SAND2021-8176C

A data-driven peridynamic continuum model for upscaling molecular dynamics

Huaiqian You, Lehigh University & Sandia National Lab, CA

Joint work with **Yue Yu** (Lehigh University)
Stewart Silling (Sandia National Lab, NM)
Marta D'Elia (Sandia National Lab, CA)

Present at the 2021 MLDL workshop

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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.

Outline

- Introduction to the nonlocal modeling
- Challenges in MD (molecular dynamics) and nonlocal model
- A nonlocal operator regression algorithm
- Nonlocal surrogates for MD simulations

What is a nonlocal model?

- **Basic concepts**

1) The state of a system at any material point depends on the state in a neighborhood of points
2) Interactions can occur at distance without contact
3) The solutions has a lower requirement of regularity: non-differentiable, singular, discontinuous

- **Why nonlocal model?** The nonlocal model can capture the feature that traditional PDE fails to capture

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

- **Nonlocal operator with kernel K**

$$\mathcal{L}_K(u) = \int_{B_\delta(x)} K(x,y)(u(y)-u(x))\mathrm{d}y$$

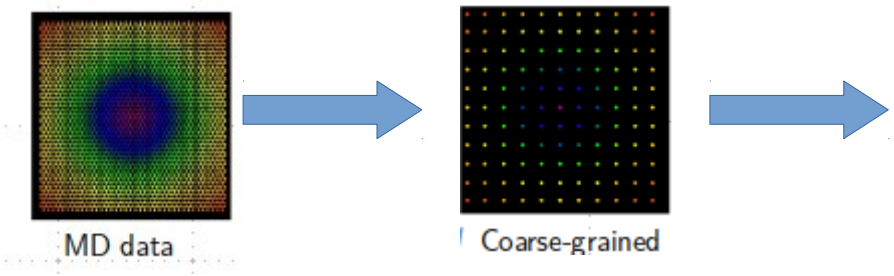
Challenges in MD and nonlocal model

Existing challenges in MD and nonlocal model:

- MD has become a fundamental tool for inventing new materials, but it suffers from the its computational limitations in length and time scale it can address.
- We need to upscale MD to the continuum level, and nonlocal models exhibit all desired properties: 1. can handle discontinuities; 2. compatible with molecular scale forces; 3. can capture multiscale behavior
- **Knowledge gap**: Nonlocal kernels are often chosen as a posteriori. It's difficult to derive them from the physical laws.

Our strategy:

1) Collect high-fidelity data from MD simulations
2) Use machine learning algorithm to learn a nonlocal kernel as well as the material properties from coarse-grained MD data
3) Learned nonlocal operator needs to satisfy some solvability constraints to ensure well-posedness.



Coarse-grained MD data + ML = new advanced material model

Learning process

- **Goal** : Identify the kernel function as well as the material properties
- **Given** : A collection of samples of coarse-grained MD displacement and forcing

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

- **Model :** Linearized peridynamic solid (LPS) model

$$\mathcal{L}_K \mathbf{u}(\mathbf{x}) := -\frac{C_\alpha}{m(\delta)} \int_{B_\delta(\mathbf{x})} (\lambda - \mu)\, K(|\mathbf{y}-\mathbf{x}|)\, (\mathbf{y}-\mathbf{x})\, (\theta(\mathbf{x}) + \theta(\mathbf{y}))\, d\mathbf{y}$$
$$- \frac{C_\beta}{m(\delta)} \int_{B_\delta(\mathbf{x})} \mu K(|\mathbf{y}-\mathbf{x}|) \frac{(\mathbf{y}-\mathbf{x}) \otimes (\mathbf{y}-\mathbf{x})}{|\mathbf{y}-\mathbf{x}|^2} (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y} = \mathbf{b}(\mathbf{x}),$$
$$\theta(\mathbf{x}) := \frac{2}{m(\delta)} \int_{B_\delta(\mathbf{x})} K(|\mathbf{y}-\mathbf{x}|)(\mathbf{y}-\mathbf{x}) \cdot (\mathbf{u}(\mathbf{y}) - \mathbf{u}(\mathbf{x}))\, d\mathbf{y},$$

and K is approximated by Bernstein polynomials:

$$K(\mathbf{x}, \mathbf{y}) = \sum_{k=0}^{M} \frac{D_k}{|\mathbf{x}-\mathbf{y}|^\alpha} B_{k,M}\left(\frac{|\mathbf{x}-\mathbf{y}|}{\delta}\right).$$

