Appl. Math. Mech. -Engl. Ed., **44**(7), 1125–1150 (2023)

APPLIED MATHEMATICS AND MECHANICS (ENGLISH EDITION)

https://doi.org/10.1007/s10483-023-2996-8

Towards a unified nonlocal, peridynamics framework for the coarse-graining of molecular dynamics data with fractures*

H. Q. YOU[1], X. XU[2], Y. YU[1,†], S. SILLING[3], M. D'ELIA[4], J. FOSTER[2]

1. Department of Mathematics, Lehigh University, Bethlehem, PA 18015, U. S. A.;
2. Department of Petroleum and Geosystems Engineering, The University of Texas at Austin, Austin, TX 78712-1139, U. S. A.;
3. Center for Computing Research, Sandia National Laboratories, Albuquerque, NM 87185-5820, U. S. A.;
4. Center for Computing Research, Sandia National Laboratories, Livermore, CA 94551-0969, U. S. A.

(Received Dec. 17, 2022 / Revised May 17, 2023)

Abstract Molecular dynamics (MD) has served as a powerful tool for designing materials with reduced reliance on laboratory testing. However, the use of MD directly to treat the deformation and failure of materials at the mesoscale is still largely beyond reach. In this work, we propose a learning framework to extract a peridynamics model as a mesoscale continuum surrogate from MD simulated material fracture data sets. Firstly, we develop a novel coarse-graining method, to automatically handle the material fracture and its corresponding discontinuities in the MD displacement data sets. Inspired by the weighted essentially non-oscillatory (WENO) scheme, the key idea lies at an adaptive procedure to automatically choose the locally smoothest stencil, then reconstruct the coarse-grained material displacement field as the piecewise smooth solutions containing discontinuities. Then, based on the coarse-grained MD data, a two-phase optimization-based learning approach is proposed to infer the optimal peridynamics model with damage criterion. In the first phase, we identify the optimal nonlocal kernel function from the data sets without material damage to capture the material stiffness properties. Then, in the second phase, the material damage criterion is learnt as a smoothed step function from the data with fractures. As a result, a peridynamics surrogate is obtained. As a continuum model, our peridynamics surrogate model can be employed in further prediction tasks with different grid resolutions from training, and hence allows for substantial reductions in computational cost compared with MD. We illustrate the efficacy of the proposed approach with several numerical tests for the dynamic crack propagation problem in a single-layer graphene. Our tests show that the proposed data-driven model is robust and generalizable, in the sense that it is capable of modeling the initialization and growth of fractures under discretization and loading settings that are different from the ones used during training.

* Citation: YOU, H. Q., XU, X., YU, Y., SILLING, S., D'ELIA, M., and FOSTER, J. Towards a unified nonlocal, peridynamics framework for the coarse-graining of molecular dynamics data with fractures. *Applied Mathematics and Mechanics* (*English Edition*), **44**(7), 1125–1150 (2023) https://doi.org/10.1007/s10483-023-2996-8

† Corresponding author, E-mail: yuy214@lehigh.edu

Key words nonlocal model, machine learning, homogenization, peridynamics, material fracture

Chinese Library Classification O241.3
2010 Mathematics Subject Classification 65D05

1 Introduction

Detection and prediction of material damage progression attract lots of interest from the broad scientific and engineering community[1–10]. Physically, the propagation of cracks results from a long-term physical process with its origin in the atomistic scale, which often requires a micro-structural model such as molecular dynamics (MD). However, although MD has made enormous advances in capabilities through better algorithms, better interatomic potentials, and improvements in computational power, its direct employment in treating the deformation and failure of materials at the mesoscale is still largely beyond reach. At the mesoscale and above, a continuum model of mechanics is often required in practice. This fact creates the need for homogenized models that act at larger scales and that, combined with new advanced architectures, allow for fast and accurate predictions of material deformation and fracture[11–21].

In this paper, we aim to address the question of how to extract coarse-grained measurements and a homogenized surrogate model from MD simulations, which is able to capture material deformation and the nucleation and growth of fractures. Nonlocal models are among the best candidates for this task[22]. In the context of homogenization, nonlocal models are characterized by integral operators (as opposed to differentiable operators) that embed time and length scales in their definition. Therefore, they are able to capture long-range effects that classical partial differential equation (PDE) models fail to describe[23], which makes nonlocal models viable alternatives to PDE models when the effects of the small-scale behavior of a system affect its global state[22, 24–31]. For monitoring and predicting material fractures, because the nonlocal viewpoint avoids classical notions like deformation gradient, nonlocal models allow a natural description of processes requiring reduced regularity in the relevant solution[32–33]. As such, the nonlocal continuum mechanics model, in the form of peridynamics[23, 31, 34–41], provides a unified modeling of continuum media where continuity and complex material damage modes can be captured autonomously.

In peridynamics and the general nonlocal models, constitutive laws take the form of integrand functions, whose functional form is often justified a posteriori, which makes rigorous calibration and validation challenging and time-consuming. On the other hand, although the nonlocal constitutive laws must be consistent with the classical effective properties, they contain information about the small-scale response of the system and must be chosen to reproduce this response with the greatest fidelity. Therefore, it is desired to extract an optimal integrand function from small-scale data, such that the calibrated nonlocal model reproduces the material responses and can further serve as a homogenized surrogate for future material deformation and fracture prediction tasks[42–43]. Recently, with the explosion of machine learning, optimized nonlocal models have been designed with the purpose of accurately reproducing observed coarse-grained behavior and predicting unseen behavior with the learnt model. We refer the readers to the works[44–49] for several examples of the use of optimization-based machine learning for the design of homogenized nonlocal operators and the rigorous analysis of its learning theory[49–50].

Although successful in providing optimal nonlocal surrogates to the homogenization problem, to the authors' best knowledge, none of these approaches addresses the challenge of capturing the main features of dynamic fracture that are seen in small-scale data. Fundamental challenges are still present, mainly due to the two difficulties. First, when mapping the MD measurements onto a coarser grid, coarse-graining methods can use the mean atomic velocities weighted by a smoothing function[51]. In a nonlocal setting, a smoothed displacement field can be shown to evolve according to the peridynamic linear momentum balance[43, 46]. However,

once the material fracture occurs, such a weighted average approach might overly smooth the displacement field and introduce errors near cracks in the coarse-grained data set. Second, in peridynamics the material damage is often described by breaking bonds. Therefore, the integrand functions present jumps near the damage criterion, which results in nonsmooth losses in the optimization problem and hinders the application of a suite of continuous optimization techniques. Herein, we address these two challenges and present a complete workflow demonstrating how to obtain large-scale nonlocal descriptions that capture MD behavior with fractures.

To accomplish this, we develop a novel coarse-graining method which is inspired by the weighted essentially non-oscillatory (WENO) scheme, and extend the machine learning technique in our previous work[46] to identify a smoothed damage criterion together with optimal nonlocal kernel functions. We summarize our main contributions as below.

(I) We develop a novel coarse-graining approach from micro-scale fracture measurements, to automatically choose a locally smoothest stencil and capture the displacement discontinuities. While the existing coarse-graining techniques use the same smoothing stencil everywhere[43, 46, 51], our proposed method chooses the locally smoothest stencil to construct the coarse-grained material displacement field. As such, coarse-grained measurements are obtained as piecewise smooth solutions containing discontinuities without overly smoothing the crack pattern.

(II) We propose a new two-step optimization strategy, and identify the best upscaled surrogate in the form of peridynamics. Without prior knowledge of the material properties, the resultant peridynamics model describes the material deformation together with the nucleation and growth of fractures.

(III) The optimal nonlocal model generalizes well to fracture patterns that are substantially different from the ones used for training. The optimal model also enables extrapolation to longer time simulations and a multiscale capability to predictions across resolutions.

We adopt peridynamics as the continuum model because of its natural compatibility with the physical nature of cracks as discontinuities. In contrast, the local theory of solid mechanics uses PDEs, which fail to apply on the surfaces of an emergent or growing crack, on which the deformation field is not differentiable. The basic equations of other nonlocal theories, such as Eringen's[30], similarly contain partial derivatives with respect to the spatial coordinates, and therefore are also not directly applicable to fracture. The field equations of peridynamics do not contain these partial derivatives, and thus treat fracture on the same mathematical basis as continuous deformations.

This paper is organized as follows. Section 2 shows how to obtain an adaptive stencil in the form of a smoothness indicator function, to extract coarse-grained measurements from MD displacements with fractures. In Section 3, we summarize the linear peridynamic solid (LPS) model, the treatment of material fracture and the handling of free surfaces, and the discretization technique used in this work. Section 4 presents our two-step learning approach consisting of a kernel learning step and a damage criterion learning step. Section 5 verifies the learning technique for MD displacements and studies the generalizability of the resultant model. On a single-layer graphene, we demonstrate the efficacy of our workflow by identifying an optimal two-dimensional (2D) nonlocal model and employing this model in complex prediction tasks. In particular, we illustrate several properties including generalization with respect to loadings, domain settings, crack shapes, and grid resolutions. Section 6 summarizes our contributions and provides future research ideas.

2 Coarse-graining of MD displacement with damage

In this section, we introduce the coarse-graining method to map the displacement field data from MD simulations into a larger-scale discretized data cloud. For materials without fracture, in Refs. [43] and [46], a nonlocal coarse-graining method was proposed. In this method, the

coarse-grained displacement for each particle is defined as a weighted average of the microscale displacements in its neighborhood. As such, a smoothed displacement field is obtained, which preserves a linear momentum balance as a consequence of the momentum balance for the atoms.

However, this coarse-graining method hides a pitfall: once the material fracture occurs when introducing discontinuities in the displacement field, the weighted average approach would overly smooth the displacement field and smudge the crack pattern. To resolve this challenge, in this section, we will extend the coarse-graining method to an adaptive approach, so as to automatically handle the material fracture and its corresponding discontinuities in the MD displacement data set.

To introduce the coarse-graining method, we consider the MD data set as an assembly of S mutually interacting particles. Then, we define the mass of each mutually interacting particle as M_ε ($\varepsilon = 1, 2, \cdots, S$), the reference positions of these particles as $\boldsymbol{X}_\varepsilon$, and the displacement vectors as $\boldsymbol{U}_\varepsilon(t)$. Each particle is subject to a prescribed external force $\boldsymbol{B}_\varepsilon(t)$.

The coarse-grained measurements can be defined by choosing a compactly supported function $\omega(\boldsymbol{x}, \cdot)$ for each material point $\boldsymbol{x} \in \mathbb{R}^d$, such that

$$\int_{\mathbb{R}^d} \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \mathrm{d}\boldsymbol{x} = 1, \quad \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) = 0 \quad \text{if} \quad |\boldsymbol{x} - \boldsymbol{X}_\varepsilon| > R. \tag{1}$$

Here, R is a pre-chosen hyperparameter, representing the coarse-grained radius. Then, the smoothed material density and body force density are, respectively, expressed as

$$\rho(\boldsymbol{x}) = \sum_{\varepsilon=1}^{S} \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) M_\varepsilon, \quad \boldsymbol{b}(\boldsymbol{x}, t) = \sum_{\varepsilon=1}^{S} \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \boldsymbol{B}_\varepsilon. \tag{2}$$

Correspondingly, the smoothed displacement field at the material point $\boldsymbol{x}$ is obtained by

$$\boldsymbol{u}(\boldsymbol{x}, t) = \frac{1}{\rho(\boldsymbol{x})} \sum_{\varepsilon=1}^{S} \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) M_\varepsilon \boldsymbol{U}_\varepsilon(t). \tag{3}$$

In Ref. [46], the authors proposed to employ a general cone-shaped weighted function for all material points, by defining

$$\omega(\boldsymbol{x}_i, \boldsymbol{X}_\varepsilon) := \frac{\tau(\boldsymbol{x}_i, \boldsymbol{X}_\varepsilon)}{\sum\limits_j \tau(\boldsymbol{x}_j, \boldsymbol{X}_\varepsilon)}, \tag{4}$$

where $\tau(\boldsymbol{x}, \boldsymbol{X}) = \max\{0, R - |\boldsymbol{X} - \boldsymbol{x}|\}$. Such an approach is found to be effective for materials without fracture, where the displacement field is continuous. Then, in Ref. [46], a data-driven surrogate model was built from this coarse-grained displacement field, which has successfully captured the material properties as well as a constitutive law acting at a larger scale.

However, once the material fracture occurs, the smooth weight function such as the one in Eq. (4) would smudge the crack pattern and hence may compromise the reliability of the resultant surrogate model. As shown in Fig. 1(a), coarse-grained points appear on the middle of the crack, showing the effect of overly-smoothing the displacement field.

To provide coarse-grained displacement fields for both damaged and undamaged material regions, we propose to adjust the smoothing function $\omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ when the material point $\boldsymbol{x}$ is close to the crack. Intuitively, the weight function should choose the locally smoothest stencil and avoid crossing discontinuities in the averaging procedure as much as possible. The similar idea was employed in the essentially non-oscillatory (ENO) and WENO methods[52–53], to develop finite difference schemes for PDE problems with piecewise smooth solutions containing discontinuities. Inspired by the WENO methods, our key idea is to assign an additional

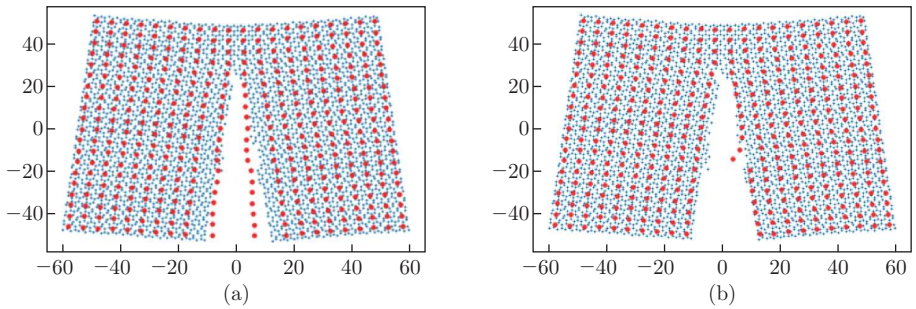


Fig. 1 An example of the vanilla coarse-graining method developed in Ref. [43] and our proposed approach, in handling the MD measurements with a crack. Small blue points represent the MD particles, and red dots stand for coarse-grained points. The results of (a) are obtained from the vanilla coarse-graining method with a weight function ω, and those of (b) are obtained from our proposed coarse-graining method with an adaptive weight function $\hat{\omega}$ (color online)

smoothness indicator function $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ to each pair of continuum material point $\boldsymbol{x}$ and MD particle $\boldsymbol{X}_\varepsilon$, and modify the weight function $\omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ as

$$\hat{\omega}(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) = \frac{\omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)}{\int_{\mathbb{R}^d} \omega(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)\mathrm{d}\boldsymbol{x}}. \tag{5}$$

When the crack intersects with the bond between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$, the displacement field between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$ contains discontinuity. One should use less information from $\boldsymbol{X}_\varepsilon$ to calculate the weighted average on $\boldsymbol{x}$, by taking the smoothness indicator function $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \approx 0$. That means, an adaptive procedure is required, to detect if there is a displacement jump between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$.

