the inverse problem should also be **well-posed and resolution independent**.

$$\{b_1(\mathbf{x}_i), f_1(\mathbf{x}_i), u_1(\mathbf{x}_i)\}$$
$$\{b_2(\mathbf{x}_i), f_2(\mathbf{x}_i), u_2(\mathbf{x}_i)\}$$
$$\cdots$$
$$\{b_N(\mathbf{x}_i), f_N(\mathbf{x}_i), u_N(\mathbf{x}_i)\}$$

Training Samples

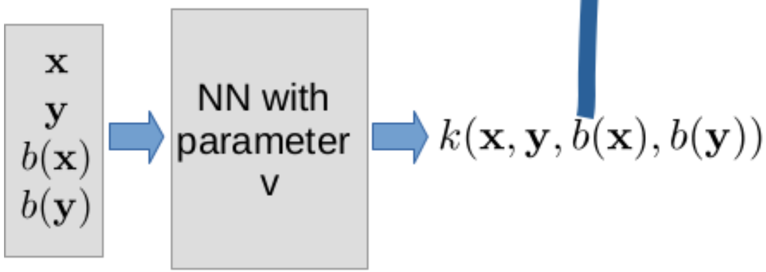


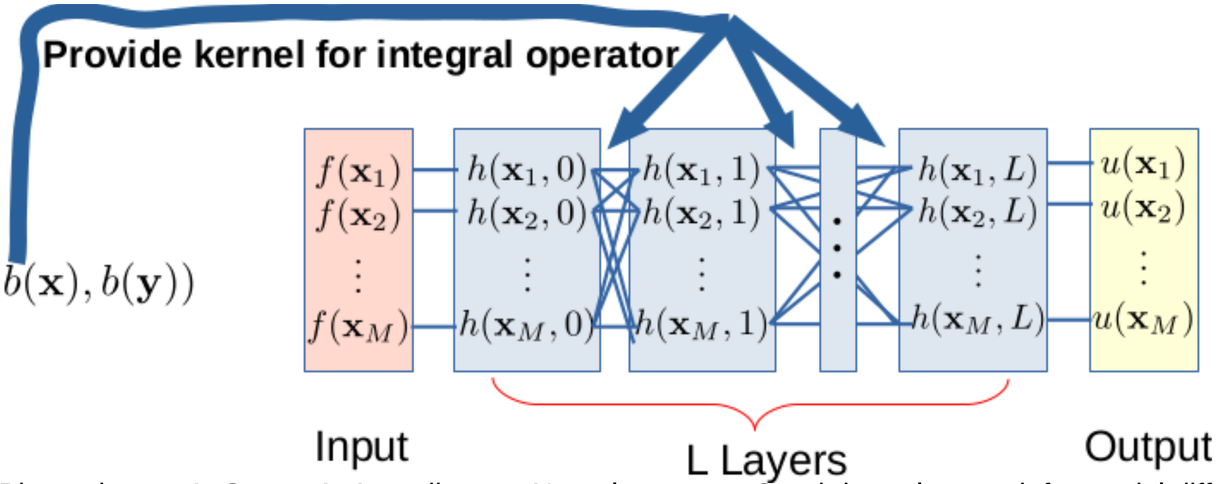
What is nonlocal model?

Goal: learn nonlocal constitutive laws for material modeling

- Desired properties: 1. the learnt model should be **generalizable to future prediction tasks**.
  2. the inverse problem should also be **well-posed and resolution independent**.

Propose: Learning the nonlocal kernel/integrant!





[1]Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, Neural operator: Graph kernel network for partial differential equations, arXiv preprint arXiv:2003.03485.
[2]H. You, Y. Yu, N. Trask, M. Gulian, M. D'Elia, "Data-driven learning of nonlocal physics from high-fidelity synthetic data", Computer Methods in Applied Mechanics and Engineering, Volume 374, 113553, 2021.

Part I

Learning Nonlocal Kernel for Homogenized Models

[1] H. You, Y. Yu*, S. Silling, M. D'Elia, "A data-driven peridynamic continuum model for upscaling molecular dynamics". CMAME, 2022.

[2] F. Lu, Q. An, Y. Yu*, "Nonparametric learning of kernels in nonlocal operators". Submitted.

[3] H. You, Y. Yu, S. Silling, M. D'Elia, "Data-driven learning of nonlocal models: from high-fidelity simulations to constitutive laws". AAAI Spring Symposium: MLPS, 2021

[4] H. You, Y. Yu, N. Trask, M. Gulian, M. D'Elia, "Data-driven learning of nonlocal physics from high-fidelity synthetic data", CMAME, 2021.

[5] H. You, L. Zhang, Y. Yu, "A meta-learnt nonlocal operator regression approach for metamaterial modeling". MRS Communications, 2022.

[6] Fan Y., D'Elia M, Yu Y, Najm H., Silling S. "Bayesian Nonlocal Operator Regression (BNOR): A Data-Driven Learning Framework of Nonlocal Models with Uncertainty Quantification". Submitted, 2022

Nonlocal Operator Regression (NOR)

Propose: a linear nonlocal constitutive law for homogenization

- **Goal:** identify a nonlocal kernel k in $\mathcal{L}_K u(x) = \int_{B_\delta(x)} (u(y) - u(x)) k(x, y; \mu) dy$

$$\begin{cases} \ddot{u} = \mathcal{L}u + g & \text{in } \Omega \\ u = u_{bc} & \text{on the nonlocal boundary} \end{cases}$$

here, $f := \ddot{u} - g$

1) **Collect measurements** of solution and forcing term: $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$

training set: measurements or high fidelity simulations

2) **Approximate the kernel** with a parameterization: $k(x, y) = \sum_{m=1}^{M} c_m \phi_m(|x - y|)$

3) **Minimize the residual** $\mathcal{E}_\lambda(k) = \frac{1}{N} \sum_{i=1}^{N} \|L_k[u_i] - f_i\|_{L^2}^2 + \lambda \mathcal{R}(k)$

subject to solvability and physical constraints.

outcome: coefficients c_m

Step forward towards learning constitutive behavior of heterogeneous materials

→ **Decrease reliance on lab testing.**

Nonlocal Operator Regression (NOR)

Propose: a linear nonlocal constitutive law for homogenization

- **Goal:** identify a nonlocal kernel k in $\mathcal{L}_K u(x) = \int_{B_\delta(x)} (u(y) - u(x)) k(x, y; \mu) dy$

1) Collect measurements of solution and forcing term: $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$

training set: measurements or high fidelity simulations

2) Approximate the kernel with a parameterization: $k(x, y) = \sum_{m=1}^{M} c_m \phi_m(|x - y|)$

3) Minimize the residual $\mathcal{E}_\lambda(k) = \frac{1}{N} \sum_{i=1}^{N} \|L_k[u_i] - f_i\|_{L^2}^2 + \lambda \mathcal{R}(k)$

subject to solvability and physical constraints.

outcome: coefficients c_m

Key Algorithm Features/Contributions:

- One can selects a set of basis functions for a hypothesis space.
- Learns the functional form of the kernel (previous works only identify discrete parameters!).

➡ **Resolution independent Estimator (Kernel k)**

Lu, F., An, Q., & Yu, Y. (2022). Nonparametric learning of kernels in nonlocal operators. arXiv preprint arXiv:2205.11006.

Nonlocal Operator Regression (NOR)

Propose: a linear nonlocal constitutive law for homogenization

- **Goal:** identify a nonlocal kernel k in $\mathcal{L}_K u(x) = \int_{B_\delta(x)} (u(y) - u(x)) k(x, y; \mu) dy$

1) Collect measurements of solution and forcing term: $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$
2) Approximate the kernel with a parameterization: $k(x, y) = \sum_{m=1}^{M} c_m \phi_m(|x - y|)$
3) Minimize the residual $\mathcal{E}_\lambda(k) = \frac{1}{N} \sum_{i=1}^{N} \|L_k[u_i] - f_i\|_{L^2}^2 + \lambda \mathcal{R}(k)$

   subject to solvability and physical constraints.

training set: measurements or high fidelity simulations

outcome: coefficients c_m

Key Algorithm Features/Contributions:

- The linear model form guarantees physical laws (e.g., linear/angular momentum conservation)
- Constraints can be applied to **guarantee that the resultant surrogate model is well-posed.**

→ **Generabizable to Different Prediction Tasks**

H. You, Y. Yu, S. Silling, M. D'Elia, "Data-driven learning of nonlocal models: from high-fidelity simulations to constitutive laws". AAAI: MLPS, 2021

Nonlocal Operator Regression (NOR)

Propose: a linear nonlocal constitutive law for homogenization

- **Goal:** identify a nonlocal kernel k in $\mathcal{L}_K u(x) = \int_{B_\delta(x)} (u(y) - u(x)) k(x, y; \mu) dy$

1) Collect measurements of solution and forcing term: $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$

training set: measurements or high fidelity simulations

2) Approximate the kernel with a parameterization: $k(x, y) = \sum_{m=1}^{M} c_m \phi_m(|x - y|)$

3) Minimize the residual $\mathcal{E}_\lambda(k) = \frac{1}{N} \sum_{i=1}^{N} \|L_k[u_i] - f_i\|_{L^2}^2 + \lambda \mathcal{R}(k)$

subject to solvability and physical constraints.

outcome: coefficients c_m

Key Algorithm Features/Contributions:

- A regularization term is often necessary, to guarantee that we can find the unique minimizer in the function space of identifiability (FSOI) as $\Delta x \to 0$ and noise reduces.