- **Optimization**: Learn K, Young's modulus E, Poisson ratio ν, and α by minimizing the residual $\frac{1}{N}||\mathcal{L}_K \mathbf{u}_i - \mathbf{b}_i||_2^2$ subject to solvability constraints

Solvability constraints

- When the kernel function K is positive and not too singular, the LPS model is guaranteed to be solvable
- To better improve the accuracy of the nonlocal model, we allow the kernel to be partially negative.
- While there's no known theory on sign-changing kernel available for the LPS model, we impose the solvability constraints in a discrete manner

Theorem : the discrete LPS system is solvable if

$$\min eig(A) > 0, \text{coercivity},$$
$$\min eig(BA^{-1}B^t) > 0, \text{inf-sup},$$
$$\min eig(A - 2B^tB) \geq 0, \text{Cauchy Schwars},$$

where A and B are the discrete operators

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

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

The two-phase algorithm

- Phase 1: Constrain the parameters $\{D_k\}$ to be non-negative, and minimize the residual

$$\{D_k^{pre}, E^{pre}, \nu^{pre}, \alpha^{pre}\} = argmin \frac{1}{N} ||\mathcal{L}_\delta \mathbf{u}_i - \mathbf{b}_i||_2^2$$

- Phase 2: Allow the kernel to be negative and apply the solvability constraints while minimizing the residual

$$\{D_k, E, \nu, \alpha\} = argmin \frac{1}{N} ||\mathcal{L}_\delta \mathbf{u}_i - \mathbf{b}_i||_2^2$$

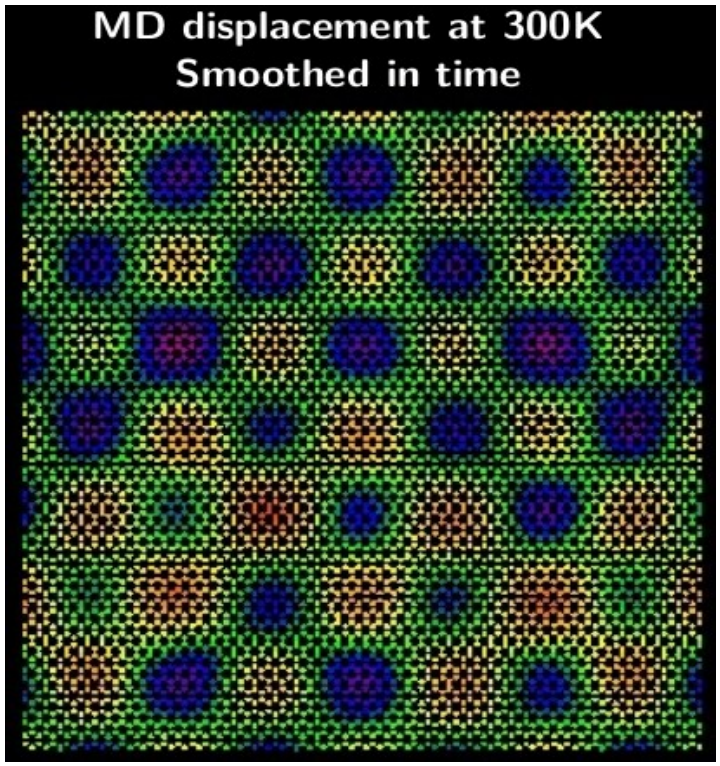
$$\min eig(A) > 0$$

$$\min eig(BA^{-1}B^t) > 0$$

$$\min eig(A - 2B^tB) \geq 0$$

MD simulations of graphene: Training set

- Perfrom MD simulation of a perfect sheet under the loads **b** at 0K and 300K
- Compute the smoothed (coarse-grained) displacements for **70** training samples, grid size is 5Å