As demonstrated in Fig. 2, we construct the smoothness indicator function $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ with the following procedure. First, we project all the MD particles within a distance of R from $\boldsymbol{x}$ to the line segment that connects $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$. For each MD particle $\boldsymbol{X}_i$, we denote the projected point as $\widetilde{\boldsymbol{X}}_i$, and calculate its projected position variable d_i, as the distance is from $\boldsymbol{x}$ to $\widetilde{\boldsymbol{X}}_i$. The displacement vector $\boldsymbol{U}_i(t)$ on $\boldsymbol{X}_i$ is also projected, and its component along the segment $(\boldsymbol{x} - \boldsymbol{X}_\varepsilon)$ is denoted as U_i. Next, we select all the particles with their projections lying between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$, to form a set of data pairs $\mathcal{D} = \{(d_i, U_i)\}$. When plotting U_i as a function of d_i, the curve will present a jump when there is a crack intersecting the segment between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$, and such a jump would naturally divide the set $\mathcal{D}$ as two sets, with each set representing a smooth curve. Therefore, our goal is then to identify the discontinuity of $U(d)$ and define the smoothness indicator function according to it. Numerically, we loop over all possible combinations of splitting the data pair set $\mathcal{D}$ into two sets, $\mathcal{D}_1$ and $\mathcal{D}_2$, such that $\mathcal{D}_1 \cup \mathcal{D}_2 = \mathcal{D}$ and $\mathcal{D}_1 \cap \mathcal{D}_2 = \varnothing$. Then, we perform linear regressions on $\mathcal{D}_1$ and $\mathcal{D}_2$,

$$k_\beta, b_\beta = \underset{k,b}{\operatorname{argmin}} \sum_{(d_i, U_i) \in \mathcal{D}_\beta} |kd_i + b - U_i|^2, \quad \beta = 1, 2 \tag{6}$$

to obtain a fitted line for each set. In the meantime, we also perform a linear regression on the entire displacement data set $\mathcal{D}$, and obtain the fitted parameter set (k, b). Denote the total squared error $\epsilon(\mathcal{D}_1, \mathcal{D}_2)$ associated with $\mathcal{D}_1$ and $\mathcal{D}_2$ as

$$\epsilon(\mathcal{D}_1, \mathcal{D}_2) := \sum_{\beta=1}^{2} \sum_{\{(d_i, U_i)\} \in \mathcal{D}_\beta} (k_\beta d_i + b_\beta - U_i)^2, \tag{7}$$

and a squared error for the whole data set $\mathcal{D}$ as

$$\epsilon_0 := \sum_{\{(d_i, U_i)\} \in \mathcal{D}} (kd_i + b - U_i)^2. \tag{8}$$

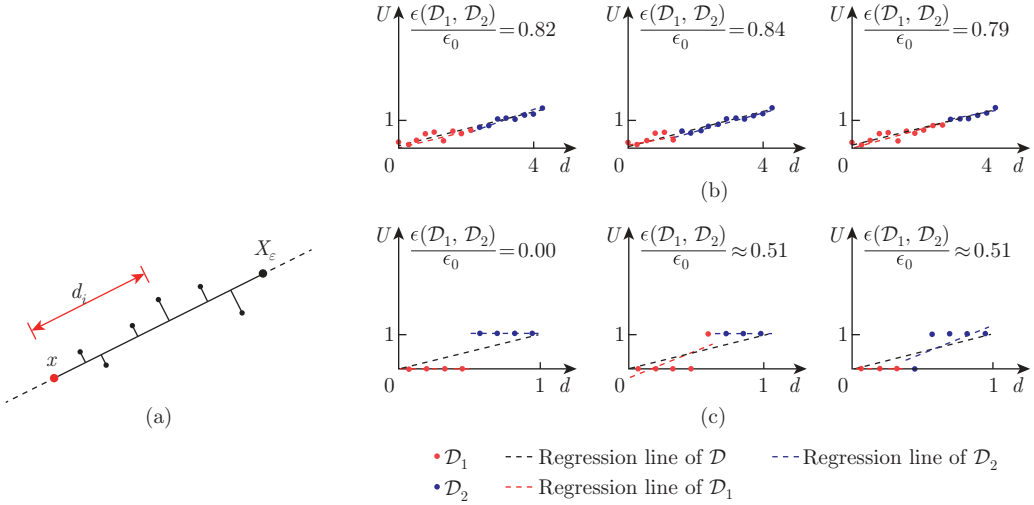


Fig. 2 Schematics and examples of calculation for the smoothness indicator function $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$: (a) a demonstration of the projection of MD points between $\boldsymbol{x}$ and $\boldsymbol{X}_\varepsilon$; (b) when the projected displacement field is continuous, the smoothness indicator $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ stays close to 1 since $\epsilon(\mathcal{D}_1, \mathcal{D}_2)$ is close to ϵ_0; (c) when material fracture occurs and the projected displacement field is discontinuous (with a jump at $d = 0.5$), $\dfrac{\epsilon(\mathcal{D}_1, \mathcal{D}_2)}{\epsilon_0}$ reaches the minimum when $\mathcal{D}_1$ and $\mathcal{D}_2$ both consist of a smooth curve, and we have $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \approx 0$ (color online)

Then, we define the smoothness indicator function $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ as

$$\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) := \frac{\min\limits_{(\mathcal{D}_1, \mathcal{D}_2)} \epsilon(\mathcal{D}_1, \mathcal{D}_2)}{\epsilon_0}. \tag{9}$$

Intuitively, when there is no material fracture, and hence $U(d)$ is a smooth curve without discontinuity, we anticipate to have $(k_\beta, b_\beta) \approx (k, b)$ for $\beta = 1, 2$. As a result, we have $\epsilon(\mathcal{D}_1, \mathcal{D}_2) \approx \epsilon_0$ and $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \approx 1$. In this case, the adjusted weight function $\hat{\omega}(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ will stay the same as the original weight function, and hence our smoothness indicator function will not alter the coarse-graining approach for materials without fracture. Figure 2(b) shows an example with continuous displacement. It can be observed that $\epsilon(\mathcal{D}_1, \mathcal{D}_2)$ stays roughly the same and close to ϵ_0 for different partitions. On the other hand, when fracture occurs, $\epsilon(\mathcal{D}_1, \mathcal{D}_2)$ would achieve its minimum when both $\mathcal{D}_1$ and $\mathcal{D}_2$ consist of a smooth curve. That means, $\mathcal{D}_1$ and $\mathcal{D}_2$ are separated by the crack. In this case, we will have $\min\limits_{(\mathcal{D}_1, \mathcal{D}_2)} \epsilon(\mathcal{D}_1, \mathcal{D}_2) < \epsilon_0$ and hence $\alpha(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) < 1$. Therefore, the adjusted weight function $\hat{\omega}(\boldsymbol{x}, \boldsymbol{X}_\varepsilon)$ would automatically reduce the weights of those particle points crossing discontinuities, so as to reduce the overly smoothing near cracks. Figure 2(c) presents an example where the displacement is piecewise constant with a jump at $d = 0.5$, demonstrating that the smooth indicator would reach its minimum when neither $\mathcal{D}_1$ nor $\mathcal{D}_2$ contains the displacement jump.

Once the adjusted weight functions are obtained, the smoothed mass density, body force density, and displacement can be, respectively, calculated as

$$\rho(\boldsymbol{x}) = \sum_{\varepsilon=1}^{S} \hat{\omega}(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) M_\varepsilon, \tag{10}$$

$$\boldsymbol{b}(\boldsymbol{x}, t) = \sum_{\varepsilon=1}^{S} \hat{\omega}(\boldsymbol{x}, \boldsymbol{X}_\varepsilon) \boldsymbol{B}_\varepsilon, \tag{11}$$

$$\boldsymbol{u}(\boldsymbol{x},t)=\frac{1}{\rho(\boldsymbol{x})}\sum_{\varepsilon=1}^{S}\hat{\omega}(\boldsymbol{x},\boldsymbol{X}_{\varepsilon})M_{\varepsilon}\boldsymbol{U}_{\varepsilon}(t). \tag{12}$$

The result of this modified weight function is demonstrated in Fig. 1(b). One can see that the coarse-grained points (red dots) are almost aligned with the crack interface, verifying the efficacy of our modified coarse-graining method in handling the MD data set with cracks.

Similar to the derivation in Ref. [46], we point out that our coarse-grained formulation naturally induces a nonlocal equation of $\boldsymbol{u}$. The goal of the present work is therefore to identify an optimal nonlocal model in the form of peridynamics, which faithfully represents given MD displacements under a given set of loading conditions, and is generalizable to further prediction tasks for analysis of material deformation and crack propagation phenomena.

3 A peridynamics model with brittle fracture

In the previous section, a data set of function trios, $\mathcal{T}:=\{(\rho^m,\boldsymbol{u}^m,\boldsymbol{b}^m)\}_{m=1}^{M}$, were derived from our coarse-grained formulation, such that each trio contains a coarse-grained density field $\rho^m(\boldsymbol{x})$, a body force density $\boldsymbol{b}^m(\boldsymbol{x},t)$, and their corresponding displacement field $\boldsymbol{u}^m(\boldsymbol{x},t)$ for the material point $\boldsymbol{x}\in\Omega^m\subset\mathbb{R}^d$ and $t\in[0,T^m]$. Herein, we note that each sample may have different spatial domains Ω^m and observation ranges T^m. We propose to learn a nonlocal momentum balance equation based on these function trios in the form of peridynamic equation of motion[23], to provide a continuum model with direct description of fracture within the basic field equations. In peridynamics, each $\boldsymbol{x}$ interacts through bond forces with other material points $\boldsymbol{y}$ within a neighborhood with the radius δ known as the family of $\boldsymbol{x}$, denoted by $B_\delta(\boldsymbol{x})$. Here, the horizon δ determines the extent of the nonlocal interactions. The equation of motion for the material point $\boldsymbol{x}$ is given as

$$\rho(\boldsymbol{x})\frac{\partial^2\boldsymbol{u}(\boldsymbol{x},t)}{\partial t^2}=\int_{B_\delta(\boldsymbol{x})}\boldsymbol{f}(\boldsymbol{y},\boldsymbol{x},t)\mathrm{d}\boldsymbol{y}+\boldsymbol{b}(\boldsymbol{x},t). \tag{13}$$

A material model in peridynamics supplies values of $\boldsymbol{f}(\boldsymbol{y},\boldsymbol{x},t)$ in terms of the deformations of the families of $\boldsymbol{x}$ and $\boldsymbol{y}$ and any other relevant variables such as temperature[54]. Peridynamics can model fracture because the equation of motion (see Eq. (13)) is an integro-differential equation that does not involve the partial derivatives of displacement with respect to position, which leads to a lower requirement of the solution regularity. Moreover, many material models have been developed for peridynamics, and any material model from the local theory can be translated into a peridynamic form[55]. For a more thorough introduction and review about peridynamics, we refer interested readers to Ref. [31].

Peridynamics treats the internal forces in a body as a network of nonlocal bond forces. Although the stress tensor is not used explicitly in peridynamics, it is often useful to compute a stress tensor from these bond forces. This can be accomplished using the peridynamic stress tensor[56]. An approximate form of this tensor that is less general but easier to compute is called the partial stress tensor[57]. Either of these tensors, in essence, sums up the bond force vectors per unit area through any plane that contains a given point $\boldsymbol{x}$.

In this work, we aim to address the question of how to use MD to obtain a peridynamic material model that is able to treat material deformation and the nucleation and growth of fractures. With a purpose of demonstration and without loss of generality, we consider a 2D simulation problem ($d=2$) in a single-layer graphene, and concern small deformations in the linear regime of material response, although the algorithm may be generalized to finite deformations and three-dimensional (3D) cases. In Ref. [46], a data-driven 2D LPS model[54] under the plane-stress assumption was found to adequately represent the material response from MD simulations on a graphene sheet without fracture. Inspired by such preliminary results, in

this work, we also employ the LPS model as the base model, and further consider learning of material failure from coarse-grained MD data sets. In this section, we first briefly introduce the LPS model without material fracture in Subsection 3.1. Then, we extend the model to describe material fracture and handle the imposition of traction loads as fracture surfaces open up in Subsection 3.2. Then, in the next section, we will describe our learning algorithm.

3.1 LPS model

In this section, we summarize the mathematical formulation for the LPS model[58–60]. The LPS model is a prototypical state-based model which can be seen as a nonlocal extension of the linear elasticity model. It is suitable to describe isotropic elastic materials under infinitesimal deformation. Compared with the previously developed bond-based peridynamic models[23, 61], the LPS model has advantages in the fact that it is not restricted to a Poisson's ratio of 1/4, which is important for our application since Poisson's ratio of graphene is found to be negative from MD and molecular statistics simulations[62–63].