Identifiability and Robustness to Noise

Lu, F., An, Q., & Yu, Y. (2022). Nonparametric learning of kernels in nonlocal operators. arXiv preprint arXiv:2205.11006.

NOR: Convergence and Robustness to Noise

- **Training set:** $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$, generated from the nonlocal equation $\mathcal{L}_K u(x) = f(x)$ where $\mathcal{L}_K$ is associated to a manufactured kernel $k_{true}(x,y) := k_{true}(|x-y|)$

- **Manufactured kernel:** $k_{true}(r) = c_{d,s} \frac{1}{r^{d+2s}} \mathbf{1}_{[0.1,6]}(x) + \frac{1}{0.1^{d+2s}} \mathbf{1}_{[0,0.1]}(x)$ where $d = 1, s = 0.5.$

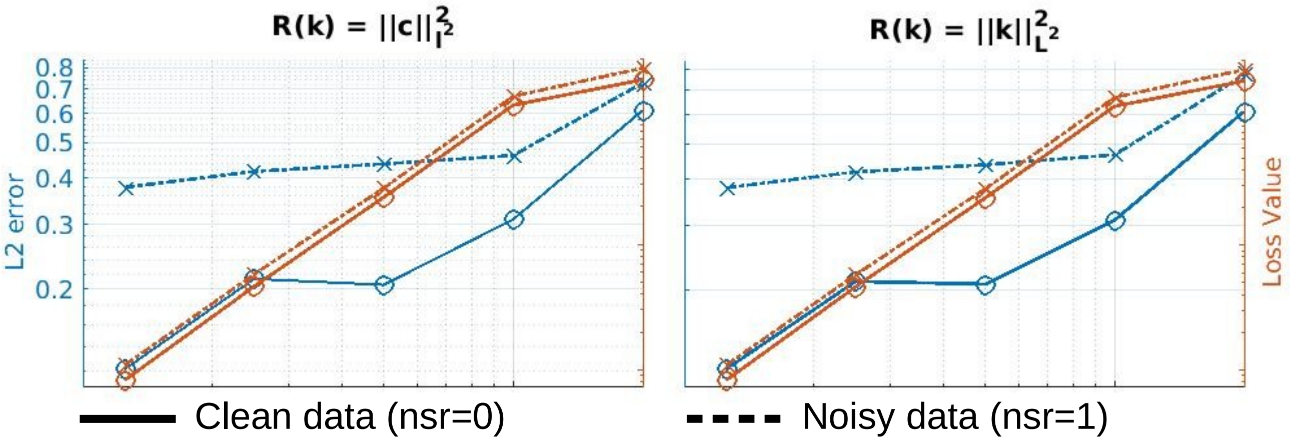
- **Optimization-based learning:** $\min_{c_m} \frac{\Delta x}{N} \sum_{i=1}^{N} \sum_{j=1}^{J} |L_k[u_i](x_j) - f_i(x_j)|^2 + \lambda \mathcal{R}(k)$

where k is approximated by B-splines: $k(x,y) = k(|x-y|) = k(r) = \sum_{m=1}^{M} c_m \phi_m(r)$

When taking the classical Tikhonov regularization:

$\mathcal{R}(k) = \|c\|_{l^2}^2$ or $\mathcal{R}(k) = \|k\|_{L^2}^2$

Convergence of function estimator as the data mesh-size Δx decreases from 0.2 to 0.0125:



NOR: Convergence and Robustness to Noise

- **Training set:** $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$, generated from the nonlocal equation $\mathcal{L}_K u(x) = f(x)$ where $\mathcal{L}_K$ is associated to a manufactured kernel $k_{true}(x,y) := k_{true}(|x-y|)$

Theorem (Function space of identifiability) [Lu, An, Yu, 2022]:
Consider the problem of identifying the kernel k, the function space of identifiability, in which the true kernel is the unique minimizer of the loss functional, is an RKHS (denoted by H_G) with reproducing kernel:

$$\bar{G}(r,s) = \frac{G(r,s)}{\rho'_N(r)\rho'_N(s)}, \text{ where } G(r,s) = \frac{1}{N}\sum_{i=1}^N \int_{|\eta|=1}\int_{|\xi|=1}\left[\int [u_i(x+r\xi)-u_i(x)][u_i(x+s\eta)-u_i(x)]dx\right] d\xi d\eta$$

where ρ'_N is the density of an empirical probability density $\rho_N(dr) = \frac{1}{ZN}\sum_{i=1}^N \int_\Omega \int_\Omega \delta_{|x-y|}(r)|u_i(x)-u_i(y)|dxdy$.

Theorem (Characterization of the RKHS space) [Lu, An, Yu, 2022]:
The RKHS H_G with $\bar{G}$ as reproducing kernel satisfies $H_G = \mathcal{L}_{\bar{G}}^{1/2}(L^2(\rho_N))$, where $L_{\bar{G}}$ is an integral operator defined by

$$\mathcal{L}_{\bar{G}} k(r) = \int_0^\infty k(s)\bar{G}(r,s)\rho_N(s)ds$$

The eigenvalues of $L_{\bar{G}}$ converges to zero, and its eigen-functions $\{\psi_l(r)\}$ can form a complete orthonormal basis of $L^2(\rho_N)$. The optimal kernel satisfies: $\hat{k} = \mathcal{L}_{\bar{G}}^{-1} P k_N^f$.

Lu, F., An, Q., & Yu, Y. (2022). Nonparametric learning of kernels in nonlocal operators. arXiv preprint arXiv:2205.11006.

NOR: Convergence and R

Two fundamental challenges:

1. The inverse problem is well-defined, but only in the function space of identifiability.
2. Outside the function space of identifiability, it is ill-posed.

- **Training set:** $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^N$, generated from t
where $\mathcal{L}_K$ is associated to a manufactured kernel k_{t}

Theorem (Function space of identifiability) [Lu, An
Consider the problem of identifying the kernel k, the f
true kernel is the unique minimizer of the loss functio
reproducing kernel:

$$\bar{G}(r,s) = \frac{G(r,s)}{\rho'_N(r)\rho'_N(s)}, \text{ where } G(r,s) = \frac{1}{N}\sum_{i=1}^{N}\int_{|\eta|=1}\int_{|\xi|=1}\left[\int [u_i(x+r \ldots (x+s\eta)-u_i(x)]dx\right] d\xi d\eta$$

where ρ'_N is the density of an empirical probability density $\rho_N(\ldots \frac{1}{N}\sum_{i=1}^{N}\int_\Omega\int_\Omega \delta_{|x-y|}(r)|u_i(x)-u_i(y)|dxdy$.

Theorem (Characterization of the RKHS space) [Lu, An, Yu 2022]:
The RKHS H_G with $\bar{G}$ as reproducing kernel satisfies $H_G = \mathcal{L}_{\bar{G}}^{1/2}(L^2(\rho_N))$, where $L_{\bar{G}}$ is an integral operator defined by

$$\mathcal{L}_{\bar{G}}k(r) = \int_0^\infty k(s)\bar{G}(r,s)\rho_N(s)ds$$

The eigenvalues of $L_{\bar{G}}$ converges to zero, and its eigen-functions $\{\psi_l(r)\}$ can form a complete orthonormal basis of $L^2(\rho_N)$. The optimal kernel satisfies: $\hat{k} = \mathcal{L}_{\bar{G}}^{-1} P k_N^f$.

Lu, F., An, Q., & Yu, Y. (2022). Nonparametric learning of kernels in nonlocal operators. arXiv preprint arXiv:2205.11006.