The domain is a [-50Å,50Å]^2 square, the given loads **b** are

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

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

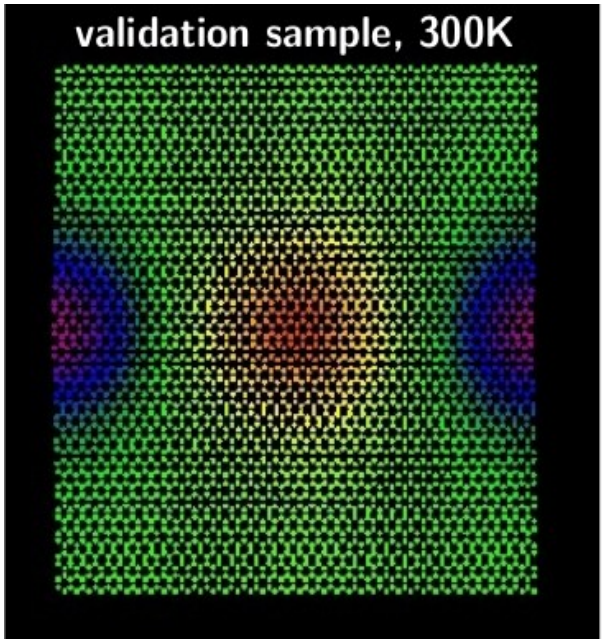
Where $k_1, k_2 \in \{0, \pi/50, 2\pi/50, \ldots, 5\pi/50\}$

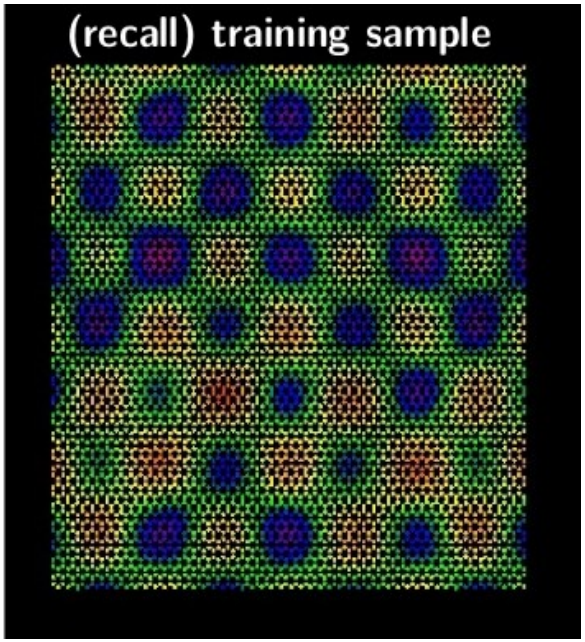
$\mathbb{C}$ Is chosen such that the resulting strain are within the small strain regime (1%)

MD simulations of graphene: Validation set

- Perfrom MD simulation of a perfect sheet under the loads **b** at 0K and 300K
- Compute the smoothed (coarse-grained) displacements for **10** training samples

$$\mathbf{b}(\mathbf{x}) = \sum_{n=-1}^{1} \exp\left\{ \frac{-1}{1 - \frac{(x_1 - na)^2 + x_2^2}{r^2}} \right\}$$





The training and validation datasets have the same domain but different loading conditions

Optimal choice of hyperparameters (δ, M)

- To find the optimal hyperparameters, we compare the optimal δ for each given kernel order M, using two metrics on training and validation datasets.