Consider a body occupying the domain $\Omega \subset \mathbb{R}^d$, and let θ be the nonlocal dilatation, generalizing the local divergence of the displacement. In this section, we consider the material without damage, with fully prescribed Dirichlet-type boundary conditions, and will further extend the discussion to more general boundary conditions and brittle fractures in Subsection 3.2. Here, we note that in nonlocal problems, unless otherwise stated, the boundary conditions should no longer be prescribed on the sharp interface $\partial\Omega$, but on a collar of thickness of at least δ surrounding the domain Ω, which we denote as

$$\mathcal{B}\Omega := \{\boldsymbol{x} \notin \Omega | \mathrm{dist}(\boldsymbol{x}, \partial\Omega) < 2\delta\}.$$

Given nonlocal boundary conditions prescribed on the nonlocal volumetric boundary domain (or simply nonlocal boundary),

$$\boldsymbol{u}(\boldsymbol{x}, t) := \boldsymbol{u}_D(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \mathcal{B}\Omega, \quad t \in [0, T],$$

and the initial velocity $\phi(\boldsymbol{x})$ and displacement $\psi(\boldsymbol{x})$ for $\boldsymbol{x} \in \Omega \cup \mathcal{B}\Omega$ at $t = 0$, the peridynamic operator in the LPS model is given by

$$\begin{aligned}\mathcal{L}_K[\boldsymbol{u}](\boldsymbol{x}, t) := &-\frac{C_1}{m}\int_{B_\delta(\boldsymbol{x})}(\lambda-\mu)K(|\boldsymbol{y}-\boldsymbol{x}|)(\boldsymbol{y}-\boldsymbol{x})(\theta(\boldsymbol{x}, t)+\theta(\boldsymbol{y}, t))\mathrm{d}\boldsymbol{y}\\ &-\frac{C_2}{m}\int_{B_\delta(\boldsymbol{x})}\mu K(|\boldsymbol{y}-\boldsymbol{x}|)\frac{(\boldsymbol{y}-\boldsymbol{x})\otimes(\boldsymbol{y}-\boldsymbol{x})}{|\boldsymbol{y}-\boldsymbol{x}|^2}(\boldsymbol{u}(\boldsymbol{y}, t)-\boldsymbol{u}(\boldsymbol{x}, t))\mathrm{d}\boldsymbol{y},\end{aligned} \tag{14}$$

and the nonlocal dilatation is defined via

$$\theta(\boldsymbol{x}, t) := \frac{d}{m}\int_{B_\delta(\boldsymbol{x})} K(|\boldsymbol{y}-\boldsymbol{x}|)(\boldsymbol{y}-\boldsymbol{x})\cdot(\boldsymbol{u}(\boldsymbol{y}, t)-\boldsymbol{u}(\boldsymbol{x}, t))\mathrm{d}\boldsymbol{y}, \tag{15}$$

where $m := \int_{B_\delta(\boldsymbol{0})} K(|\boldsymbol{z}|)|\boldsymbol{z}|^2\mathrm{d}\boldsymbol{z}$ is the weighted volume, λ is Lamé's first parameter, and μ is the shear modulus. To recover parameters for 3D linear elasticity, one should take $C_1 = 3$, $C_2 = 30$; whereas for 2D problems, $C_1 = 2$, $C_2 = 16$. Here, we note that m is determined by the horizon size δ and the influence function K. We use the subscript K in the nonlocal operator $\mathcal{L}_K[\boldsymbol{u}](\boldsymbol{x})$ to emphasize the operator's dependence on the influence function K. Then, the time-dependent LPS problem is given by

$$\begin{cases}\rho(\boldsymbol{x})\dfrac{\partial^2\boldsymbol{u}(\boldsymbol{x}, t)}{\partial t^2}+\mathcal{L}_K[\boldsymbol{u}](\boldsymbol{x}, t) = \boldsymbol{b}(\boldsymbol{x}, t), \quad (\boldsymbol{x}, t) \in \Omega\times[0, T],\\ \boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_D(\boldsymbol{x}, t), \quad (\boldsymbol{x}, t) \in \mathcal{B}\Omega\times[0, T],\\ \boldsymbol{u}(\boldsymbol{x}, 0) = \psi(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega\cup\mathcal{B}\Omega,\\ \dot{\boldsymbol{u}}(\boldsymbol{x}, 0) = \phi(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega\cup\mathcal{B}\Omega.\end{cases} \tag{16}$$

Note that our learning algorithm is compatible with other types of boundary conditions in $\mathcal{B}\Omega$. Here, we focus on the Dirichlet-type boundary condition in this work for its simplicity.

To discretize the above LPS model, we employ the optimization-based meshfree quadrature rule developed in Refs. [64]–[71]. Suppose that the values of function trios $\rho(\boldsymbol{x})$, $\boldsymbol{u}(\boldsymbol{x},t)$, and $\boldsymbol{b}(\boldsymbol{x},t)$ are provided on a set of coarse-grained material points $\chi := \{\boldsymbol{x}_i\}_{i=1}^{I} \subset \Omega \cup \mathcal{B}\Omega$ and time instances $t^n = n\Delta t$, $n = 0, \cdots, T/\Delta t$. Although the machine learning algorithm as well as the quadrature rule is compatible with the general non-uniform grids, in this work we consider the uniform grids with the grid size h and the uniform time steps with the size Δt, for simplicity. We write the discretized approximation of $\mathcal{L}_K$ as

$$\begin{aligned}\mathcal{L}_K^h[\boldsymbol{u}](\boldsymbol{x}_i,t^n) := &-\frac{C_1}{m_i}\sum_{\boldsymbol{x}_j\in B_\delta(\boldsymbol{x}_i)\cap\chi}(\lambda-\mu)K_{ij}(\boldsymbol{x}_j-\boldsymbol{x}_i)(\theta^h(\boldsymbol{x}_i,t^n)+\theta^h(\boldsymbol{x}_j,t^n))W_{j,i}\\ &-\frac{C_2}{m_i}\sum_{\boldsymbol{x}_j\in B_\delta(\boldsymbol{x}_i)\cap\chi}\mu K_{ij}\frac{(\boldsymbol{x}_j-\boldsymbol{x}_i)\otimes(\boldsymbol{x}_j-\boldsymbol{x}_i)}{|\boldsymbol{x}_j-\boldsymbol{x}_i|^2}(\boldsymbol{u}(\boldsymbol{x}_j,t^n)-\boldsymbol{u}(\boldsymbol{x}_i,t^n))W_{j,i}, \quad (17)\end{aligned}$$

$$\theta^h(\boldsymbol{x}_i,t^n) := \frac{d}{m_i}\sum_{\boldsymbol{x}_j\in B_\delta(\boldsymbol{x}_i)\cap\chi}K_{ij}(\boldsymbol{x}_j-\boldsymbol{x}_i)\cdot(\boldsymbol{u}(\boldsymbol{x}_j,t^n)-\boldsymbol{u}(\boldsymbol{x}_i,t^n))W_{j,i}, \quad (18)$$

where $K_{ij} := K(|\boldsymbol{x}_j-\boldsymbol{x}_i|)$, and $m_i := \sum_{\boldsymbol{x}_j\in B_\delta(\boldsymbol{x}_i)\cap\chi} K_{ij}|\boldsymbol{x}_j-\boldsymbol{x}_i|^2W_{j,i}$. The quadrature weights $W_{j,i}$ are associated with a local neighborhood of particles for each discretization point $\boldsymbol{x}_i$, generated by local optimizations to make the approximation rule exact for certain classes of functions. For each $\boldsymbol{x}_i \in \chi\cap\Omega$, we solve for $W_{j,i}$ via

$$\underset{\{\omega_{j,i}\}}{\operatorname{argmin}}\sum_{\boldsymbol{x}_j\in\chi\cap B_\delta(\boldsymbol{x}_i)\setminus\{\boldsymbol{x}_i\}}W_{j,i}^2 \quad \text{s.t.} \sum_{\boldsymbol{x}_j\in B_\delta(\boldsymbol{x}_i)}q(\boldsymbol{x}_i,\boldsymbol{x}_j)W_{j,i} = \int_{B_\delta(\boldsymbol{x}_i)}q(\boldsymbol{x}_i,\boldsymbol{y})\mathrm{d}\boldsymbol{y}, \quad \forall\, q\in\boldsymbol{V}_{\boldsymbol{x}_i}, \quad (19)$$

where "s.t." represents "subject to", and $\boldsymbol{V}_{\boldsymbol{x}_i}$ denotes the space of functions which should be integrated exactly. Following Ref. [65], in this work, we take $\boldsymbol{V}_{\boldsymbol{x}_i} := \big\{q(\boldsymbol{y}-\boldsymbol{x}_i) = \frac{p(\boldsymbol{y}-\boldsymbol{x}_i)}{|\boldsymbol{y}-\boldsymbol{x}_i|^3}\big| p\in \boldsymbol{P}_5(\mathbb{R}^d)$ such that $\int_{B_\delta(\boldsymbol{0})}q(\boldsymbol{y})\mathrm{d}\boldsymbol{y}<\infty\big\}$, and denote $\boldsymbol{P}_5(\mathbb{R}^d)$ as the space of quintic polynomials. As the horizon size δ vanishes, this discretization preserves the consistency in the limit to the local solution[64–65]. Moreover, we point out that the quadrature weights, $W_{j,i}$, only depend on the grid set χ, and they are invariant of the influence function K. Hence, in our learning algorithm, one only needs to generate the quadrature weights and solves the local optimization problem (19) once in the preprocessing step.

For the dynamic peridynamics model, to discretize in time we apply the central difference time stepping scheme. With the time step size Δt, at the $(n+1)$th time step, one can solve for the displacement $\boldsymbol{u}_i^{n+1} \approx \boldsymbol{u}(\boldsymbol{x}_i,t^{n+1})$ following

$$\begin{cases}\rho(\boldsymbol{x}_i)\ddot{\boldsymbol{u}}_i^n + \mathcal{L}_K^h[\boldsymbol{u}](\boldsymbol{x}_i,t^n) = \boldsymbol{b}(\boldsymbol{x}_i,n\Delta t) \quad \text{for} \quad \boldsymbol{x}_i \quad \text{in} \quad \Omega\cap\chi,\\ \boldsymbol{u}_i^{n+1} = \boldsymbol{u}_D(\boldsymbol{x}_i,(n+1)\Delta t) \quad \text{for} \quad \boldsymbol{x}_i \quad \text{in} \quad \mathcal{B}\Omega\cap\chi,\end{cases} \quad (20)$$

where $\mathcal{L}_K^h$ is the discretized nonlocal operator as defined in Eq. (17), and the acceleration $\ddot{\boldsymbol{u}}_i^n$ is estimated via the central difference scheme,

$$\ddot{\boldsymbol{u}}_i^n := \frac{\boldsymbol{u}_i^{n+1}-2\boldsymbol{u}_i^n+\boldsymbol{u}_i^{n-1}}{\Delta t^2}. \quad (21)$$

As the initial conditions, we set $\boldsymbol{u}_i^0 = \psi(\boldsymbol{x}_i)$ and $\dfrac{\boldsymbol{u}_i^1-\boldsymbol{u}_i^0}{\Delta t} = \phi(\boldsymbol{x}_i)$ for $\boldsymbol{x}_i \in (\mathcal{B}\Omega\cup\Omega)\cap\chi$.

3.2 Peridynamics formulation for brittle fractures

One of the main appeals of peridynamics is to handle fracture problems, where free surfaces are associated with the evolution of a fracture surface. In this section, we consider the LPS model with free surfaces, then apply it to the treatment of brittle fractures.

To describe the free surfaces associated with the time evolution of a fracture surface, we now consider general mixed boundary conditions: $\partial\Omega = \partial\Omega_D \cup \partial\Omega_N$ and $(\partial\Omega_D)^{\circ} \cap (\partial\Omega_N)^{\circ} = \varnothing$. Here, $\partial\Omega_D$ and $\partial\Omega_N$ are both curves. $\partial\Omega_N$ is the (possibly time-dependent) sharp crack surface evolving with the material fractures, and a free surface boundary condition is applied on it. To define a Dirichlet-type constraint, we denote

$$\mathcal{B}\Omega_D := \{\boldsymbol{x} \notin \Omega | \mathrm{dist}(\boldsymbol{x}, \partial\Omega_D) < 2\delta\},$$

and assume that the value of $\boldsymbol{u}(\boldsymbol{x}, t) = \boldsymbol{u}_D(\boldsymbol{x}, t)$ is given on $\boldsymbol{x} \in \mathcal{B}\Omega_D$. For notation simplicity, we denote $\Omega_D := \Omega \cup \mathcal{B}\Omega_D$. To apply the free surface boundary condition, we denote

$$\mathcal{I}\Omega_N := \{\boldsymbol{x} \in \Omega | \mathrm{dist}(\boldsymbol{x}, \partial\Omega_N) < \delta\}, \quad \mathcal{I}\Omega := \{\boldsymbol{x} \in \Omega | \mathrm{dist}(\boldsymbol{x}, \partial\Omega) < \delta\}.$$

Unless stated otherwise, in this paper we further assume sufficient regularity in the boundary region $\mathcal{I}\Omega$ that there exists a unique orthogonal projection of $\boldsymbol{x}$ onto $\partial\Omega$, which is the closest point on $\partial\Omega$ to $\boldsymbol{x}$, and we denote this projection as $\overline{\boldsymbol{x}}$. Then, one has $\overline{\boldsymbol{x}} - \boldsymbol{x} = s_x \boldsymbol{n}(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{I}\Omega_N$, where $0 < s_x < \delta$. Here, $\boldsymbol{n}$ denotes the normal direction pointing out of the domain for each $\boldsymbol{x} \in \mathcal{I}\Omega_N$, and let $\boldsymbol{p}$ denote the tangential direction. In our numerical solver, we treat $\boldsymbol{x}$ with the free surface boundary condition if the projection of $\boldsymbol{x}$ is in $\partial\Omega_N$. Otherwise, we use the Dirichlet-type boundary condition at $\boldsymbol{x}$.

In peridynamics, material damage is incorporated into the constitutive model by allowing the bonds of material points to break irreversibly. To model brittle fracture in the LPS model, we employ a smoothed critical stretch criterion, where weakening occurs when a bond is extended beyond some predetermined critical bond deformed length[39, 65, 72]. In particular, a scalar state function $\gamma(\boldsymbol{x}, \boldsymbol{y}, t)$ is defined and takes values in the interval $[0, 1]$, to describe the bond weakening and breakage through the crack growing,

$$\gamma(\boldsymbol{x}, \boldsymbol{y}, t) := \frac{1}{2}\Big(- \tanh\Big(\frac{\max_{\tau\in[0,t]} S(\boldsymbol{x}, \boldsymbol{y}, \tau) - s_0}{\eta}\Big) + 1\Big), \tag{22}$$

where

$$S(\boldsymbol{x}, \boldsymbol{y}, \tau) := \frac{|\boldsymbol{x} - \boldsymbol{y} + \boldsymbol{u}(\boldsymbol{x}, \tau) - \boldsymbol{u}(\boldsymbol{y}, \tau)|}{|\boldsymbol{x} - \boldsymbol{y}|} - 1, \tag{23}$$

and s_0 is the critical stretch criterion depending on the material. $\gamma(\boldsymbol{x}, \boldsymbol{y}, t)$ is a history-dependent function, i.e., a bond can never recover once it exceeds the critical stretch criterion. An illustration of γ can be visualized in Fig. 3, where a hyperparameter $\eta \ll 1$ can be tuned to control the level of smoothness. When $\gamma(\boldsymbol{x}, \boldsymbol{y}, t) = 1$, the bond between material points $\boldsymbol{x}$ and $\boldsymbol{y}$ is considered "intact", and the change of displacement on material point $\boldsymbol{y}$ may have an impact on the displacement at $\boldsymbol{x}$. When the stretch $S(\boldsymbol{x}, \boldsymbol{y}, \tau)$ exceeds the critical criterion s_0 for some time $\tau < t$, the material gets damaged, and we have $\gamma(\boldsymbol{x}, \boldsymbol{y}, t) < 1$. As the stretch further increases, finally $\gamma(\boldsymbol{x}, \boldsymbol{y}, t) = 0$, and we consider the bonds between $\boldsymbol{x}$ and $\boldsymbol{y}$ as fully "broken". Instead of defining γ as a step function following Ref. [65], in Eq. (22) we allow the weakening of force scalar within small ranges of excessive bond stretch values, and set γ as a smoothed step function. As shown in Ref. [39], such a smoothed state function would impose the continuity of the learnt bond force $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{y}, t)$ in our peridynamics model, and guarantee the well-posedness of the peridynamics model as a dynamic system. On the other hand, a continuous formulation of the damage factor γ would result in a continuous optimization problem, and allow generic optimization routines to be used in the training procedure.