NOR: Convergence and Robustness to Noise

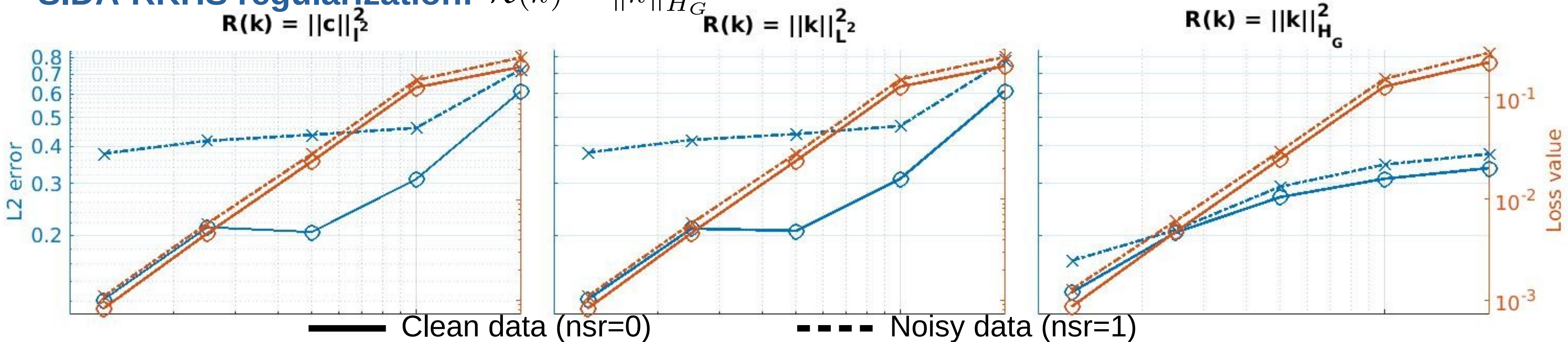
- **Training set:** $\mathcal{D} = \{u_i(x), f_i(x)\}_{i=1}^{N}$, generated from the nonlocal equation $\mathcal{L}_K u(x) = f(x)$ where $\mathcal{L}_K$ is associated to a manufactured kernel $k_{true}(x,y) := k_{true}(|x-y|)$

- **Manufactured kernel:** $k_{true}(r) = c_{d,s}\frac{1}{r^{d+2s}}\mathbf{1}_{[0.1,6]}(x) + \frac{1}{0.1^{d+2s}}\mathbf{1}_{[0,0.1]}(x)$ where $d=1, s=0.5.$

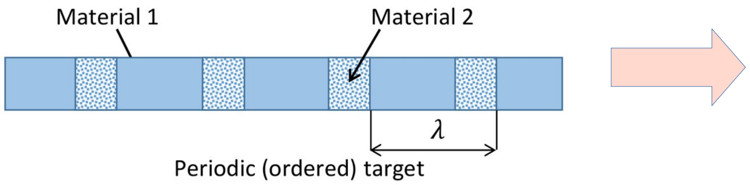
- **Optimization-based learning:** $\min_{c_m} \frac{\Delta x}{N}\sum_{i=1}^{N}\sum_{j=1}^{J} |L_k[u_i](x_j) - f_i(x_j)|^2 + \lambda\mathcal{R}(k)$

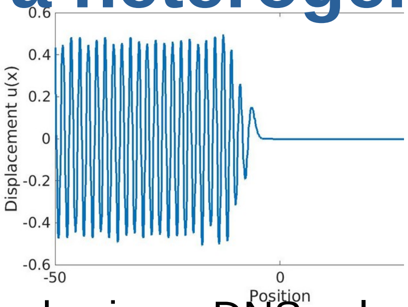
  where k is approximated by B-splines: $k(x,y) = k(|x-y|) = k(r) = \sum_{m=1}^{M} c_m\phi_m(r)$

- **SIDA-RKHS regularization:** $\mathcal{R}(k) = \|k\|_{H_G}^2$



NOR: Wave propagation in a heterogeneous bar

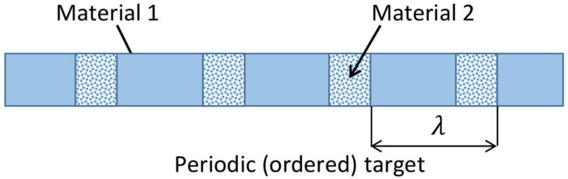


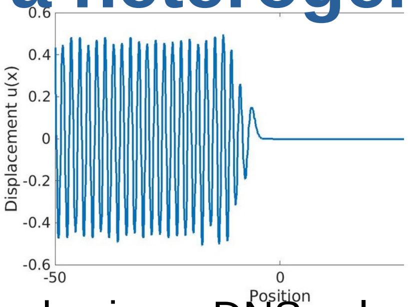


- **Training set:** oscillating source and plane wave obtained using a DNS solver that computes the velocity exactly, with t from 0 to 2.
  Oscillating source: $\Omega = [-50, 50]$, $g(x,t) = \exp^{-(\frac{2x}{5jL})^2} \exp^{-(\frac{t-0.8}{0.8})^2} \cos^2(\frac{2\pi x}{jL})$, for $j = 1, 2, \cdots, 20.$
  Plane wave 1: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \cos(jt)$ for $j = 0.35, 0.7, \cdots, 3.85.$
  Plane wave 2: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \sin(jt)$ for $j = 0.35, 0.7, \cdots, 3.85.$

- **Experiments:**
  Coarse data set 1: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 1.
  Coarse data set 2: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 2.
  Fine data set: we train the estimator using ``fine'' dataset (Δx=0.025) of oscillating source and plane wave 1.

NOR: Wave propagation in a heterogeneous bar



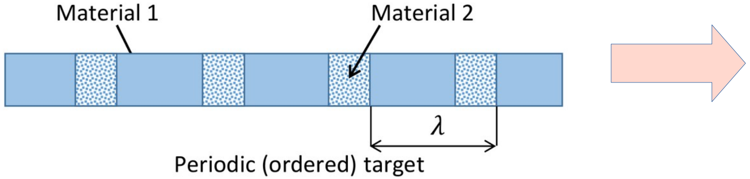


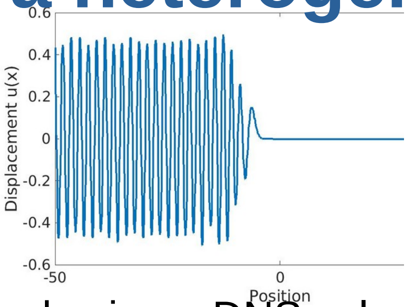
- **Training set:** oscillating source and plane wave obtained using a DNS solver that computes the velocity exactly, with t from 0 to 2.
  Oscillating source: $\Omega = [-50, 50]$, $g(x,t) = \exp^{-(\frac{2x}{5jL})^2} \exp^{-(\frac{t-0.8}{0.8})^2} \cos^2(\frac{2\pi x}{jL})$, for $j = 1, 2, \cdots, 20$.
  Plane wave 1: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \cos(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.
  Plane wave 2: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \sin(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.

- **Experiments:**
  Coarse data set 1: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 1.
  Coarse data set 2: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 2.
  Fine data set: we train the estimator using ``fine'' dataset (Δx=0.025) of oscillating source and plane wave 1.

investigate the **sensitivity** of the inverse problem.

NOR: Wave propagation in a heterogeneous bar





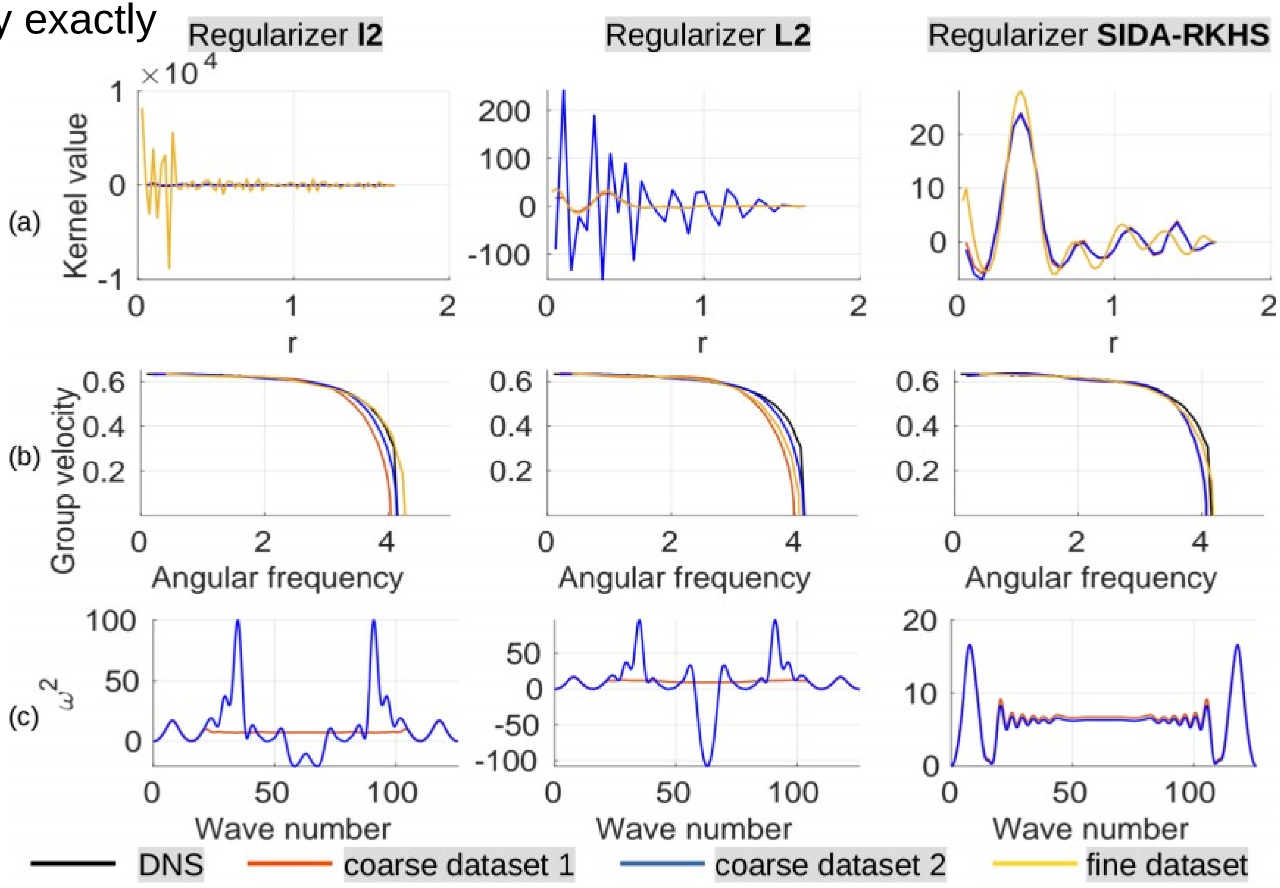
- **Training set:** oscillating source and plane wave obtained using a DNS solver that computes the velocity exactly, with t from 0 to 2.
  Oscillating source: $\Omega = [-50, 50]$, $g(x,t) = \exp^{-(\frac{2x}{5jL})^2} \exp^{-(\frac{t-0.8}{0.8})^2} \cos^2(\frac{2\pi x}{jL})$, for $j = 1, 2, \cdots, 20$.
  Plane wave 1: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \cos(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.
  Plane wave 2: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \sin(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.