  Error in residual : $average\{||\mathcal{L}_\delta \mathbf{u}_i - \mathbf{b}_i||\}_{i=1}^N$

  Error in solution : $average\{||\mathcal{L}_\delta^{-1} \mathbf{b}_i - \mathbf{u}_i||\}_{i=1}^N$

- Results:

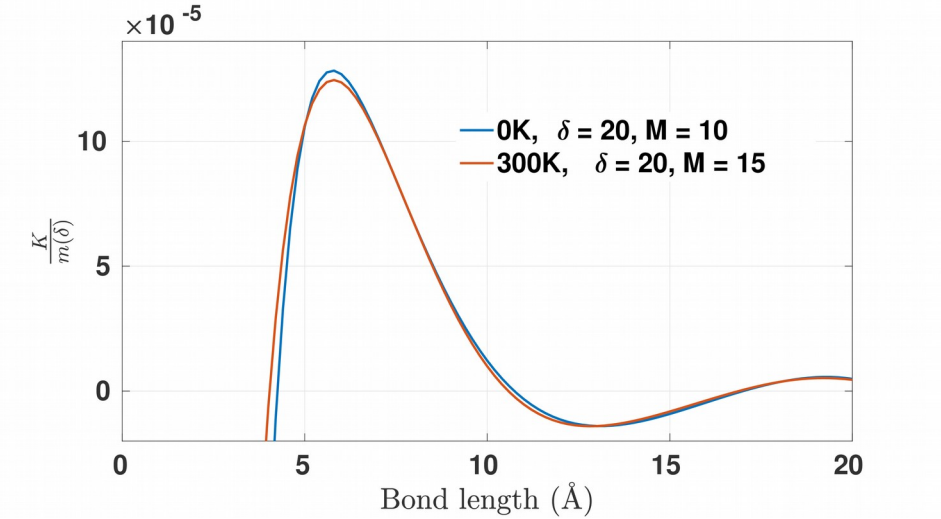
data set	M	$\delta_M^\star$	E_{Res}^{train}	$E_{\mathbf{u}}^{train}$	E_{Res}^{val}	$E_{\mathbf{u}}^{val}$	AvgE
$0K$	0	12.5Å	13.91%	17.54%	16.31%	14.49%	1
	5	12.5Å	10.42%	12.19%	13.02%	7.69%	0.6933
	10	**20Å**	9.81%	11.72%	13.28%	7.16%	**0.6704**
	15	22.5Å	9.80%	11.61%	13.50%	7.22%	0.6731
	20	25Å	9.75%	11.89%	13.53%	7.00%	0.6729
$300K$	0	12.5Å	13.46%	31.33%	20.15%	14.86%	1
	5	12.5Å	10.50%	13.80%	17.83%	9.66%	0.6784
	10	20Å	9.79%	13.32%	18.11%	9.08%	0.6549
	15	**20Å**	9.82%	13.16%	18.08%	8.88%	**0.6505**
	20	25Å	9.81%	13.36%	18.34%	9.23%	0.6609

Kernel function and material properties

- We compare the material properties obtained from 0K and 300K

Dataset	E (Tpa)	Poisson ratio ν	α
0K	0.91	-0.4297	2.8335
300K	0.90	-0.4196	2.5946

- Optimal kernel K:

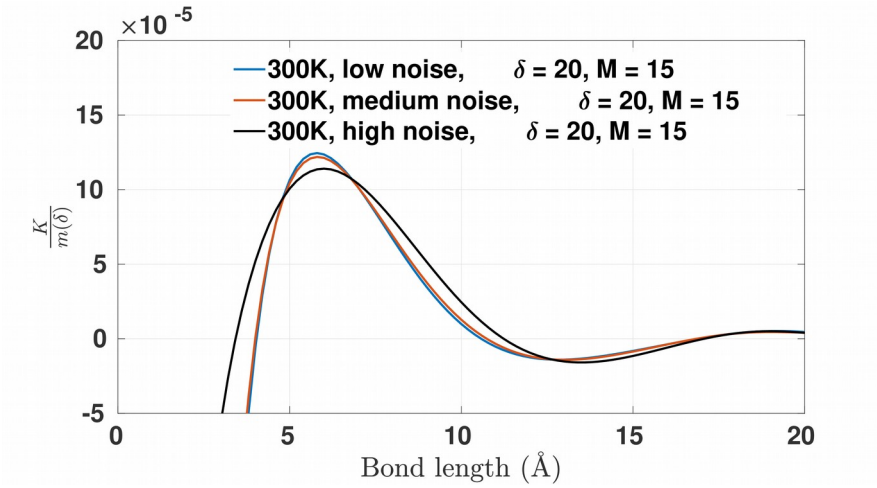


Kernel functions and material properties are not sensitive to temperature

Kernel function and material properties at 300K

- We compare the material properties at 300K with different noise level.