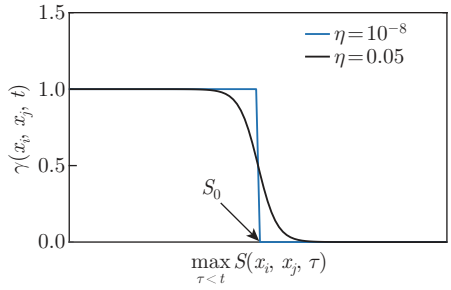


Fig. 3 An illustration of the smoothed scalar state function γ with the tunable parameter $\eta = \{0.05, 10^{-8}\}$ (color online)

With the state function γ, we treat the time-evolving fracture as free surfaces and employ the following formulations:

$$\begin{cases} \rho(x)\dfrac{\partial^2 \boldsymbol{u}(\boldsymbol{x},t)}{\partial t^2} + \mathcal{L}_{KN}[\boldsymbol{u}](\boldsymbol{x},t) = \boldsymbol{b}(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \Omega \times [0,T], \\ \boldsymbol{u}(\boldsymbol{x},t) = \boldsymbol{u}_D(\boldsymbol{x},t), \quad (\boldsymbol{x},t) \in \mathcal{B}\Omega_D \times [0,T], \\ \boldsymbol{u}(\boldsymbol{x},0) = \psi(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega \cup \mathcal{B}\Omega, \\ \dot{\boldsymbol{u}}(\boldsymbol{x},0) = \phi(\boldsymbol{x}), \quad \boldsymbol{x} \in \Omega \cup \mathcal{B}\Omega. \end{cases} \tag{24}$$

Here, the modified LPS operator $\mathcal{L}_{KN}$ follows the formulation[65, 71],

$$\begin{aligned} \mathcal{L}_{KN}[\boldsymbol{u}](\boldsymbol{x},t) := & -\frac{C_1}{m}\int_{B_\delta(\boldsymbol{x})} (\lambda-\mu)K(|\boldsymbol{y}-\boldsymbol{x}|)\gamma(\boldsymbol{x},\boldsymbol{y},t)(\boldsymbol{y}-\boldsymbol{x})(\theta^{\mathrm{corr}}(\boldsymbol{x},t)+\theta^{\mathrm{corr}}(\boldsymbol{y},t))\mathrm{d}\boldsymbol{y} \\ & -\frac{C_2}{m}\int_{B_\delta(\boldsymbol{x})} \mu K(|\boldsymbol{y}-\boldsymbol{x}|)\gamma(\boldsymbol{x},\boldsymbol{y},t)\frac{(\boldsymbol{y}-\boldsymbol{x})\otimes(\boldsymbol{y}-\boldsymbol{x})}{|\boldsymbol{y}-\boldsymbol{x}|^2}(\boldsymbol{u}(\boldsymbol{y},t)-\boldsymbol{u}(\boldsymbol{x},t))\mathrm{d}\boldsymbol{y} \\ & -\frac{2C_1\theta^{\mathrm{corr}}(\boldsymbol{x},t)}{m}\int_{B_\delta(\boldsymbol{x})} (\lambda-\mu)K(|\boldsymbol{y}-\boldsymbol{x}|)(1-\gamma(\boldsymbol{x},\boldsymbol{y},t))(\boldsymbol{y}-\boldsymbol{x})\mathrm{d}\boldsymbol{y} \\ & -\frac{C_2\theta^{\mathrm{corr}}(\boldsymbol{x},t)}{2m}\int_{B_\delta(\boldsymbol{x})} (\lambda+2\mu)K(|\boldsymbol{y}-\boldsymbol{x}|)(1-\gamma(\boldsymbol{x},\boldsymbol{y},t)) \\ & \cdot\frac{((\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{n})((\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{p})^2}{|\boldsymbol{y}-\boldsymbol{x}|^2}\boldsymbol{n}\mathrm{d}\boldsymbol{y} \\ & +\frac{C_2\theta^{\mathrm{corr}}(\boldsymbol{x},t)}{2m}\int_{B_\delta(\boldsymbol{x})} \lambda K(|\boldsymbol{y}-\boldsymbol{x}|)(1-\gamma(\boldsymbol{x},\boldsymbol{y},t))\frac{((\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{n})^3}{|\boldsymbol{y}-\boldsymbol{x}|^2}\boldsymbol{n}\mathrm{d}\boldsymbol{y} \end{aligned} \tag{25}$$

with

$$\theta^{\mathrm{corr}}(\boldsymbol{x},t) := \frac{d}{m}\int_{B_\delta(\boldsymbol{x})} K(|\boldsymbol{y}-\boldsymbol{x}|)\gamma(\boldsymbol{x},\boldsymbol{y},t)(\boldsymbol{y}-\boldsymbol{x})\cdot\boldsymbol{M}(\boldsymbol{x})\cdot(\boldsymbol{u}(\boldsymbol{y},t)-\boldsymbol{u}(\boldsymbol{x},t))\mathrm{d}\boldsymbol{y}, \tag{26}$$

$$\boldsymbol{M}(\boldsymbol{x},t) := \Big(\frac{d}{m}\int_{B_\delta(\boldsymbol{x})} K(|\boldsymbol{y}-\boldsymbol{x}|)\gamma(\boldsymbol{x},\boldsymbol{y},t)\,(\boldsymbol{y}-\boldsymbol{x})\otimes(\boldsymbol{y}-\boldsymbol{x})\,\mathrm{d}\boldsymbol{y}\Big)^{-1}. \tag{27}$$

As such, the LPS model provides an approximation for the corresponding linear elastic model with free surfaces in the case of linear displacement fields. We notice that when all bonds in $B_\delta(\boldsymbol{x})$ are intact, i.e., the material point $\boldsymbol{x}$ is sufficiently far away from the free surface, we have $\gamma(\boldsymbol{x},\boldsymbol{y},t) = 1$ for all $\boldsymbol{y} \in B_\delta(\boldsymbol{x})$. Then, Eq. (25) yields $\mathcal{L}_{KN} = \mathcal{L}_K$, and the original momentum balance and nonlocal dilatation formulation in the LPS model are obtained. Therefore, Eq. (25)

provides a unified mathematical framework which automatically captures material deformation and the evolution of cracks as free surfaces.

We now extend the optimization-based quadrature rule and the central difference time-stepping method introduced in Subsection 3.1, to the LPS model (25) with fracture. Particularly, at the $(n+1)$th time step, we approximate the state function $\gamma(\boldsymbol{x}_i, \boldsymbol{x}_j, t^n)$ via

$$\gamma_{ij}^n := \frac{1}{2}\Big(- \tanh\Big(\frac{\max\limits_{0\leqslant m\leqslant n} S_{ij}^m - s_0}{\eta}\Big) + 1\Big), \tag{28}$$

where $S_{ij}^m := \frac{|\boldsymbol{x}_i - \boldsymbol{x}_j + \boldsymbol{u}_i^m - \boldsymbol{u}_j^m|}{|\boldsymbol{x}_i - \boldsymbol{x}_j|} - 1$. Then, the approximated displacement field $\boldsymbol{u}_i^{n+1} \approx \boldsymbol{u}(\boldsymbol{x}_i, t^{n+1})$ can be solved via the following formulations:

$$\begin{cases} \rho(\boldsymbol{x}_i)\ddot{\boldsymbol{u}}_i^n + \mathcal{L}_{KN}^h[\boldsymbol{u}](\boldsymbol{x}_i, t^n) = \boldsymbol{b}(\boldsymbol{x}_i, n\Delta t) \quad \text{for} \quad \boldsymbol{x}_i \quad \text{in} \quad \Omega \cap \chi, \\ \boldsymbol{u}_i^{n+1} = \boldsymbol{u}_D(\boldsymbol{x}_i, (n+1)\Delta t) \quad \text{for} \quad \boldsymbol{x}_i \quad \text{in} \quad \mathcal{B}\Omega_D \cap \chi, \end{cases} \tag{29}$$

where

$$\begin{aligned} \mathcal{L}_{KN}^h[\boldsymbol{u}](\boldsymbol{x}_i, t^n) := & - \frac{C_1}{m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} (\lambda - \mu) K_{ij} \gamma_{ij}^n (\boldsymbol{x}_j - \boldsymbol{x}_i) \left((\theta^{\text{corr}})_i^n + (\theta^{\text{corr}})_j^n \right) W_{j,i} \\ & - \frac{C_2}{m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} \mu K_{ij} \gamma_{ij}^n \frac{(\boldsymbol{x}_j - \boldsymbol{x}_i) \otimes (\boldsymbol{x}_j - \boldsymbol{x}_i)}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^2} (\boldsymbol{u}_j^n - \boldsymbol{u}_i^n) W_{j,i} \\ & - \frac{2C_1 (\theta^{\text{corr}})_i^n}{m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} (\lambda - \mu) K_{ij} (1 - \gamma_{ij}^n)(\boldsymbol{x}_j - \boldsymbol{x}_i) W_{j,i} \\ & - \frac{C_2 (\theta^{\text{corr}})_i^n}{2m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} (\lambda + 2\mu) K_{ij} (1 - \gamma_{ij}^n) \\ & \cdot \frac{((\boldsymbol{x}_j - \boldsymbol{x}_i) \cdot \boldsymbol{n}_i^n)((\boldsymbol{x}_j - \boldsymbol{x}_i) \cdot \boldsymbol{p}_i^n)^2}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^2} \boldsymbol{n}_i^n W_{j,i} \\ & + \frac{C_2 (\theta^{\text{corr}})_i^n}{2m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} \lambda K_{ij} (1 - \gamma_{ij}^n) \frac{((\boldsymbol{x}_j - \boldsymbol{x}_i) \cdot \boldsymbol{n}_i^n)^3}{|\boldsymbol{x}_j - \boldsymbol{x}_i|^2} \boldsymbol{n}_i^n W_{j,i} \end{aligned} \tag{30}$$

with

$$(\theta^{\text{corr}})_i^n := \frac{d}{m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} K_{ij} \gamma_{ij}^n (\boldsymbol{x}_j - \boldsymbol{x}_i) \cdot \boldsymbol{M}_i^n \cdot (\boldsymbol{u}_j^n - \boldsymbol{u}_i^n) W_{j,i}, \tag{31}$$

$$\boldsymbol{M}_i^n := \Big(\frac{d}{m_i} \sum_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)\cap\chi} K_{ij} \gamma_{ij}^n (\boldsymbol{x}_j - \boldsymbol{x}_i) \otimes (\boldsymbol{x}_j - \boldsymbol{x}_i) W_{j,i} \Big)^{-1}. \tag{32}$$

Here, we note that both the free surface $\partial\Omega_N$ and the normal vector $\boldsymbol{n}(\boldsymbol{x})$ on free surfaces change as the fracture evolves. To numerically approximate $\boldsymbol{n}(\boldsymbol{x}_i, t^n)$ at each time step, we update it via

$$\boldsymbol{n}_i^n = - \frac{\sum\limits_{\boldsymbol{x}_j \in \chi \cap B_\delta(\boldsymbol{x}_i)} (\boldsymbol{x}_j - \boldsymbol{x}_i) W_{j,i} \gamma_{ij}^n}{\Big| \sum\limits_{\boldsymbol{x}_j \in \chi \cap B_\delta(\boldsymbol{x}_i)} (\boldsymbol{x}_j - \boldsymbol{x}_i) W_{j,i} \gamma_{ij}^n \Big|}, \tag{33}$$

and the tangential vector $\boldsymbol{p}_i^n$ is calculated as the orthogonal direction to $\boldsymbol{n}_i^n$. The correction tensor should be invertible to ensure that the correction dilatation can be computed. This holds as long as the bonds in the horizon are non-colinear. For fracture cases resulting in bond break, leaving an isolated particle, we replace the matrix inverse with the pseudo-inverse.