- **Experiments:**
  Coarse data set 1: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 1.
  Coarse data set 2: we train the estimator using ``coarse'' dataset (Δx=0.05) of oscillating source and plane wave 2.
  Fine data set: we train the estimator using ``fine'' dataset (Δx=0.025) of oscillating source and plane wave 1.

investigate the **convergence** of the inverse problem.

NOR: Wave propagation in a heterogeneous bar

- **Training set:** oscillating source and plane wave obtained using a DNS solver that computes the velocity exactly



Kernel convergence

matching DNS indicates physical consistency

>0 indicates physical stability

NOR: Wave propagation in a heterogeneous bar

- **Training set:** oscillating source and plane wave obtained using a DNS solver that computes the velocity exactly, with t from 0 to 2.
  Oscillating source: $\Omega = [-50, 50]$, $g(x,t) = \exp^{-(\frac{2x}{5jL})^2} \exp^{-(\frac{t-0.8}{0.8})^2} \cos^2(\frac{2\pi x}{jL})$, for $j = 1, 2, \cdots, 20$.
  Plane wave 1: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \cos(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.
  Plane wave 2: $\Omega = [-50, 50]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-50,t) = \sin(jt)$ for $j = 0.35, 0.7, \cdots, 3.85$.

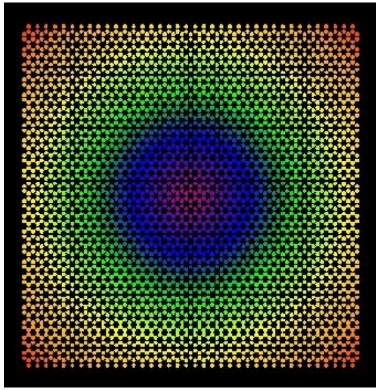
- **Test set:** wave packet obtained using a DNS solver with a different loading and domain, from the training dataset, and with a much longer simulation time (t from 0 to 100).
  Wave packet: $\Omega = [-133.3, 133.3]$, $g(x,t) = 0$, $u(x,0) = 0$, $v(-133.3,t) = \sin(jt)\exp\left(-(t/5-3)^2\right)$, for $j = 1, 2, 3$.

The relative L2 errors of long term (T=100) displacement prediction on the test dataset:

Resolution	l2	L2	SIDA-RKHS
Coarse ($\Delta x = 0.05$)	23.5%	28.4%	**21.8%**
Fine ($\Delta x = 0.025$)	INF	23.4%	**19.2%**

NOR: Coarse-grained MD model for graphene

- Given: a collection of samples of coarse-grained MD displacements and forcing $\{(\mathrm{u}_i, \mathbf{f}_i)\}_{i=1}^N$

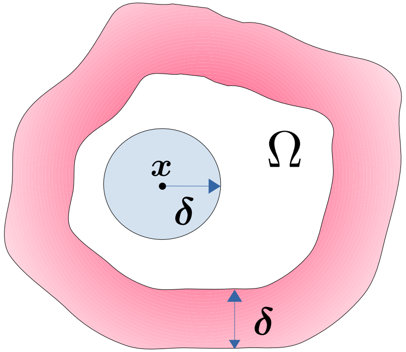


- Model: linearized peridynamic solid (LPS) model

$$\mathcal{L}_\delta \mathbf{u} := -\frac{C_\alpha}{m(\delta)} \int_{B_\delta(\mathbf{x})} (\lambda - \mu) K(|\mathbf{y}-\mathbf{x}|)(\mathbf{y}-\mathbf{x})(\theta(\mathbf{x}) + \theta(\mathbf{y}))\, d\mathbf{y}$$
$$- \frac{C_\beta}{m(\delta)} \int_{B_\delta(\mathbf{x})} \mu K(|\mathbf{y}-\mathbf{x}|) \frac{(\mathbf{y}-\mathbf{x}) \otimes (\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2} (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y} = \mathbf{f}(\mathbf{x}) \quad \mathbf{x} \in \Omega,$$
$$\theta(\mathbf{x}) := \frac{d}{m(\delta)} \int_{B_\delta(\mathbf{x})} K(|\mathbf{y}-\mathbf{x}|)(\mathbf{y}-\mathbf{x}) \cdot (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y}, \quad \mathbf{x} \in \Omega,$$
$$\mathcal{B}_I \mathbf{u}(\mathbf{x}) = \mathbf{q}(\mathbf{x}) \quad \mathbf{x} \in \Omega_I.$$

where the kernel K is approximated by Bernstein polynomials:

$$\mathrm{K}(|\mathbf{y}-\mathbf{x}|) = \sum_{m=0}^{M} \frac{C_m}{\delta^{d+2-\alpha} |\mathbf{y}-\mathbf{x}|^\alpha} B_{m,M}\left(\left|\frac{\mathbf{y}-\mathbf{x}}{\delta}\right|\right) \text{ when } |\mathbf{y}-\mathbf{x}| < \delta$$



- Goal: approximate the kernel K(|y-x|), the Youngs modulus E and the Poisson ratio ν subject to solvability constraints.

H. You, Y. Yu*, S. Silling, M. D'Elia, "A data-driven peridynamic continuum model for upscaling molecular dynamics". CMAME, 2022.

NOR: Coarse-grained MD model for graphene

- When the kernel K is non-negative and not too singular, this linearized model is guaranteed to be solvable.
- However, the non-negative assumption is too restricted.
- We numerically discretize the model with the meshfree quadrature rule, then imposed the solvability constraint in a discrete manner:

Theorem (Well-posedness of the discretized nonlocal model):
The discrete nonlocal coercivity and inf-sup conditions are satisfied if

(Coercivity) $\quad eig(A) > 0,$

(Inf-Sup) $\quad eig(BA^{-1}B^t) > 0,$

where A and B are the discrete matrices of the following nonlocal operators

$$A\mathbf{u} \approx -\frac{C_\beta}{m(\delta)} \int_{B_\delta(\mathbf{x})} K(|\mathbf{y}-\mathbf{x}|) \frac{(\mathbf{y}-\mathbf{x}) \otimes (\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2} (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y},$$

$$B\mathbf{u} \approx \frac{d}{m(\delta)} \int_{B_\delta(\mathbf{x})} K(|\mathbf{y}-\mathbf{x}|)(\mathbf{y}-\mathbf{x}) \cdot (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y}.$$

H. You, Y. Yu*, S. Silling, M. D'Elia, "A data-driven peridynamic continuum model for upscaling molecular dynamics". CMAME, 2022.