Datasets	E (Tpa)	ν	α	Residual (train)	Solution (train)	Residual (Val)	Solution (Val)
low noise	0.90	-0.4196	2.5946	9.82%	13.16%	18.08%	8.88%
med noise	0.87	-0.4422	2.5197	14.52%	28.27%	18.34%	9.82%
high noise	0.95	-0.3106	1.9365	27.64%	48.86%	23.73%	17.54%



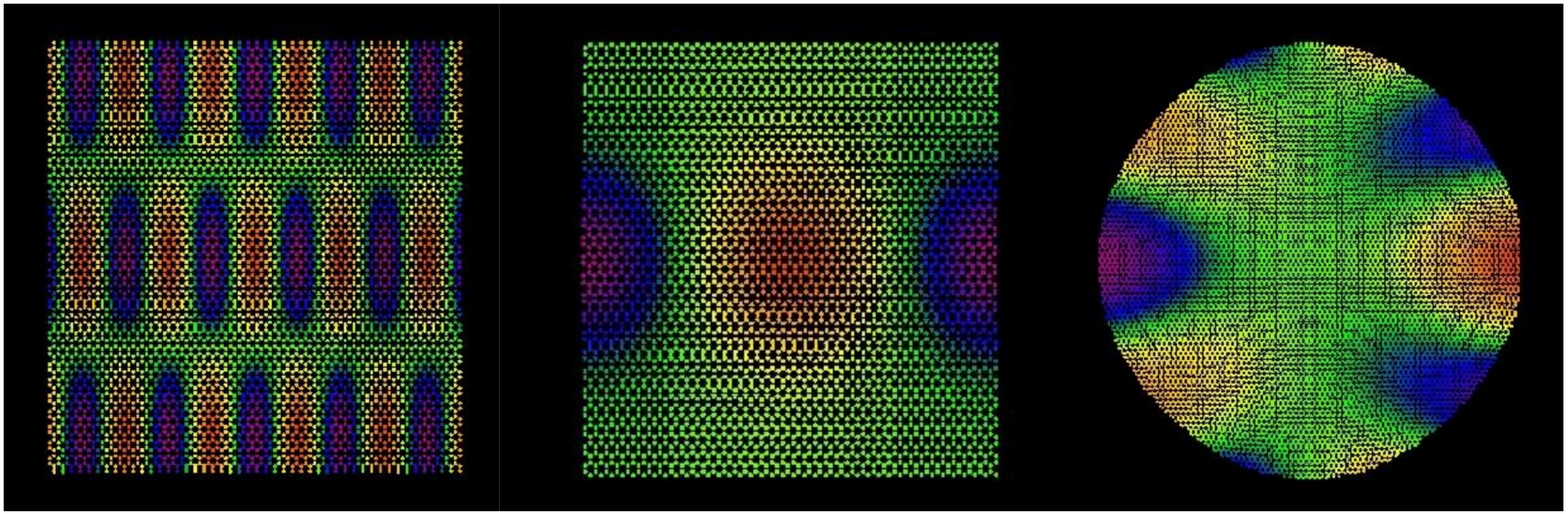
Our learning strategy is robust with the existence of thermal noise.