4 Learning algorithm

Let $\mathcal{T} := \{\rho^m(\boldsymbol{x}_{i,m}), \boldsymbol{u}^m(\boldsymbol{x}_{i,m}, t_m^n), \boldsymbol{b}^m(\boldsymbol{x}_{i,m}, t_m^n)\}$, $m = 1, 2, \cdots, M$, be coarse-grained function trios available at $\boldsymbol{x}_{i,m} \in \chi_m$ and $t_m^n = n\Delta t_m$, $n = 1, 2, \cdots, N_m$. Our goal is to identify an optimal constitutive relation on the basis of MD data sets. Here, we use χ_m and Δt_m to highlight the fact that in our learning algorithm, each sample can be of different spatial/temporal domains and resolutions. In the following content, we will skip the subscript m and denote the function trios as $\rho^m(\boldsymbol{x}_i)$, $\boldsymbol{u}^m(\boldsymbol{x}_i, t^n)$, and $\boldsymbol{b}^m(\boldsymbol{x}_i, t^n)$ for simplicity. Let $\mathcal{L}_{KN}$ be the LPS operator defined in Eq. (30). Then, we aim to learn an optimal continuum model in the form of LPS models, where the optimal model consists of the influence function K, which may be sign-changing, and the parameters λ, μ, and s_0, such that the action of $\mathcal{L}_{KN}$ most closely satisfies (30) for all s. Formally, the optimal influence function and parameters $(\lambda^*, \mu^*, s_0^*, K^*)$ are the solutions to the following optimization problem:

$$\begin{aligned}&(\lambda^*, \mu^*, s_0^*, K^*)\\ &= \underset{\lambda,\mu,s_0,K}{\operatorname{argmin}} \frac{1}{M} \sum_{m=1}^{M} \sum_{n=1}^{N_m-1} \Delta t_m \big\| \rho^m(\boldsymbol{x}_i)(\ddot{\boldsymbol{u}}^m)_i^n + \mathcal{L}_{KN}^h[\boldsymbol{u}^m](\boldsymbol{x}_i, t^n) - \boldsymbol{b}^m(\boldsymbol{x}_i, t^n) \big\|_{\ell^2(\chi_m)}^2. \end{aligned} \tag{34}$$

The influence function $K(|\boldsymbol{x} - \boldsymbol{y}|)$ will now be parameterized. Following Ref. [73], in this work, the interacting kernel function $K(|\boldsymbol{x} - \boldsymbol{y}|)$ is taken as a radial function compactly supported on the δ-ball $B_\delta(\boldsymbol{x})$ with αth-order singularity,

$$K(|\boldsymbol{x} - \boldsymbol{y}|) = \sum_{k=0}^{P} \frac{D_k}{|\boldsymbol{x} - \boldsymbol{y}|^\alpha} B_{k,P}\Big(\frac{|\boldsymbol{x} - \boldsymbol{y}|}{\delta}\Big). \tag{35}$$

Here, the Bernstein polynomials are defined as

$$B_{k,P}(r) = \binom{P}{k} r^k (1 - r)^{P-k} \quad \text{for} \quad 0 \leqslant r \leqslant 1. \tag{36}$$

Following the arguments in Refs. [66] and [46], in the learning algorithm, we require the fractional order α to be bounded by 3 and allow $D_k \in \mathbb{R}$ for all k with sufficient well-posedness conditions embedded for the discretized operator. Here, we note that in the samples with material fracture, some particles might become isolated due to fragmentation, and hence it would be impossible to require solvability constraints. Therefore, we only apply the solvability constraints to the model without fracture. With the analysis in Ref. [46], given a tolerance parameter $\zeta > 0$, we apply the following solvability constraints:

$$\begin{cases} \lambda + \mu > 0,\ \mu > 0,\ \alpha < 3,\ \Lambda_{\min}(\Gamma_{(\alpha,\boldsymbol{D},\delta,P)}) \geqslant \zeta, \\ \Lambda_{\min}(\Phi_{(\alpha,\boldsymbol{D},\delta,P)} \Gamma^\dagger_{(\alpha,\boldsymbol{D},\delta,P)} \Phi^t_{(\alpha,\boldsymbol{D},\delta,P)}) \geqslant \zeta, \\ \Lambda_{\min}(\Gamma_{(\alpha,\boldsymbol{D},\delta,P)} - 2\Phi^t_{(\alpha,\boldsymbol{D},\delta,P)} \Phi_{(\alpha,\boldsymbol{D},\delta,P)}) \geqslant 0. \end{cases} \tag{37}$$

Here, Γ and Φ are the matrices that correspond to the deviatoric and dilatation contributions of the deformation, and $\Lambda_{\min}(\boldsymbol{A})$ denotes the smallest nonzero eigenvalue of a matrix $\boldsymbol{A}$.

The overall formulation of the constrained optimization problem is as follows. Given a collection of training samples $\{\rho^m(\boldsymbol{x}_i), \boldsymbol{u}^m(\boldsymbol{x}_i, t^n), \boldsymbol{b}^m(\boldsymbol{x}_i, t^n)\}$, $m = 1, 2, \cdots, M$, we seek to learn the parameters λ and μ, the Bernstein polynomial coefficients $\boldsymbol{D} = [D_0, D_1, \cdots, D_P] \in \mathbb{R}^{P+1}$, the order α, the horizon δ, the polynomial order P, and the damage criterion s_0, by

minimizing the mean square loss (MSL) of the LPS equation,

$$\begin{cases} (\lambda^*, \mu^*, \boldsymbol{D}^*, \alpha^*, \delta^*, P^*, s_0^*) = \underset{\lambda,\mu,\boldsymbol{D},\alpha,\delta,P,s_0}{\operatorname{argmin}} \dfrac{1}{M} \displaystyle\sum_{m=1}^{M} \sum_{n=1}^{N_m-1} \Delta t_m \big\| \rho^m(\boldsymbol{x}_i)(\ddot{\boldsymbol{u}}^m)_i^n \\ \qquad\qquad + \mathcal{L}_{KN}^h[\boldsymbol{u}^m](\boldsymbol{x}_i, t^n) - \boldsymbol{b}^m(\boldsymbol{x}_i, t^n) \big\|_{\ell^2(\chi_m)}^2 \\ \text{s.t. solvability constraints (37).} \end{cases} \tag{38}$$

However, numerically solving the constraint optimization problem (37) could be time-consuming and possibly unstable, due to three factors. First, as shown in Fig. 3, when s_0 is away from the optimal value, its impact on the loss function would be relatively flattened, causing the vanishing gradient issue in optimizers. Second, the update of s_0 would induce the change of correction operator (30), which increases the computational cost on each epoch. Lastly, the imposition of solvability constraints (37) would also be expensive, since it involves additional calculations (such as with the projection method) and/or subiterations (such as with the augmented Lagrangian method), together with the evaluation of eigenvalues at each epoch. To make the optimization algorithm more efficient and robust, we propose to separate the solving procedure of the damage criterion s_0 with other parameters, and propose a "two-stage" strategy. Key components are summarized in Algorithm 1. In particular, we notice that the correction operator (30) and the damage criterion s_0 are only associated with samples with material fractures, while the influence function K and other material parameters can be inferred from samples without fracture. Therefore, we divide the training data set into two sets,

$$\mathcal{T}^{\text{Non-Frac}} := \{\rho^m(\boldsymbol{x}_i), \boldsymbol{u}^m(\boldsymbol{x}_i, t^n), \boldsymbol{b}^m(\boldsymbol{x}_i, t^n)\}, \quad m = 1, 2, \cdots, M^{\text{Non-Frac}},$$

which includes all samples without fracture, and

$$\mathcal{T}^{\text{Frac}} := \{\widetilde{\rho}^m(\boldsymbol{x}_i), \widetilde{\boldsymbol{u}}^m(\boldsymbol{x}_i, t^n), \widetilde{\boldsymbol{b}}^m(\boldsymbol{x}_i, t^n)\}, \quad m = 1, 2, \cdots, M^{\text{Frac}}$$

for training samples with fracture. Then, the optimization problem (38) is also split into a non-fracture kernel learning step and a damage criterion learning step.

For the kernel learning step, we infer the influence function K and Lamé's moduli λ and μ by solving a constraint optimization problem from $\mathcal{T}^{\text{Non-Frac}}$,

$$\begin{cases} (\lambda^*, \mu^*, \boldsymbol{D}^*, \alpha^*, \delta^*, P^*) = \underset{\lambda,\mu,\boldsymbol{D},\alpha,\delta,P}{\operatorname{argmin}} \ \text{Res}(\mathcal{T}^{\text{Non-Frac}}) \\ \text{s.t. solvability constraints (37),} \end{cases} \tag{39}$$

where

$$\begin{aligned} &\text{Res}(\mathcal{T}^{\text{Non-Frac}}) \\ &:= \frac{1}{M^{\text{Non-Frac}}} \sum_{m=1}^{M^{\text{Non-Frac}}} \sum_{n=1}^{N_m-1} \Delta t_m \big\| \rho^m(\boldsymbol{x}_i)(\ddot{\boldsymbol{u}}^m)_i^n + \mathcal{L}_K^h[\boldsymbol{u}^m](\boldsymbol{x}_i, t^n) - \boldsymbol{b}^m(\boldsymbol{x}_i, t^n) \big\|_{\ell^2(\chi_m)}^2. \end{aligned} \tag{40}$$

As such, one only has to evaluate the nonlocal operator without fracture following Eq. (17), which is computationally more efficient. In this step, we treat δ and P as hyperparameters to be separately tuned, to achieve the best learning accuracy without overfitting. For each combination of δ and P, the Adam optimizer in PyTorch is employed, together with the augmented Lagrangian method to impose the inequality constraints. For further details of the optimization algorithm and settings, we refer interested readers to Ref. [46].

In the damage criterion learning step, we fix the learnt parameters $(\lambda^*, \mu^*, \boldsymbol{D}^*, \alpha^*, \delta^*, P^*)$ and search for the optimal s_0 by considering an unconstraint optimization problem on $\mathcal{T}^{\text{Frac}}$,

$$s_0^* = \underset{s_0}{\text{argmin}} \ \widetilde{\text{Res}}(\mathcal{T}^{\text{Frac}}), \tag{41}$$

where

$$\begin{aligned}&\widetilde{\text{Res}}(\mathcal{T}^{\text{Frac}})\\ &:= \frac{1}{M^{\text{Frac}}} \sum_{m=1}^{M^{\text{Frac}}} \sum_{n=1}^{N_m-1} \Delta t_m \big\| \widetilde{\rho}^m(\boldsymbol{x}_i)(\ddot{\widetilde{\boldsymbol{u}}}^m)_i^n + \mathcal{L}_{NK}^h[\widetilde{\boldsymbol{u}}^m](\boldsymbol{x}_i, t^n) - \widetilde{\boldsymbol{b}}^m(\boldsymbol{x}_i, t^n) \big\|_{\ell^2(\chi_m)}^2. \end{aligned} \tag{42}$$

Algorithm 1 Workflow for learning the LPS model from MD data sets

(i) To obtain samples without material fracture, generate relatively small MD displacements on fine grids $\{\boldsymbol{X}_\varepsilon^m\}$ using different external forces and domain configurations, then group the samples into two data sets, $\mathcal{M}_{\text{train}}^{\text{Non-Frac}}$ for training the nonlocal kernel and $\mathcal{M}_{\text{val}}^{\text{Non-Frac}}$ for hyperparameter tuning,

$$\mathcal{M}_{\text{train/val}}^{\text{Non-Frac}} := \{M_\varepsilon^m, \boldsymbol{U}_\varepsilon^m(t), \boldsymbol{B}_\varepsilon^m(t)\}, \quad m = 1, 2, \cdots, M_{\text{train/val}}^{\text{Non-Frac}}.$$

(ii) Generate MD displacement samples with material fracture, on fine grids $\{\widetilde{\boldsymbol{X}}_\varepsilon^m\}$ using different external forces and domain configurations, then group the samples into two data sets, $\mathcal{M}_{\text{train}}^{\text{Frac}}$ for training the damage criterion and $\mathcal{M}_{\text{test}}^{\text{Frac}}$ for test,

$$\mathcal{M}_{\text{train/test}}^{\text{Frac}} := \{\widetilde{M}_\varepsilon^m, \widetilde{\boldsymbol{U}}_\varepsilon^m(t), \widetilde{\boldsymbol{B}}_\varepsilon^m(t)\}, \quad m = 1, 2, \cdots, M_{\text{train/test}}^{\text{Frac}}.$$

(iii) Coarse grain the data sets $\mathcal{M}_{\text{train/val}}^{\text{Non-Frac}}$ and $\mathcal{M}_{\text{train/test}}^{\text{Frac}}$, then evaluate the coarse-grained data at coarser grids χ_m to obtain the function trio sets,

$$\begin{aligned}\mathcal{T}_{\text{train/val}}^{\text{Non-Frac}} &:= \{\rho^m(\boldsymbol{x}_i), \boldsymbol{u}^m(\boldsymbol{x}_i, t^n), \boldsymbol{b}^m(\boldsymbol{x}_i, t^n)\}, \quad m = 1, 2, \cdots, M_{\text{train/val}}^{\text{Non-Frac}},\\ \mathcal{T}_{\text{train/test}}^{\text{Frac}} &:= \{\widetilde{\rho}^m(\boldsymbol{x}_i), \widetilde{\boldsymbol{u}}^m(\boldsymbol{x}_i, t^n), \widetilde{\boldsymbol{b}}^m(\boldsymbol{x}_i, t^n)\}, \quad m = 1, 2, \cdots, M_{\text{train/test}}^{\text{Frac}}.\end{aligned}$$

(iv) (Kernel learning step) Solve the optimization problem based on the non-fracture data set $\mathcal{T}_{\text{train}}^{\text{Non-Frac}}$,

$$\begin{cases} (\lambda^*, \mu^*, \boldsymbol{D}^*, \alpha^*) = \underset{\lambda, \mu, \boldsymbol{D}, \alpha}{\text{argmin}} \ \text{Res}(\mathcal{T}_{\text{train}}^{\text{Non-Frac}}) \\ \text{s.t. solvability constraints (37)}, \end{cases}$$

and tune the hyperparameters δ^* and P^*, to minimize the test errors on the validation data set $\mathcal{T}_{\text{val}}^{\text{Non-Frac}}$.

(v) (Damage criterion learning step) With fixed parameters (λ^*, μ^*, $\boldsymbol{D}^*$, α^*, δ^*, P^*), train for the optimal fracture criterion parameter based on the fracture data sets $\mathcal{T}_{\text{train}}^{\text{Frac}}$,

$$s_0^* = \underset{s_0}{\text{argmin}} \ \widetilde{\text{Res}}(\mathcal{T}^{\text{Frac}}).$$

(vi) To study the generalizability on unseen external forces and fracture scenarios, use the learnt LPS model to predict the material deformation and fracture on $\mathcal{T}_{\text{test}}^{\text{Frac}}$.

In all tests, we set the smoothing parameter $\eta = 0.05$, and employ the bisection method to solve for s_0^*.

In this study, all training and tests were performed on a 26-core 2.1 GHz Xeon Gold 6230R processor. The most time-consuming step in training is the kernel learning step, since we need

to calculate the eigenvalues of three matrices of sizes $I \times I$ and $2I \times 2I$ (where $I = \#(\chi)$ is the total number of coarse-grained material points), so as to impose the solvability constraints in Eq. (37). Therefore, the computational complexity scales as $O(I^3)$. The training time cost can become extremely large and even non-feasible if we keep refining the mesh resolution and correspondingly increasing the number of coarse-grained material points. However, we point out that since a continuous kernel is obtained, our model is resolution-independent. As such, the learnt nonlocal model on a coarse grid can generalize and provide consistent simulations on finer grids. Therefore, training on a very fine mesh is often unnecessary. Such mesh generalization performances are verified in the empirical studies of the previous work[46] and Subsection 5.4 of this work.

5 Application to single-layer graphene

To illustrate the capability of our method in obtaining an optimal surrogate material damage model from coarse-grained MD displacements, we consider single-layer graphene sheets as the application. Graphene is a single layer of carbon atoms, tightly bound in a hexagonal honeycomb lattice. Up to now, much of what has been known about the mechanical and electronic properties of graphene is based on models on the atomistic scale, such as the MD simulations. However, the use of MD directly to treat the deformation and failure of materials at the mesoscale is still largely beyond reach. Hence, we aim to learn a peridynamics model by upscaling from MD to a continuum scale.