NOR: Coarse-grained MD model for g

- Perform MD modeling of a perfect graphene sheet under loads with different frequencies for **70 training samples**:

Data set generation: for $(x_1, x_2) \in [-100, 100]^2 \mathring{A}$

solve the MD problem at constant temperature (0 or 300K)

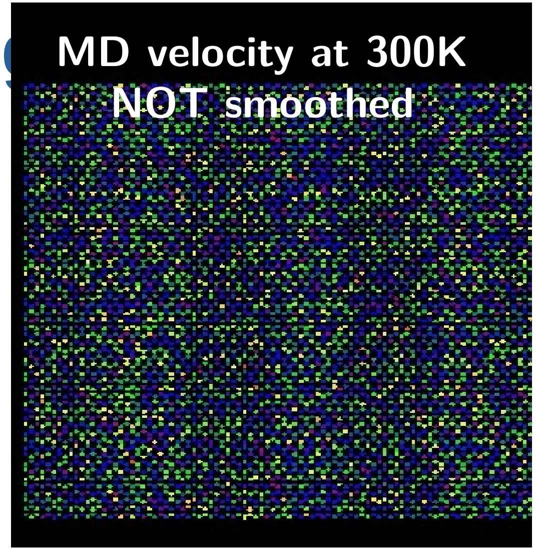
using a "thermostat" with additional external forcing

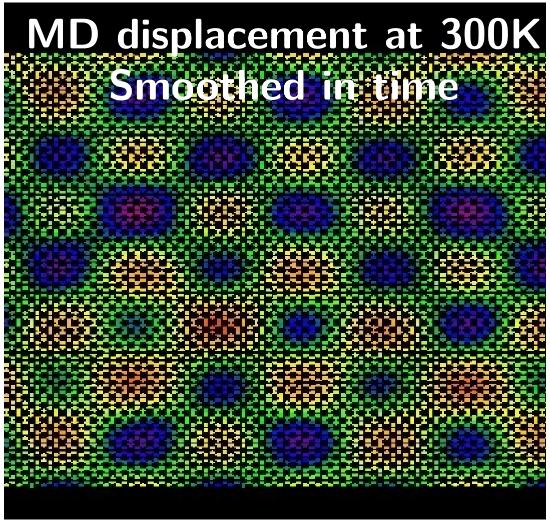
$\mathbf{f}_{k_1,k_2}(x_1, x_2) = (\mathbb{C}\cos(k_1 x_1)\cos(k_2 x_2), 0)$ or

$\mathbf{f}_{k_1,k_2}(x_1, x_2) = (0, \mathbb{C}\cos(k_1 x_1)\cos(k_2 x_2))$ where

- $k_1, k_2 \in \{0, \pi/50, 2\pi/50, \cdots, 5\pi/50\}$
- $\mathbb{C}$: such that the resulting strains are within the linear region of material response (1%).

- Compute the coarse-grained displacements with grid size Δx=5 and normalize each sample such that $||\mathbf{f}_i||_{L^2(\Omega)} = 1$





NOR: Coarse-grained MD model for graphene

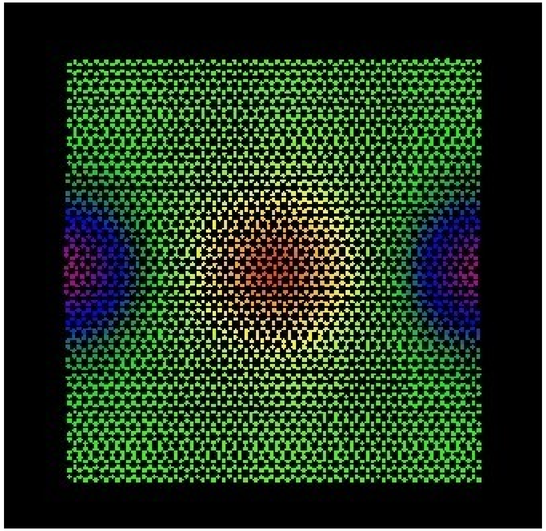
- Perform MD modeling and coarse graining of a perfect graphene sheet under point loads for **10 validation samples**:

  Data set generation: for $(x_1, x_2) \in [-100, 100]^2 \mathring{A}$

  solve the MD problem at constant temperature (0 or 300K)

  using a "thermostat" with additional external forcing

$$\sum_{n=-1}^{1} (-1)^n \exp\left(\frac{-1}{1 - \frac{(x-na)^2+y^2}{r^2}} \right)$$



- This dataset has the **same domain and grids but under substantially different loading conditions**.

NOR: Coarse-grained MD model for graphene

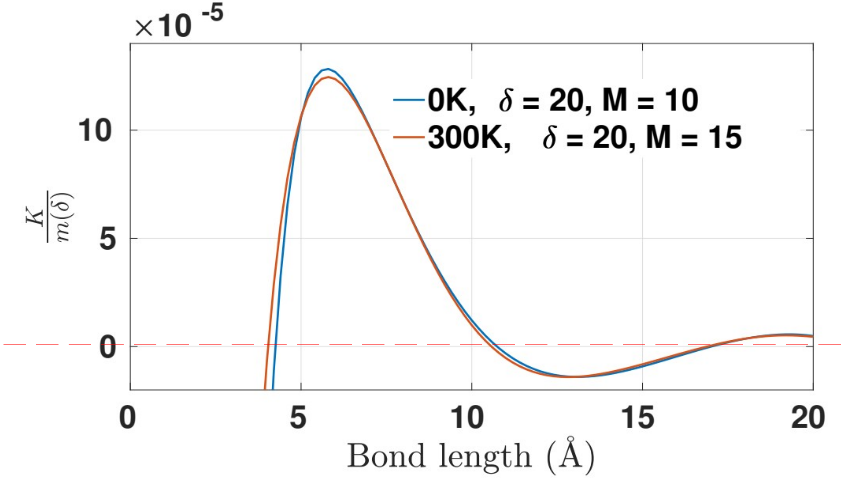
-0.38 in [1]
-0.33 in [2]

- We first study the perfect graphene crystal structu… no noise.
- Optimal parameters:

$$\delta = 20 \text{ Angstrom, M=10}$$

$$\text{Young's modulus } E = 0.91 \text{ TPa, Poisson's ratio } \nu = -0.43, \ \alpha = 2.83$$

- Optimal Kernel K:



[1] Qin, Huasong, et al. "Negative Poisson's ratio in rippled graphene." Nanoscale 9.12 (2017): 4135-4142.
[2] Jiang, Jin-Wu, et al. "Intrinsic negative Poisson's ratio for single-layer graphene." Nano letters 16.8 (2016): 5286-5290.

NOR: Coarse-grained MD model for graphene

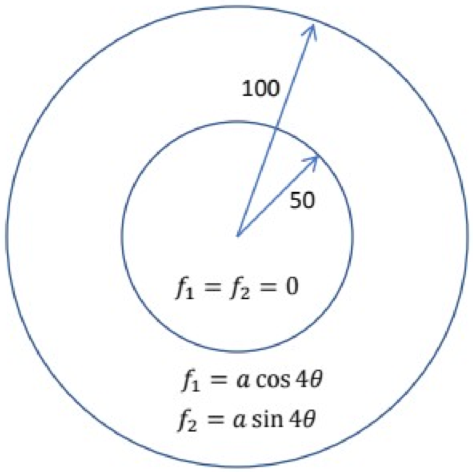
- Perform MD modeling and coarse graining of a perfect graphene sheet for 4 test samples with circular domain and zero loading:

Model parameters: 0K or 300K, $\Delta x = 5$Å

Testing domain: A circular object with radius 100Å

Sample testing forcing term:

$$\mathbf{f} = (0, 0), \text{ when } r \leq 50,$$
$$\mathbf{f} = (a\cos(4\theta), a\sin(4\theta)), \text{ when } 50 < r \leq 100.$$

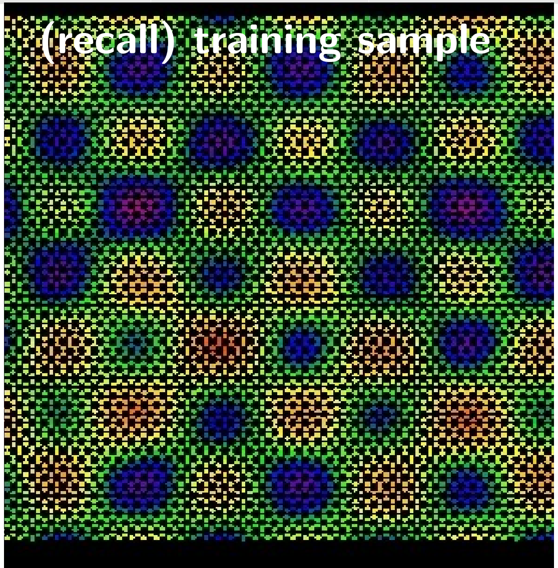


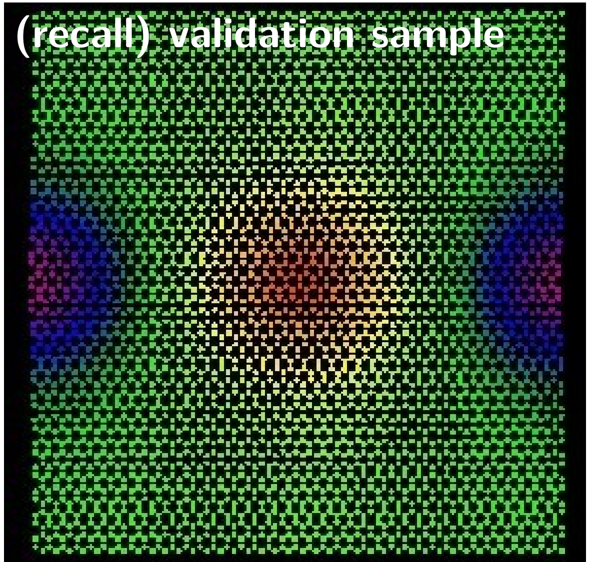
- This dataset has **substantially different domain and loading conditions**.