Generalization to different domains/loadings

- Graphene at 0K on a different domain, circular object with radius 100Å
- Forcing terms are given as

$$\mathbf{b} = (0,0) \text{for} \quad r \leq 50$$

$$\mathbf{b} = (a\cos(4\theta)\cos(\theta), a\cos(4\theta)\sin(\theta)) \text{for} \quad r > 50$$



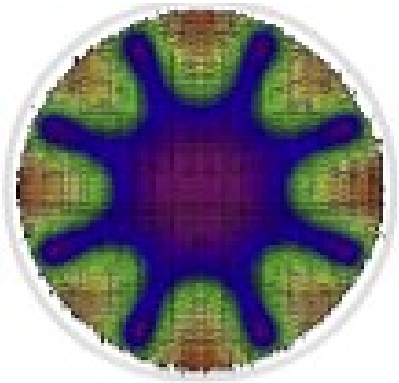
Training set | Validation set | Test set

Generalization to different domains/loadings

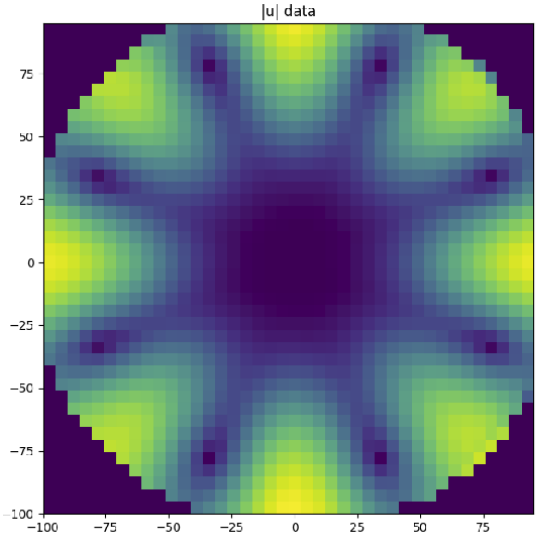
- Graphene at 0K on a different domain, circular object with radius 100Å
- Forcing terms are given as

$$\mathbf{b} = (0,0) \text{for} \quad r \leq 50$$

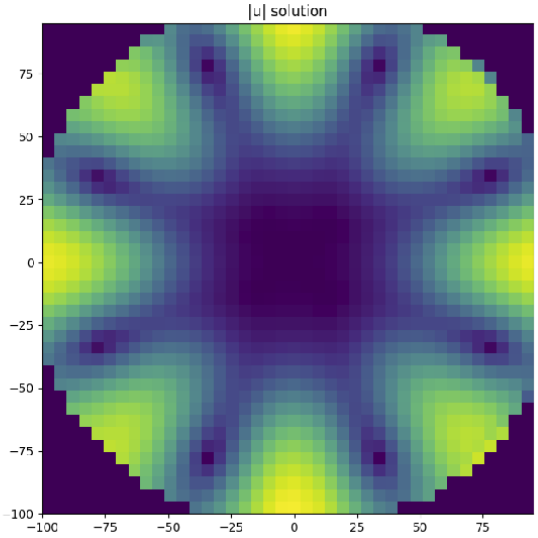
$$\mathbf{b} = (a\cos(4\theta)\cos(\theta), a\cos(4\theta)\sin(\theta)) \text{for} \quad r > 50$$



Test sample



|u| data



|u| solution (9% accuracy)

Generalization to different discretizations

- We consider training on hybrid datasets with grid size 5Å and 2.5Å, and compare the residual error and solution error on each grid size.

$\mathbb{S}_{train}$	$\mathbb{S}_{val}$ and $\mathbb{S}_{test}$	E_{Res}^{val}	$E_{\mathbf{u}}^{val}$	E_{Res}^{test}	$E_{\mathbf{u}}^{test}$
$h = 5\mathring{A}$ and $h = 2.5\mathring{A}$	$h = 2.5\mathring{A}$	16.19%	8.01%	2.95E-0	8.44%
$h = 5\mathring{A}$ and $h = 2.5\mathring{A}$	$h = 5\mathring{A}$	13.24%	9.29%	1.97E-1	7.80%

Our learning strategy can handle data from different discretizations!

Summary

- We proposed a machine learning method that extracts nonlocal models from coarse-grained MD data.
- The learned nonlocal model is guaranteed to be solvable.
- The learned nonlocal model is robust with the presence of thermal noise, and can be generalized to problems with different loadins/domains/discretizations.