For the present study, an MD model is created using the Tersoff interatomic potential[74], a widely used potential in the MD community for graphene[43]. Unstressed graphene nominally has an interatomic spacing of 1.46 Å (1 Å = 0.1 nm). Unless otherwise stated, in this study, the values of the coarse-grained data trios are evaluated on a square lattice of nodes with spacing $h = 5.0$ Å. The only exception is in Subsection 5.4, where we also consider an additional, finer data set generated with spacing 3.17 Å, to assess the generalization properties of the proposed learning approach to different grids. Without loss of generality, in this work, we consider MD simulations on the temperature 0 K. In all cases, external loading is applied to the atoms in the MD grid. For the non-fracture data sets, the magnitude of the loading is chosen so that the bond strains are no larger than 2%, which is less than the strains at which nonlinear effects appear. In all MD experiments, the atoms are initialized with positions on a hexagonal lattice in the x_1x_2-plane with an interatomic spacing of 1.46 Å. The mass of each atom is 2×10^{-26} kg. For purposes of computing stresses, the thickness of the lattice is set to 3.35 Å, which is the approximate distance between layers in multi-layer graphene. On quasi-static data sets, we smooth the MD simulation results in time as described in Ref. [46]. For the dynamic data sets, the MD time step size is set as 4.95×10^{-14} s.

5.1 Data generation and learning results

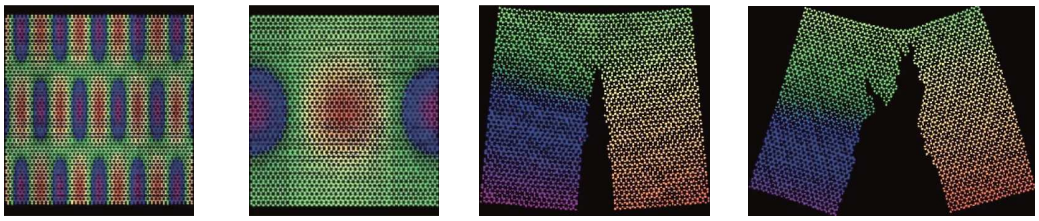
In this section, we apply the learning algorithm described in Section 4, to extract a coarse-grained model from MD simulations of a graphene sheet at 0 K. For the purpose of training, validation, and test, we generate the following four groups of MD simulations, with exemplar images showing contours of U_1, the component of atomic displacement in the x_1-direction, provided in Fig. 4.

(I) Non-fracture training data set ($\mathcal{M}^{\text{Non-Frac}}_{\text{train}}$ with 70 quasi-static MD simulation samples)

The MD domain is a 10 nm × 10 nm square, and for $k_1, k_2 \in \{0, \frac{\pi}{50}, \frac{2\pi}{50}, \cdots, \frac{5\pi}{50}\}$, the prescribed external loadings are given by

$$\boldsymbol{b}(x_1, x_2) = (C^1_{k_1,k_2} \cos(k_1 x_1) \cos(k_2 x_2), 0) \quad \text{or} \quad \boldsymbol{b}(x_1, x_2) = (0, C^2_{k_1,k_2} \cos(k_1 x_1) \cos(k_2 x_2)). \tag{43}$$

The constants $C^1_{k_1,k_2}$ and $C^2_{k_1,k_2}$ are adjusted so that the bond strains are no larger than 2%. Therefore, the deformation remains in the linear range of material response. A periodic boundary condition is employed for all samples in this data set.



(a) Non-fracture training data set $\mathcal{M}_{\text{train}}^{\text{Non-Frac}}$ (b) Non-fracture validation data set $\mathcal{M}_{\text{val}}^{\text{Non-Frac}}$ (c) Fracture training data set $\mathcal{M}_{\text{train}}^{\text{Frac}}$ (d) Fracture test data set $\mathcal{M}_{\text{test}}^{\text{Frac}}$

Fig. 4 Contours of exemplar U_1 displacement in typical MD simulations at zero temperature for the four data sets: (a) non-fracture training data set $\mathcal{M}_{\text{train}}^{\text{Non-Frac}}$ for the kernel learning step; (b) non-fracture validation data set $\mathcal{M}_{\text{val}}^{\text{Non-Frac}}$ for the kernel learning step; (c) fracture training data set $\mathcal{M}_{\text{train}}^{\text{Frac}}$ for the damage criterion learning step; (d) fracture test data set $\mathcal{M}_{\text{test}}^{\text{Frac}}$ to study the efficacy and generalizability of the overall workflow (color online)

(II) Non-fracture validation data set ($\mathcal{M}_{\text{val}}^{\text{Non-Frac}}$ with 10 quasi-static MD simulation samples)

For the same MD grid and coarse-grained nodes as in the non-fracture training data set, the applied loads in the validation data set are as follows:

$$\boldsymbol{b}(x_1, x_2) = (C_k^1, C_k^2) \sum_{j=-1}^{1} (-1)^j \cos\left(\frac{\pi}{2} \min\left\{1, \frac{r_{j,k}}{R_k}\right\}\right), \tag{44}$$

where

$$r_{j,k} = \sqrt{(x_1 - (1 - p_k)Lj)^2 + (x_2 - p_k Lj)^2}. \tag{45}$$

Here, $L = 50$, and the values of the parameters C_k^1, C_k^2, p_k, and R_k are given in Table 1. In each case, loads are applied to the atoms within three disks of radius R_k with centers at the center of the grid and at the left and right boundaries (if $p_k = 0$) or the upper and lower boundaries (if $p_k = 1$). The loads in all cases are self-equilibrated and periodic. A periodic boundary condition is employed for all samples in this data set.

Table 1 Parameters used in the MD loading in the 10 validation tests

k	C_k^1	C_k^2	p_k	R_k
1	0.001	0	0	25
2	0	0.001	0	25
3	0.001	0	0	15
4	0	0.001	0	15
5	0.001	0	0	10
6	0.001	0	1	25
7	0	0.001	1	25
8	0.001	0	1	15
9	0	0.001	1	15
10	0.001	0	1	10

(III) Fracture training data set ($\mathcal{M}_{\text{train}}^{\text{Frac}}$ with 1 dynamic MD simulation sample)

The domain of the graphene sheet is set as a square: $[-50\,\text{Å}, 50\,\text{Å}] \times [-50\,\text{Å}, 50\,\text{Å}]$. The MD grid initially contains a slit (edge crack) of length 25 Å, oriented vertically extending from the lower surface. The vertical edges of the MD grid have prescribed velocities in the x_1-direction that tend to open the crack. These prescribed velocities are applied to the atoms along the vertical edges of the MD sample. To help maintain stable crack growth, the prescribed velocities decrease linearly with x_2, thus tending to limit the crack growth velocity. A schematic of the

crack pattern at the 40th time step in the MD simulation can be found in Fig. 4(c). For the purpose of validation on different grid resolutions, the density, displacement, and external loading are computed at two sets of coarse-grained nodes, which are spaced 5 Å or 3.17 Å, apart on a square lattice.

(IV) Fracture test data set ($\mathcal{M}^{\text{Frac}}_{\text{test}}$ with 1 dynamic MD simulation sample)

To demonstrate that the learnt material model applies to different loading scenarios and crack patterns, one additional test case is considered. Here, the MD region and the pre-existing slit are the same as in the fracture training data set. However, instead of hard loading along the vertical edges, a non-zero body force is applied to the atoms in the MD grid as

$$b_1 = b_0\big(\mathrm{e}^{-t/t_r}(1-\mathrm{e}^{-t/t_r})\big)\sin\Big(\frac{\pi x_1}{L}\Big)\mathrm{e}^{-\left(\frac{1}{2}+\frac{x_2}{L}\right)},$$

where $L\,(=100\,\text{Å})$ is the edge length of the sample, t_r is constant pulse duration time, and b_0 is a positive constant. The origin is at the center of the sample. This loading exerts a pulse that tends to open the crack. Unlike the fracture training case, the vertical edges are free in the fracture validation simulations. The resulting crack pattern, which includes branching, is substantially different from that occurring in the training data. A view of the crack pattern at the 40th time step in the MD simulation can be found in Fig. 4(d).

As metrics of accuracy on tests, we compare the prediction from the learnt peridynamics model with the ground-truth data from coarse-grained MD measurements. Solution contours are provided as a qualitative validation. With the purpose of providing a quantitative comparison, we also calculate the averaged (in time) mean square errors (MSEs) of the displacement field and the damage field. To provide a fair comparison between different sets, all these qualitative accuracy metrics are normalized with respect to the ground-truth data.

For the kernel learning step, we have followed a similar procedure as in Ref. [46]. The learnt influence function K is plotted in Fig. 5(a), and the optimal material parameters are obtained as $\lambda = -0.479\,6\,\text{TPa}$, $\mu = 0.797\,8\,\text{TPa}$, with Poisson's ratio $\nu = -0.429\,7$, and the horizon size $\delta = 20\,\text{Å}$. Then, for the damage criterion learning step, since the crack initiates at the 5th time step, we use the fracture training data set from the 5th time step till the 20th time step to learn the damage criterion s_0, then solve for the optimal s_0 by minimizing the loss $\widetilde{\text{Res}}(\mathcal{T}^{\text{Frac}})$ in Eq. (41). Note that when calculating the loss function, we apply the Dirichlet-type boundary conditions on a layer of particles near the boundary of our square domain, and hence only the particles in $[-50\,\text{Å}+2\delta, 50\,\text{Å}-2\delta]\times[-50\,\text{Å}+2\delta, 50\,\text{Å}-2\delta]$ are considered in Eq. (41). This setting differs from the settings in non-fracture data sets, where periodic boundary conditions are considered for all samples. This is due to the fact that it is generally non-realistic to prescribe the periodic boundary condition in the problem with a crack, since the crack itself does not satisfy the periodic condition. This fact also highlights the generalizability of the proposed approach: our homogenized surrogate model can handle data sets with different domains, loadings, and also boundary conditions. A demonstration of the loss function for different values of s_0 is provided in Fig. 5(b). The optimal damage criterion is obtained as $s_0^* = 0.11$, which is consistent with the result $s_0 = 0.145$ inferred directly from the MD data set in Ref. [43].

5.2 Extrapolation to longer time simulations

Next, we validate the learnt model, by using it in a longer term simulation on the fracture training data set, to predict the material deformation and crack propagation upto the 40th time step. Note that we have used the data upto the 20th time step for the purpose of training, and therefore this test can be seen as an investigation on the long-term extrapolation capability of our coarse-grained surrogate model. To solve for the displacement field from the LPS model, at the nth time step, we first assume that there is no broken bond, and solve for the displacement field $\hat{\boldsymbol{u}}^{n+1}$, then we update the bond-stretch for each connecting bond, and we keep solving for

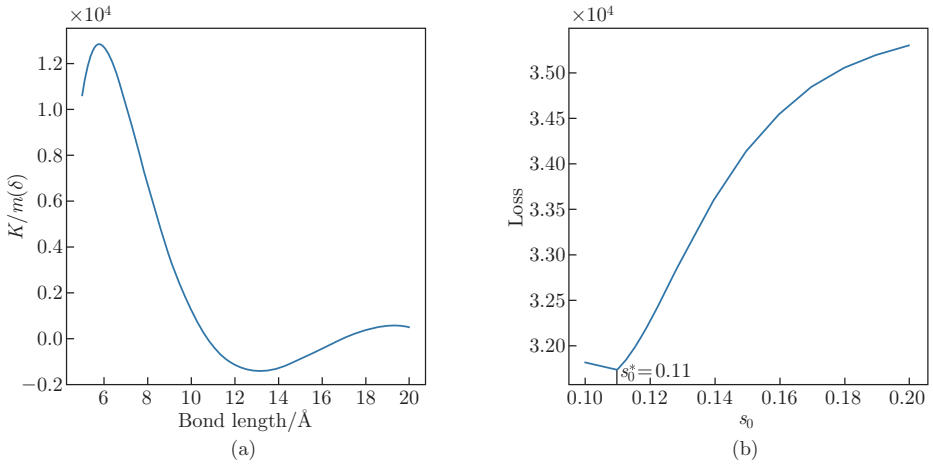


Fig. 5 Learning results on a single-layer graphene sheet, (a) the optimal influence function K for the LPS model and (b) the optimal damage criterion s_0^* obtained at 0.11 (color online)

the displacement until there is no new bond breaking. Then, we define the damage profile at each particle $\boldsymbol{x}_i$ at the time step n as

$$\phi(\boldsymbol{x}_i)^n = 1 - \frac{\sum\limits_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)} \gamma^n(\boldsymbol{x}_i, \boldsymbol{x}_j)}{\sum\limits_{\boldsymbol{x}_j \in B_\delta(\boldsymbol{x}_i)} 1}. \tag{46}$$

Figure 6 shows the comparison of displacement and damage fields at the time steps 20, 30, and 40. It is observed that the prediction not only matches the data within the training set (the step 20) but also exhibits good agreement at the time steps 30 and 40, which are not included in the training set. This result suggests that our learnt damage criterion s_0 is applicable to longer term simulations out of the training data set. For the first 40 steps, we have obtained a relative error of 27% for the prediction of the displacement field and a relative error of 9% for the damage field.

5.3 Generalization to different body forces and crack patterns

In this section, we use the learnt LPS surrogate to model the same graphene sheet subject to a different body force load as described in the fracture test data set. Different from the settings in the training data set, in this data set, the graphene sheet is subject to a nonzero body load, with its crack pattern at the 40th time step, as illustrated in Fig. 4(d). Compared with the crack pattern in the training data set (see Fig. 4(c)), the crack path in this test data set is less symmetric and bifurcates at the middle of the domain. Hence, with this example we not only investigate the extrapolation capability of the learnt model by making a longer time (the step 40) prediction, but also aim to verify its generalizability, since both the loading scenario and crack pattern from this test data set are not covered in the training data. All these factors make the validation more challenging. In Fig. 7, we show the prediction of displacement and damage fields from the learnt LPS model upto the 40th step. Visually good agreement is observed between the coarse-grained data and LPS prediction. This example has qualitatively validated that the learnt material damage model can be directly applied to the problems with different body forces. For the first 40 steps, we have obtained a relative error of 58% for the prediction of the displacement field and a relative error of 15% for the damage field.