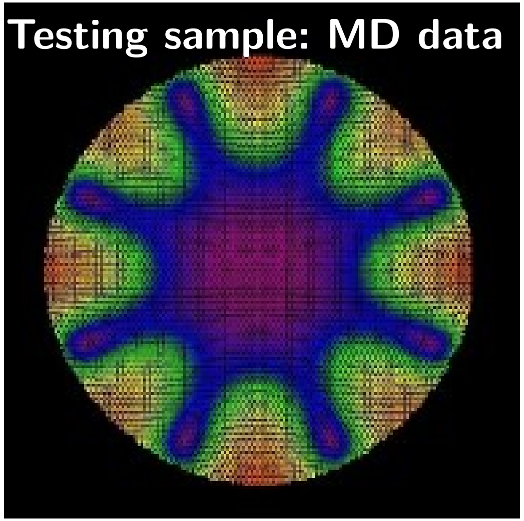
NOR: Coarse-grained MD model for graphene

- Perform MD modeling and coarse graining of a perfect graphene sheet for 4 test samples with circular domain and zero loading:

Training set	Young's modulus	Poisson ratio	α	Training Loss	Training error in u	Validation Loss	Validation error in u	Test error in u
0K	0.91 TPa	-0.43	2.8	9.81%	11.72%	13.28%	7.16%	**6.75%**
300K, Low	0.90 TPa	-0.42	2.6	9.82%	13.16%	18.08%	8.88%	**9.21%**





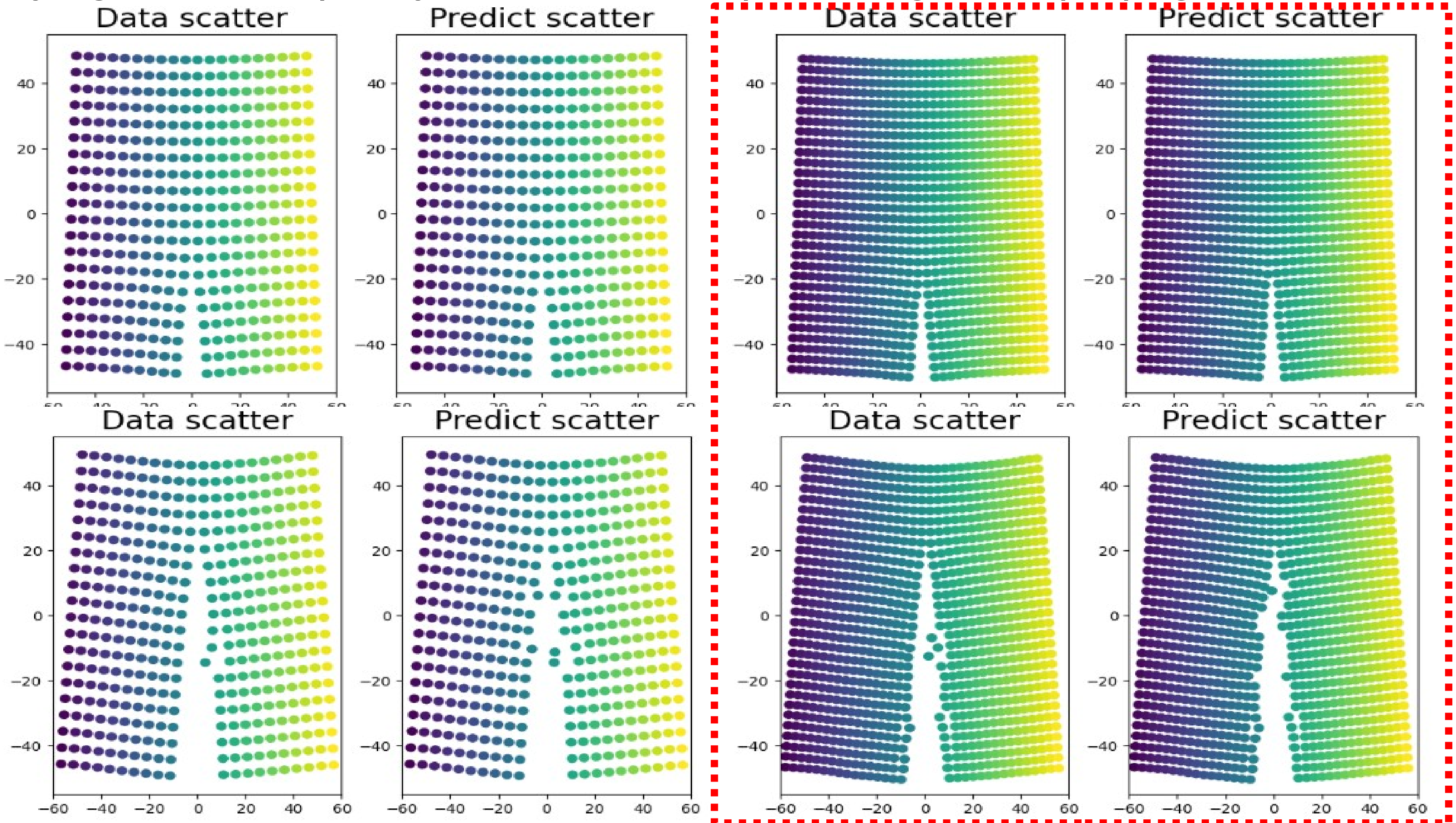


NOR: Coarse-grained MD model for graphene

- Employing the learnt peridynamic model in predicting crack propagation.

Prediction at 20 steps

Prediction at 40 steps



Learn the kernel and critical stretch ratio on Δx=5, and apply on Δx=2.5

Part II
Learning Nonlocal Neural Operators for Heterogeneous Models

[1] N. Liu, Y. Yu*, H. You, N. Tatikola. "INO: Invariant Neural Operator for Learning Complex Physical Systems with Momentum Conservation", AISTATS, 2023

[2] H. You, **Y. Yu***, M. D'Elia, T. Gao, S. Silling, "Nonlocal Kernel Network (NKN): a stable and resolution independent deep neural network". JCP, 2022

[3] L. Zhang, H. You, T. Gao, M. Yu, C-H. Lee, **Y. Yu***, "MetaNO: How to Transfer Your Knowledge on Learning Hidden Physics", Under Review, 2023.

[4] H. You, Q. Zhang, C. Ross, C-H. Lee, **Y. Yu***, "Learning Deep Implicit Fourier Neural Operators (IFNOs) with Applications to Heterogeneous Material Modeling". CMAME, 2022.

[5] H. You, Q. Zhang, C. Ross, C-H. Lee, M-C. Hsu, **Y. Yu***, "A Physics-Guided Neural Operator Learning Approach to Model Biological Tissues from Digital Image Correlation Measurements". Journal of Biomechanical Engineering, 2022.

Nonlocal Neural Operators

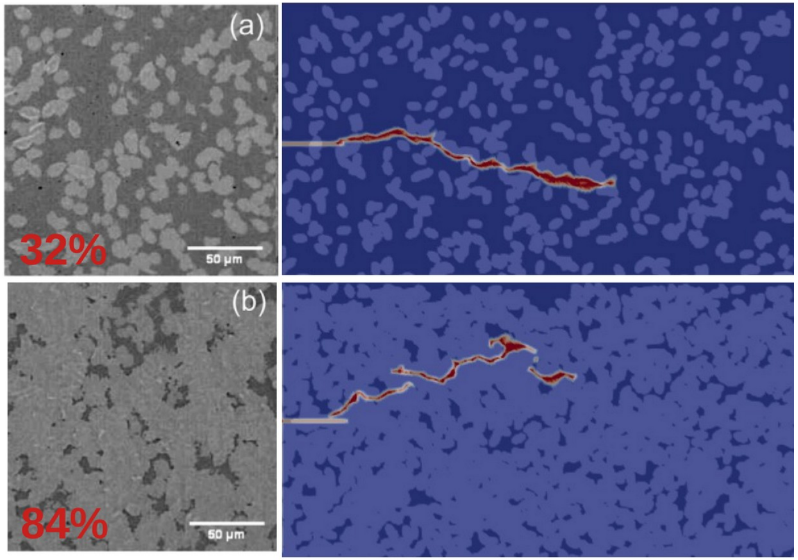
Propose: a nonlocal neural constitutive law for nonlinear and heterogeneous materials

- **Idea**: the material response is governed by a constitutive law, parameterized as **a neural operator:**

$$\mathcal{G}[\mathbf{u}](\mathbf{x}) = \mathbf{f}(\mathbf{x})$$

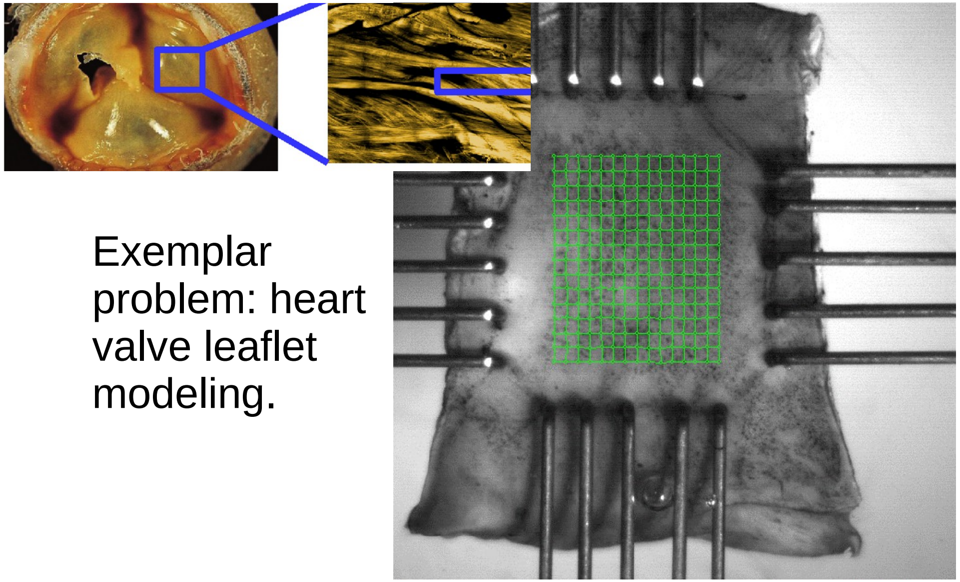
where f(x) is the external loading, and u(x) is the corresponding material responses.

Exemplar problem: crack on glass-ceramics.



Crack propagation simulations using peridynamics.

Exemplar problem: heart valve leaflet modeling.



Mechanical Testing of heart valve leaflet

Nonlocal Neural Operators

Propose: a nonlocal neural constitutive law for nonlinear and heterogeneous materials

- Assume: an **unknown** governing equation

$$\mathcal{G}[\mathbf{u}](\mathbf{x}) = \mathbf{f}(\mathbf{x}), \quad \mathbf{x} \in D,$$

$$\mathbf{u}(\mathbf{x}) = \mathbf{u}_{bc}(\mathbf{x}), \quad \mathbf{x} \in \partial D,$$

- Learn the neural operator $\mathcal{G} : \mathcal{A} \times \Theta \to \mathcal{U}$, such that for each data pairs, G[u]=f.
- **Advantages:**

  1. Only require observed data pairs $\{(\mathbf{f}_j, \mathbf{u}_j)\}_{j=1}^N$, and hence can be applied when the underlying constitutive law is unknown.

  2. G allows nonlinear and heterogeneous material responses.

  3. No further modification or tuning will be required for different resolutions and discretizations.
- **Cons:**

  1. Does not guarantee well-posedness nor physical laws.

[1]L. Lu, P. Jin, G. Pang, Z. Zhang, G. E. Karniadakis, Learning nonlinear operators via deeponet based on the universal approximation theorem of operators, Nature Machine Intelligence 3 (3) (2021) 218–229.
[2]Z. Li, N. Kovachki, K. Azizzadenesheli, B. Liu, K. Bhattacharya, A. Stuart, A. Anandkumar, Neural operator: Graph kernel network for partial differential equations, arXiv preprint arXiv:2003.03485.
[3]Chen, Ke, Chunmei Wang, and Haizhao Yang. "Deep Operator Learning Lessens the Curse of Dimensionality for PDEs." arXiv preprint arXiv:2301.12227 (2023).

INO: Neural Operator with Conservation Laws

- **Question**: How to impose basic physical laws into neural operators?
  Approach 1: As an additional penalization term: PINO, Physics-informed DeepONet, PG-IFNO, etc.
  Approach 2: Hard-coded into the NN architecture.

- **Propose**: a neural operator in the form of a state-based peridynamics formulation

$$\mathcal{G}[\mathbf{u}](x) := -\int_{\Omega} \mathbf{T}[x](\mathbf{u}(y) - \mathbf{u}(x), y - x) - \mathbf{T}[y](\mathbf{u}(x) - \mathbf{u}(y), x - y)dy = \mathbf{f}(x)$$

  which guarantees the **balances of total force and torque**.

How to design the stress state operator, T?

INO: Neural Operator with Conservation Laws

- **Question**: How to impose basic physical laws into neural operators?

  Approach 1: As an additional penalization term: PINO, Physics-informed DeepONet, PG-IFNO, etc.

  Approach 2: Hard-coded into the NN architecture.

Noether's theorem (Connections between symmetry and conservation laws):

Consider a system whose dynamical state at a given instant of time t can be described by a set of generalized coordinates $\mathbf{x}=[x_1,x_2,\ldots x_f]$, and a set of generalized velocities $\mathbf{p}=[p_1,p_2,\ldots,p_f]$, and for which there exists a Lagrangian function $L(t,\mathbf{x},\mathbf{p})$ which, when substituted into Lagrange's equations of motion, determines the dynamical behavior of the system.

1) If the Lagrangian function, L, is invariant under a translation in a particular direction, the total linear momentum of the system is a constant of the motion.

2) If the Lagrangian is invariant under a rotation in space, then the angular momentum of the system is a constant of the motion.

Desloge, E. A., & Karch, R. I. (1977). Noether's theorem in classical mechanics. American Journal of Physics, 45(4), 336-339.

INO: Neural Operator with Conservation Laws

- **Question**: How to impose basic physical laws into neural operators?

  Approach 1: As an additional penalization term: PINO, Physics-informed DeepONet, PG-IFNO etc.

  Approach 2: Hard-coded into th

E.g., on material displacement modeling:

Translational Invariant → Linear Momentum Conservation

Rotational Equivariant → Angular Momentum Conservation

Noether's theorem (Connecti

Consider a system whose dynamical state at a g f time t can be described by a set of generalized coordinates $\mathbf{x}=[x_1,x_2,\ldots x_f]$, and a set of ge velocities $\mathbf{p}=[p_1,p_2,\ldots,p_f]$, and for which there exists a Lagrangian function $L(t,\mathbf{x},\mathbf{p})$ which, w ubstituted into Lagrange's equations of motion, determines the dynamical behavior of the system.

1) If the Lagrangian function, L, is invariant under a translation in a particular direction, the total linear momentum of the system is a constant of the motion.

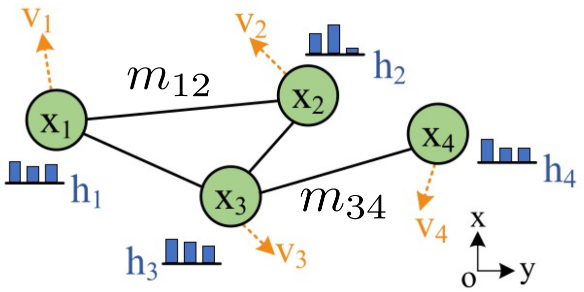
2) If the Lagrangian is invariant under a rotation in space, then the angular momentum of the system is a constant of the motion.

Desloge, E. A., & Karch, R. I. (1977). Noether's theorem in classical mechanics. American Journal of Physics, 45(4), 336-339.

EGNN: Equivariance in GNNs

- **Equivariant Graph Neural Network(EGNN)**: learn graph neural networks equivariant to rotations, translations, reflections and permutations
- **h=node features, m=edge features**

GNN: $$m_{ij} = \phi_e(\mathbf{h}_i^l, \mathbf{h}_j^l, a_{ij}), \; m_i = \sum m_{ij}, \; \mathbf{h}_i^{l+1} = \phi_h(\mathbf{h}_i^l, m_i)$$



EGNN: $$\mathbf{x}_i^{l+1} = \mathbf{x}_i^l + C\sum_{j\neq i}(\mathbf{x}_i^l - \mathbf{x}_j^l)\phi_x(m_{ij})$$

$$m_{ij} = \phi_e(\mathbf{h}_i^l, \mathbf{h}_j^l; |\mathbf{x}_i^l - \mathbf{x}_j^l|^2; a_{ij}), \; m_i = \sum m_{ij}, \; \mathbf{h}_i^{l+1} = \phi_h(\mathbf{h}_i^l, m_i)$$

[1]Satorras, V. G., Hoogeboom, E., & Welling, M. (2021). E (n) equivariant graph neural networks. In International conference on machine learning (pp. 9323-9332). PMLR.

INO: Neural Operator with Conservation Laws

- **Question**: How to impose basic physical laws into neural operators?