The very small time scales that characterize MD (picoseconds) create challenges in the simulation of the mechanics of real materials, including both continuous deformation and fracture. For example, in polymers, the deformation of long molecules in response to an applied load

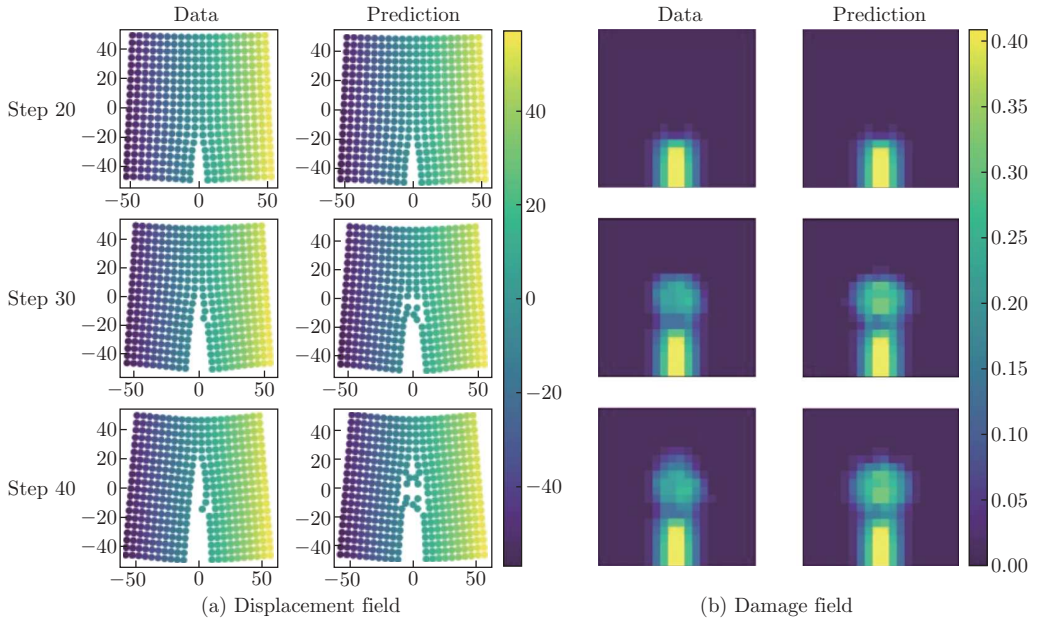


Fig. 6 Comparison of the prediction and the ground truth measurement from the MD data set at time steps 20, 30, and 40 on the fracture training data set, where the graphene sheet is subject to a zero body force. Here, we use the first 20 steps, then we use the learnt model to predict for the next 20 steps. (a) Comparison on the displacement field, where the color of the particles represents the horizontal displacement and (b) comparison on the damage field (color online)

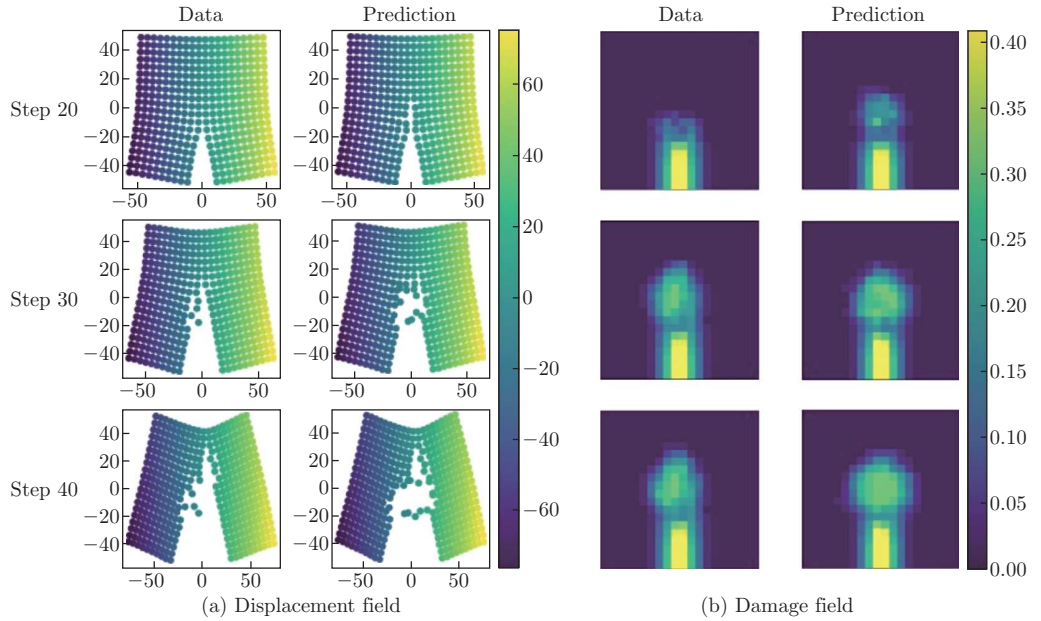
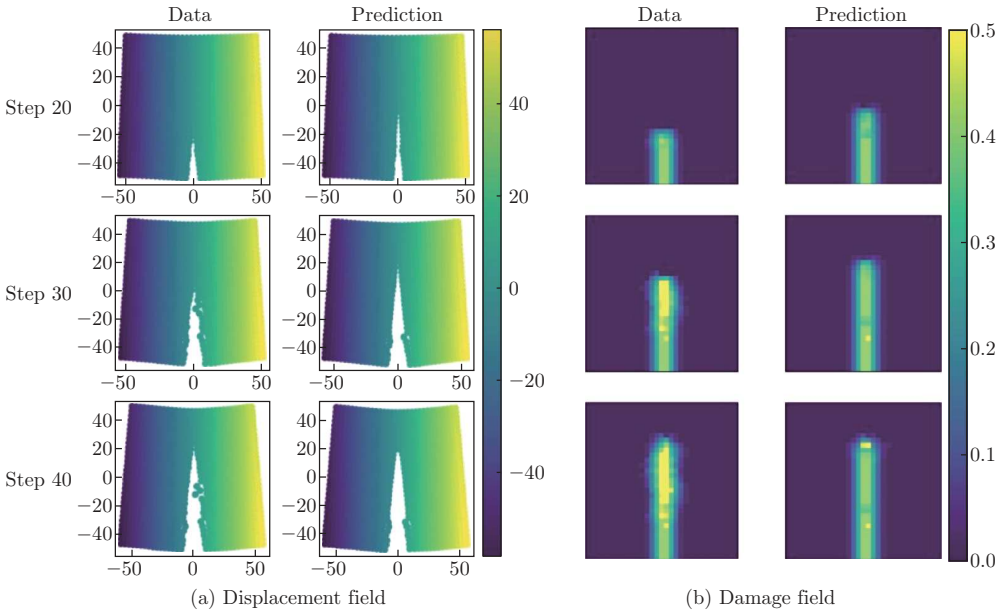


Fig. 7 Comparison of the prediction and the ground truth measurement from the MD data set at time steps 20, 30, and 40 on the fracture test data set, where the graphene sheet is subject to an unseen and nonzero body force. (a) Comparison on the displacement field, where the color of the particles represents the horizontal displacement and (b) comparison on the damage field (color online)

is partially accomplished through random processes that are strongly affected by thermal oscillations. Similar considerations apply in metals, in which dislocation motion can occur over a relatively long time scale. However, graphene is an almost ideally brittle material in which breakage of the covalent bonds in the lattice structure is governed primarily by the state of strain in the vicinity of an atom, with relatively few effects of thermal oscillations. The effects of strain rate and temperature on graphene have been studied in detail[75], confirming its relative insensitivity up to moderate temperatures. For other materials in which longer time scales apply, acceleration techniques have been developed within MD, for example, see Ref. [76]. We expect that the techniques proposed in the present paper will continue to be valid without change when such an MD acceleration is included.

5.4 Generalization to different resolutions

Last but not least, we study the resolution generalizability of our learning algorithm. Specifically, we use the same MD data as the training data, but evaluate the density, displacement, and force loading on a coarse-grained grid with a smaller grid size $h = 3.17\,\text{Å}$. Since all training data sets are with a fixed grid size $h = 5\,\text{Å}$, with this study we aim to investigate if the learnt surrogate model allows the grid size to be rescaled, providing a multiscale capability and allowing for flexible solver resolution and reductions in computational cost. As suggested in Ref. [46], we scale the horizon size δ proportionally with the grid size h to provide a fixed horizon/grid size ratio. In particular, we take $\delta = 4h = 12.68\,\text{Å}$. Then, the optimal damage criterion is also scaled correspondingly to guarantee a consistent critical release rate. As proved in Ref. [72], the damage criterion and horizon size should satisfy the relation $s_0 \propto \frac{1}{\sqrt{\delta}}$ in the LPS model. Thus, we use $s_0 := 0.11\sqrt{\frac{5}{3.17}}$ for our fine scale simulation. In Fig. 8, we show the displacement and damage field prediction results upto the 40th step, demonstrating qualitative agreement between the coarse-grained MD data and our numerical prediction. For the displacement field,



(a) Displacement field (b) Damage field

Fig. 8 Comparison of the prediction and the ground truth measurement from the MD data set at time steps 20, 30, and 40 on the fracture training data set with fine grids, where the graphene sheet is subject to a zero body force. Here, we use coarser grids data in the first 20 steps for training, then we use the learnt model to predict for the next 20 steps on a finer resolution. (a) Comparison on the displacement field, where the color of the particles represents the horizontal displacement and (b) comparison on the damage field (color online)

we have obtained a relative error of 19% in average for the first 40 steps, which is even smaller than the prediction error without resolution alternation on the same data set (27% as shown in Subsection 5.2). For the damage field, one can see that the crack pattern predicted by our surrogate model grows faster than the crack from the MD data set. Therefore, a larger prediction error, 30% average for the damage field in the first 40 steps, is obtained. This example suggests that the surrogate model can provide the qualitatively consistent displacement prediction on different resolutions. On the other hand, the prediction on the damage field is sub-optimal, possibly due to the fact that the material crack originates from microscale phenomena, and hence is more sensitive to the prediction scales. To improve the prediction accuracy on the damage field across different resolutions, practitioners might consider performing the damage criterion learning step on the new resolution, to provide a correction for the damage criterion.

6 Conclusions

In this paper, we demonstrate a data-driven workflow to extract a coarse-grained surrogate model from MD data with fracture. Firstly, to handle the discontinuities induced by material fracture in the MD displacement measurements, a smoothness indicator function is introduced, to automatically choose the locally smoothest stencil from the neighborhood of each coarse-grained grid. As such, the coarse-grained measurements are built based on this adaptive stencil, to automatically handle the discontinuities in the MD displacement data set without overly smoothing the crack pattern. It is shown that this novel adaptive procedure significantly improves the capability of capturing the location of crack interfaces. Then, based on the coarse-grained data set, we propose to extract a peridynamics surrogate, which is a continuum mechanics model that allows a natural treatment of discontinuities by replacing spatial derivatives of stress tensors with integrals of force density functions. By learning the kernel function of the integral and the damage criterion with a two-step optimization approach, we obtain an LPS model which provides good agreement with nanoscale test data while being capable to provide further material deformation and fracture predictions under unseen domain settings, loading scenarios, and even different grid resolutions. These features greatly reduce the cost of the calculation in comparison with MD, especially when used together with different discretization resolutions.

Although the present work focuses on relatively small deformations and a linear peridynamics model, the results suggest that this method may impact a broader range of materials and applications. As another natural follow-up work, one may further combine the nonlocal model with the approximation power of neural networks, to obtain a nonlinear peridynamics model in the form of integral neural operators[77–80].

Conflict of interest All authors declare no conflict of interest.

Acknowledgements H. Q. YOU and Y. YU would like to acknowledge the projects support by the National Science Foundation (No. DMS-1753031) and the Air Force Office of Scientific Research (No. FA9550-22-1-0197). Portions of this research were conducted on Lehigh University's Research Computing Infrastructure partially supported by the National Science Foundation (No. 2019035).

S. SILLING and M. D'ELIA would like to acknowledge the support of the Sandia National Laboratories (SNL) Laboratory-directed Research and Development Program and the U. S. Department of Energy (DOE), Office of Advanced Scientific Computing Research (ASCR) under the Collaboratory

on Mathematics and Physics-Informed Learning Machines for Multiscale and Multiphysics Problems (PhILMs) project. This article has been authored by an employee of National Technology and Engineering Solutions of Sandia, LLC under contract No. DE-NA0003525 with the U. S. DOE. The employee owns all right, title, and interest in and to the article and is solely responsible for its contents. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article or allow others to do so, for United States Government purposes. The DOE will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan https://www.energy.gov/downloads/doe-public-access-plan.

References

[1] ZOHDI, T. and STEIGMANN, D. The toughening effect of microscopic filament misalignment on macroscopic ballistic fabric response. *International Journal of Fracture*, **118**(4), 71–76 (2002)

[2] WRIGGERS, P., ZAVARISE, G., and ZOHDI, T. A computational study of interfacial debonding damage in fibrous composite materials. *Computational Materials Science*, **12**(1), 39–56 (1998)

[3] PRUDENCIO, E. E., BAUMAN, P. T., WILLIAMS, S., FAGHIHI, D., RAVI-CHANDAR, K., and ODEN, J. T. A dynamic data driven application system for real-time monitoring of stochastic damage. *Procedia Computer Science*, **18**, 2056–2065 (2013)

[4] SU, Z., YE, L., and LU, Y. Guided Lamb waves for identification of damage in composite structures: a review. *Journal of Sound and Vibration*, **295**, 753–780 (2006)

[5] 2014 technical strategic plan. *Technical Report*, The Air Force Office of Scientific Research, Arlington, Virginia (2014)

[6] TALREJA, R. and VARNA, J. *Modeling Damage, Fatigue and Failure of Composite Materials*, Elsevier, Cambridge (2015)

[7] SORIĆ, J., WRIGGERS, P., and ALLIX, O. *Multiscale Modeling of Heterogeneous Structures*, Springer Cham, Switzerland (2018)

[8] PIJAUDIER-CABOT, G. and DUFOUR, F. *Damage Mechanics of Cementitious Materials and Structures*, John Wiley & Sons, U. S. A. (2013)

[9] MOURLAS, C., MARKOU, G., and PAPADRAKAKIS, M. Accurate and computationally efficient nonlinear static and dynamic analysis of reinforced concrete structures considering damage factors. *Engineering Structures*, **178**, 258–285 (2019)

[10] MARKOU, G., GARCIA, R., MOURLAS, C., GUADAGNINI, M., PILAKOUTAS, K., and PAPADRAKAKIS, M. A new damage factor for seismic assessment of deficient bare and FRP-retrofitted RC structures. *Engineering Structures*, **248**, 113152 (2021)

[11] ZOHDI, T. I. Homogenization methods and multiscale modeling. *Encyclopedia of Computational Mechanics Second Edition*, John Wiley & Sons, U. S. A., 1–24 (2017)

[12] BENSOUSSAN, A., LIONS, J. L., and PAPANICOLAOU, G. *Asymptotic Analysis for Periodic Structures*, American Mathematical Society, U. S. A. (2011)

[13] WEINAN, E. and ENGQUIST, B. Multiscale modeling and computation. *Notices of the AMS*, **50**(9), 1062–1070 (2003)