EGNN

$$\mathbf{x}_i^{l+1} = \mathbf{x}_i^l + C\sum_{j\neq i}(\mathbf{x}_i^l - \mathbf{x}_j^l)\phi_x(m_{ij})$$

$$m_{ij} = \phi_e(\mathbf{h}_i^l, \mathbf{h}_j^l, |\mathbf{x}_i^l - \mathbf{x}_j^l|^2, a_{ij}), \; m_i = \sum m_{ij}, \; \mathbf{h}_i^{l+1} = \phi_h(\mathbf{h}_i^l, m_i)$$

- **Propose**: a **invariant neural operator** in the form of a state-based peridynamics formulation

$$\mathcal{G}[\mathbf{u}](x) := -\int_\Omega \mathbf{T}[x](\mathbf{u}(y) - \mathbf{u}(x), y - x) - \mathbf{T}[y](\mathbf{u}(x) - \mathbf{u}(y), x - y)dy = \mathbf{f}(x)$$

where

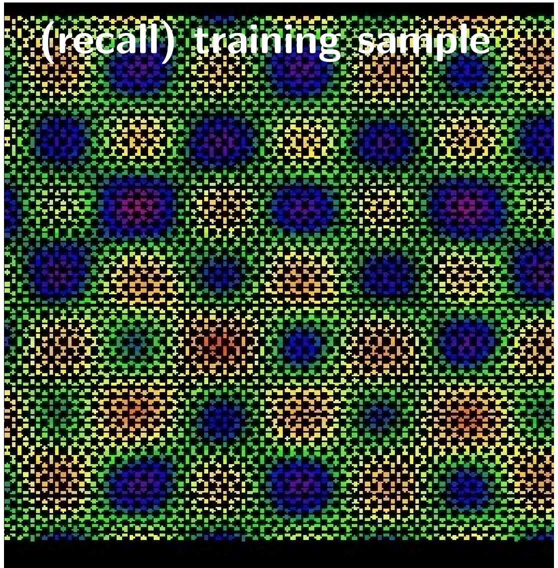
$$\mathbf{T}[x](\mathbf{u}(y) - \mathbf{u}(x), y - x) := (\mathbf{u}(y) - \mathbf{u}(x) + y - x)\phi(\mathbf{m}(|x-y|, \theta; v), \mathbf{h}(x), |y - x + \mathbf{u}(y) - \mathbf{u}(x)|, |y-x|; w)$$

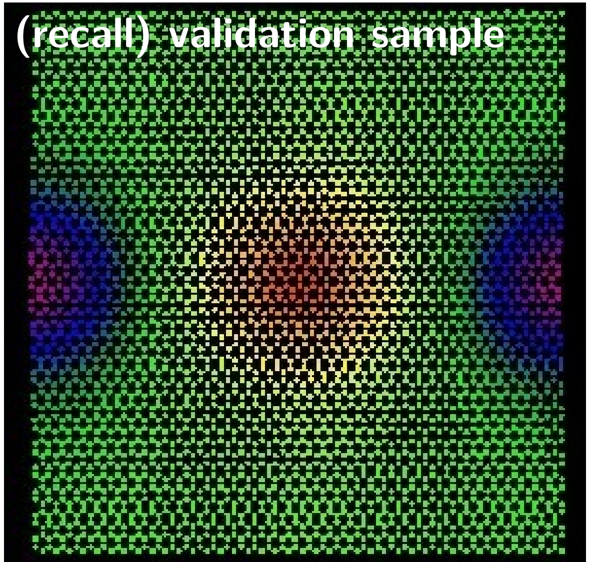
$$\mathbf{h}(x) := \int_\Omega \mathbf{m}(|x-y|, \theta; v)(|y - x + \mathbf{u}(y) - \mathbf{u}(x)| - |y-x|)|y-x|dy$$

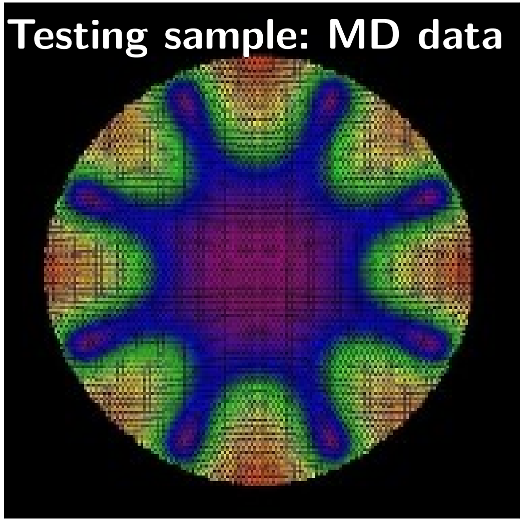
INO example 1: MD dataset

- Perform MD modeling and coarse graining of a perfect graphene sheet for 4 test samples with circular domain and zero loading:

Model	Young's modulus	Poisson ratio	α	Validation Loss	Validation error in u	Test error in u
NOR	0.91 TPa	-0.43	2.8	13.28%	7.16%	**6.75%**
INO	N/A	N/A	N/A	9.80%	3.20%	**3.40%**







INO example 2: synthetic dataset

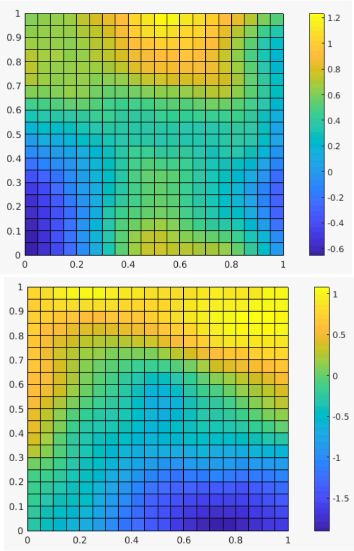
- **200 training and 25 test samples:** generated from the Holzapfel-Gasser-Odgen (HGO) model

(Ground-truth) strain energy density function:

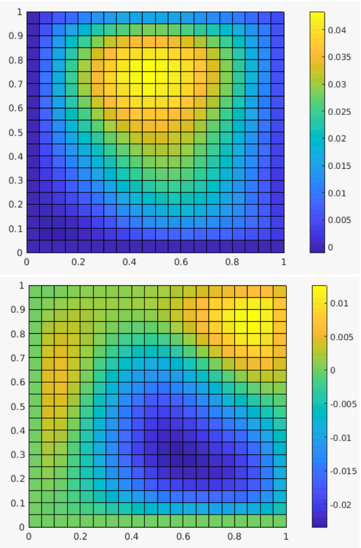
$$\frac{E}{4(1+\nu)}(\bar{I}_1-2)-\frac{E}{2(1+\nu)}\ln(J)+\frac{k_1}{2k_2}\left(\exp\left(k_2\langle S(\alpha)\rangle^2\right)+\exp\left(k_2\langle S(-\alpha)\rangle^2\right)-2\right)+\frac{E}{6(1-2\nu)}\left(\frac{J^2-1}{2}-\ln J\right).$$

with: E=0.973, ν=0.265, k_1=0.1, k_2=1.5, α=π/2. material is anisotropic and nonlinear.

Body load

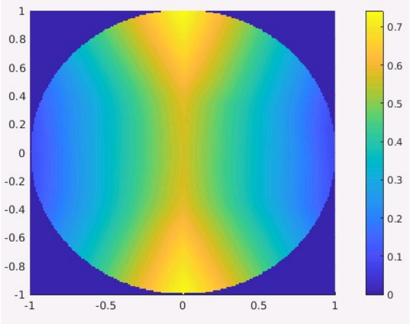


Displacement field



Test error: 4.15%

Learnt Kernel m:



Conclusion

- We proposed two new data-driven **nonlocal constitutive models**, NORs and INOs, which learns **continuous integrants** for material learning tasks.
- For **linear & homogenized model learning tasks**, the **nonlocal operator regression (NOR) model** is proposed, which learns optimal kernel functions directly from data.
- For **nonlinear & heterogeneous material modeling tasks**, the **invariant neural operator (INO) model** is proposed, which guarantees the linear and angular momentum conservation laws, and resembles nonlinear peridynamics.
- We employed NOR and INO to learn several exemplar material models directly from high-fidelity simulations/experimental measurements, and show that the learnt nonlocal operators are generalizable to different resolutions and loading scenarios.

Thank you!

- **Collaborators:**

  Huaiqian You (Ph.D. student), Siavash Jafarzadeh (postdoc), Neeraj Tatikola (master student), *Lehigh University*

  Stewart Silling, *Sandia National Lab,* Marta D'Elia, *Meta,* Ning Liu, *GEM.*

  Fei Lu, Qingci An, *JHU*

- **Funding support:**

  NSF CAREER award DMS1753031

  AFOSR YIP grant FA9550-22-1-0197

- **Computational Resources:** Lehigh HPC systems

- **References:**

- [1] Lu, F., An, Q., & Yu, Y. (2022). Nonparametric learning of kernels in nonlocal operators. arXiv preprint arXiv:2205.11006.

  [2] H. You, Y. Yu, S. Silling, M. D'Elia, "A data-driven peridynamic continuum model for upscaling molecular dynamics". CMAME, 2022.

  [3] N. Liu, Y. Yu, H. You, N. Tatikola. "INO: Invariant Neural Operator for Learning Complex Physical Systems with Momentum Conservation", AISTATS, 2023.