[14] EFENDIEV, Y., GALVIS, J., and HOU, T. Y. Generalized multiscale finite element methods (GMsFEM). *Journal of Computational Physics*, **251**, 116–135 (2013)

[15] JUNGHANS, C., PRAPROTNIK, M., and KREMER, K. Transport properties controlled by a thermostat: an extended dissipative particle dynamics thermostat. *Soft Matter*, **4**(1), 156–161 (2008)

[16] KUBO, R. The fluctuation-dissipation theorem. *Reports on Progress in Physics*, **29**(1), 255–284 (1966)

[17] SANTOSA, F. and SYMES, W. W. A dispersive effective medium for wave propagation in periodic composites. *SIAM Journal on Applied Mathematics*, **51**(4), 984–1005 (1991)

[18] DOBSON, M., LUSKIN, M., and ORTNER, C. Sharp stability estimates for the force-based quasi-continuum approximation of homogeneous tensile deformation. *Multiscale Modeling & Simulation*, **8**(3), 782–802 (2010)

[19] HUGHES, T. J., WELLS, G. N., and WRAY, A. A. Energy transfers and spectral eddy viscosity in large-eddy simulations of homogeneous isotropic turbulence: comparison of dynamic Smagorinsky and multiscale models over a range of discretizations. *Physics of Fluids*, **16**(11), 4044–4052 (2004)

[20] MOËS, N., ODEN, J. T., VEMAGANTI, K., and REMACLE, J. F. Simplified methods and a posteriori error estimation for the homogenization of representative volume elements (RVE). *Computer Methods in Applied Mechanics and Engineering*, **176**, 265–278 (1999)

[21] ORTIZ, M. A method of homogenization of elastic media. *International Journal of Engineering Science*, **25**(7), 923–934 (1987)

[22] DU, Q., ENGQUIST, B., and TIAN, X. Multiscale modeling, homogenization and nonlocal effects: mathematical and computational issues. *Contemporary Mathematics*, American Mathematical Society, U. S. A. (2020)

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

[24] BERAN, M. and MCCOY, J. Mean field variations in a statistical sample of heterogeneous linearly elastic solids. *International Journal of Solids and Structures*, **6**(8), 1035–1054 (1970)

[25] CHEREDNICHENKO, K., SMYSHLYAEV, V. P., and ZHIKOV, V. Non-local homogenised limits for composite media with highly anisotropic periodic fibres. *Proceedings of the Royal Society of Edinburgh Section A*: *Mathematics*, **136**(1), 87–114 (2006)

[26] KARAL, F. C., JR and KELLER, J. B. Elastic, electromagnetic, and other waves in a random medium. *Journal of Mathematical Physics*, **5**(4), 537–547 (1964)

[27] RAHALI, Y., GIORGIO, I., GANGHOFFER, J., and DELL'ISOLA, F. Homogenization à la Piola produces second gradient continuum models for linear pantographic lattices. *International Journal of Engineering Science*, **97**, 148–172 (2015)

[28] SMYSHLYAEV, V. P. and CHEREDNICHENKO, K. D. On rigorous derivation of strain gradient effects in the overall behaviour of periodic heterogeneous media. *Journal of the Mechanics and Physics of Solids*, **48**, 1325–1357 (2000)

[29] WILLIS, J. R. The nonlocal influence of density variations in a composite. *International Journal of Solids and Structures*, **21**(7), 805–817 (1985)

[30] ERINGEN, A. C. and EDELEN, D. G. B. On nonlocal elasticity. *International Journal of Engineering Science*, **10**(3), 233–248 (1972)

[31] BOBARU, F., FOSTER, J. T., GEUBELLE, P. H., and SILLING, S. A. *Handbook of Peridynamic Modeling*, CRC Press, Boca Raton (2016)

[32] BAŽANT, Z. P. and JIRÁSEK, M. Nonlocal integral formulations of plasticity and damage: survey of progress. *Journal of Engineering Mechanics*, **128**(11), 1119–1149 (2002)

[33] DU, Q., GUNZBURGER, M., LEHOUCQ, R. B., and ZHOU, K. A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. *Mathematical Models and Methods in Applied Sciences*, **23**(3), 493–540 (2013)

[34] SELESON, P., PARKS, M. L., GUNZBURGER, M., and LEHOUCQ, R. Peridynamics as an upscaling of molecular dynamics. *Multiscale Modeling & Simulation*, **8**(1), 204–227 (2009)

[35] PARKS, M. L., LEHOUCQ, R. B., PLIMPTON, S. J., and SILLING, S. A. Implementing peridynamics within a molecular dynamics code. *Computer Physics Communications*, **179**(11), 777–783 (2008)

[36] ZIMMERMANN, M. *A Continuum Theory with Long-Range Forces for Solids*, Ph. D. dissertation, Massachusetts Institute of Technology (2005)

[37] EMMRICH, E. and WECKNER, O. Analysis and numerical approximation of an integro-differential equation modeling non-local effects in linear elasticity. *Mathematics and Mechanics of Solids*, **12**(4), 363–384 (2007)

[38] DU, Q. and ZHOU, K. Mathematical analysis for the peridynamic nonlocal continuum theory. *ESAIM*: *Mathematical Modelling and Numerical Analysis*, **45**(2), 217–234 (2011)

[39] DU, Q., TAO, Y., and TIAN, X. A peridynamic model of fracture mechanics with bond-breaking. *Journal of Elasticity*, **132**(2), 197–218 (2018)

[40] PRAKASH, N. and SEIDEL, G. D. Electromechanical peridynamics modeling of piezoresistive response of carbon nanotube nanocomposites. *Computational Materials Science*, **113**, 154–170 (2016)

[41] PRAKASH, N. and SEIDEL, G. D. Computational electromechanical peridynamics modeling of strain and damage sensing in nanocomposite bonded explosive materials (NCBX). *Engineering Fracture Mechanics*, **177**, 180–202 (2017)

[42] SILLING, S. A. Propagation of a stress pulse in a heterogeneous elastic bar. *Journal of Peridynamics and Nonlocal Modeling*, **3**(3), 255–275 (2021)

[43] SILLING, S. A., D'ELIA, M., YU, Y., YOU, H., and FERMEN-COKER, M. Peridynamic model for single-layer graphene obtained from coarse-grained bond forces. *Journal of Peridynamics and Nonlocal Modeling*, 1–22 (2022) https://doi.org/10.1007/s42102-021-00075-w

[44] YOU, H., YU, Y., SILLING, S., and D'ELIA, M. Data-driven learning of nonlocal models: from high-fidelity simulations to constitutive laws. *arXiv Preprint*, arXiv: 2012.04157 (2020) https://doi.org/10.48550/arXiv.2012.04157

[45] YOU, H., YU, Y., TRASK, N., GULIAN, M., and D'ELIA, M. Data-driven learning of nonlocal physics from high-fidelity synthetic data. *Computer Methods in Applied Mechanics and Engineering*, **374**, 113553 (2021)

[46] YOU, H., YU, Y., SILLING, S., and D'ELIA, M. A data-driven peridynamic continuum model for upscaling molecular dynamics. *Computer Methods in Applied Mechanics and Engineering*, **389**, 114400 (2022)

[47] XU, X., D'ELIA, M., and FOSTER, J. T. A machine-learning framework for peridynamic material models with physical constraints. *Computer Methods in Applied Mechanics and Engineering*, **386**, 114062 (2021)

[48] XU, X., D'ELIA, M., GLUSA, C., and FOSTER, J. T. Machine-learning of nonlocal kernels for anomalous subsurface transport from breakthrough curves. *arXiv Preprint*, arXiv: 2201.11146 (2022) https://doi.org/10.48550/arXiv.2201.11146

[49] ZHANG, L., YOU, H., and YU, Y. MetaNOR: a meta-learnt nonlocal operator regression approach for metamaterial modeling. *MRS Communications*, **12**, 662–677 (2022)

[50] LU, F., AN, Q., and YU, Y. Nonparametric learning of kernels in nonlocal operators. *arXiv Preprint*, arXiv: 2205.11006 (2022) https://doi.org/10.48550/arXiv.2205.11006

[51] MURDOCH, A. and BEDEAUX, D. Continuum equations of balance via weighted averages of microscopic quantities. *Proceedings of the Royal Society of London Series A*: *Mathematical and Physical Sciences*, **445**, 157–179 (1994)

[52] LIU, X. D., OSHER, S., and CHAN, T. Weighted essentially non-oscillatory schemes. *Journal of Computational Physics*, **115**(1), 200–212 (1994)

[53] JIANG, G. S. and SHU, C. W. Efficient implementation of weighted ENO schemes. *Journal of Computational Physics*, **126**(1), 202–228 (1996)

[54] SILLING, S. A., EPTON, M., WECKNER, O., XU, J., and ASKARI, E. Peridynamic states and constitutive modeling. *Journal of Elasticity*, **88**(2), 151–184 (2007)

[55] SILLING, S. A. and LEHOUCQ, R. B. Peridynamic theory of solid mechanics. *Advances in Applied Mechanics*, **44**, 73–168 (2010)

[56] LEHOUCQ, R. B. and SILLING, S. A. Force flux and the peridynamic stress tensor. *Journal of the Mechanics and Physics of Solids*, **56**(4), 1566–1577 (2008)

[57] SILLING, S., LITTLEWOOD, D., and SELESON, P. Variable horizon in a peridynamic medium. *Journal of Mechanics of Materials and Structures*, **10**(5), 591–612 (2015)

[58] SILLING, S. A., EPTON, M., WECKNER, O., XU, J., and ASKARI, E. Peridynamic states and constitutive modeling. *Journal of Elasticity*, **88**(2), 151–184 (2007)

[59] EMMRICH, E. and WECKNER, O. On the well-posedness of the linear peridynamic model and its convergence towards the Navier equation of linear elasticity. *Communications in Mathematical Sciences*, **5**(4), 851–864 (2007)

[60] MENGESHA, T. and DU, Q. Nonlocal constrained value problems for a linear peridynamic Navier equation. *Journal of Elasticity*, **116**(1), 27–51 (2014)

[61] MENGESHA, T. and DU, Q. The bond-based peridynamic system with Dirichlet-type volume constraint. *Proceedings of the Royal Society of Edinburgh Section A*: *Mathematics*, **144**(1), 161–186 (2014)

[62] QIN, H., SUN, Y., LIU, J. Z., LI, M., and LIU, Y. Negative Poisson's ratio in rippled graphene. *Nanoscale*, **9**(12), 4135–4142 (2017)

[63] JIANG, J. W., CHANG, T., GUO, X., and PARK, H. S. Intrinsic negative Poisson's ratio for single-layer graphene. *Nano Letters*, **16**(8), 5286–5290 (2016)

[64] TRASK, N., YOU, H., YU, Y., and PARKS, M. L. An asymptotically compatible meshfree quadrature rule for nonlocal problems with applications to peridynamics. *Computer Methods in Applied Mechanics and Engineering*, **343**, 151–165 (2019)

[65] YU, Y., YOU, H., and TRASK, N. An asymptotically compatible treatment of traction loading in linearly elastic peridynamic fracture. *Computer Methods in Applied Mechanics and Engineering*, **377**, 113691 (2021)

[66] FAN, Y., TIAN, X., YANG, X., LI, X., WEBSTER, C., and YU, Y. An asymptotically compatible probabilistic collocation method for randomly heterogeneous nonlocal problems. *Journal of Computational Physics*, **465**, 111376 (2022)

[67] YOU, H., LU, X., TASK, N., and YU, Y. An asymptotically compatible approach for Neumann-type boundary condition on nonlocal problems. *ESAIM*: *Mathematical Modelling and Numerical Analysis*, **54**(4), 1373–1413 (2020)

[68] YOU, H., YU, Y., and KAMENSKY, D. An asymptotically compatible formulation for local-to-nonlocal coupling problems without overlapping regions. *Computer Methods in Applied Mechanics and Engineering*, **366**, 113038 (2020)

[69] FOSS, M., RADU, P., and YU, Y. Convergence analysis and numerical studies for linearly elastic peridynamics with Dirichlet-type boundary conditions. *Journal of Peridynamics and Nonlocal Modeling*, 1–36 (2022) https://doi.org/10.1007/s42102-021-00074-x

[70] FAN, Y., YOU, H. Q., and YU, Y. OBMeshfree: an optimization-based meshfree solver for nonlocal diffusion and peridynamics models. *arXiv Preprint*, arXiv: 2211.14953 (2022) https://doi.org/10.48550/arXiv.2211.14953

[71] FAN, Y., YOU, H., TIAN, X., YANG, X., LI, X., PRAKASH, N., and YU, Y. A meshfree peridynamic model for brittle fracture in randomly heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering*, **399**, 115340 (2022)

[72] ZHANG, H. and QIAO, P. A state-based peridynamic model for quantitative fracture analysis. *International Journal of Fracture*, **211**, 217–235 (2018)

[73] YOU, H., YU, Y., TRASK, N., GULIAN, M., and D'ELIA, M. Data-driven learning of robust nonlocal physics from high-fidelity synthetic data. *Computer Methods in Applied Mechanics and Engineering*, **374**, 113553 (2021)

[74] TERSOFF, J. Empirical interatomic potential for carbon, with applications to amorphous carbon. *Physical Review Letters*, **61**(25), 2879–2882 (1988)

[75] ZHAO, H. and ALURU, N. R. Temperature and strain-rate dependent fracture strength of graphene. *Journal of Applied Physics*, **108**(6), 064321 (2010)

[76] VOTER, A. F. Hyperdynamics: accelerated molecular dynamics of infrequent events. *Physical Review Letters*, **78**(20), 3908–3911 (1997)

[77] KOVACHKI, N., LI, Z., LIU, B., AZIZZADENESHELI, K., BHATTACHARYA, K., STUART, A., and ANANDKUMAR, A. Neural operator: learning maps between function spaces with applications to PDEs. *Journal of Machine Learning Research*, **24**(89), 1–97 (2023)

[78] YOU, H., ZHANG, Q., ROSS, C. J., LEE, C. H., HSU, M. C., and YU, Y. A physics-guided neural operator learning approach to model biological tissues from digital image correlation measurements. *Journal of Biomechanical Engineering*, **144**, 121012 (2022)

[79] YOU, H., YU, Y., D'ELIA, M., GAO, T., and SILLING, S. Nonlocal kernel network (NKN): a stable and resolution-independent deep neural network. *Journal of Computational Physics*, **469**, 111536 (2022)

[80] YOU, H., ZHANG, Q., ROSS, C. J., LEE, C. H., and YU, Y. Learning deep implicit Fourier neural operators (IFNOs) with applications to heterogeneous material modeling. *Computer Methods in Applied Mechanics and Engineering*, **398**, 115296 (2022)