

<https://doi.org/10.1038/s41524-024-01394-4>

# Data-driven modeling of dislocation mobility from atomistics using physics-informed machine learning

Check for updates

Yifeng Tian<sup>1</sup>✉, Soumendu Bagchi<sup>2,3</sup>, Liam Myhill<sup>4</sup>, Giacomo Po<sup>5</sup>, Enrique Martinez<sup>4</sup>, Yen Ting Lin<sup>1</sup>, Nithin Mathew<sup>2</sup> & Danny Perez<sup>2</sup>

Dislocation mobility, which dictates the response of dislocations to an applied stress, is a fundamental property of crystalline materials that governs the evolution of plastic deformation. Traditional approaches for deriving mobility laws rely on phenomenological models of the underlying physics, whose free parameters are in turn fitted to a small number of intuition-driven atomic scale simulations under varying conditions of temperature and stress. This tedious and time-consuming approach becomes particularly cumbersome for materials with complex dependencies on stress, temperature, and local environment, such as body-centered cubic crystals (BCC) metals and alloys. In this paper, we present a novel, uncertainty quantification-driven active learning paradigm for learning dislocation mobility laws from automated high-throughput large-scale molecular dynamics simulations, using Graph Neural Networks (GNN) with a physics-informed architecture. We demonstrate that this Physics-informed Graph Neural Network (PI-GNN) framework captures the underlying physics more accurately compared to existing phenomenological mobility laws in BCC metals.

Almost a century ago<sup>1–3</sup>, dislocations were identified as carriers of plastic deformation in crystalline materials. A key construct of the theory is the response of dislocations to an applied stress state, the *dislocation mobility law*, which describes the relationship between the force experienced by the dislocation ( $\mathbf{f}$ ) and the resultant velocity ( $\mathbf{v}$ ). In crystals with low lattice resistance and planar core structure, e.g., Face-Centered Cubic (FCC) metals, the mobility law is usually simplified to a viscous damping relationship  $\mathbf{v} = \mathbf{M}(T) \cdot \mathbf{f}$ , where  $\mathbf{M}$  is the matrix of mobility coefficients and  $T$  is the temperature<sup>4</sup>, under the assumption of over-damped dynamics. In metals,  $\mathbf{M}$  encodes dissipation mechanisms due to both electrons and phonons, with dominant phonon contribution except at very low temperatures. Theoretical calculations of phonon scattering predict a linear dependence of the mobility coefficients on temperature<sup>5</sup> with a temperature-independent leading term proposed recently<sup>6,7</sup>. While this picture holds for edge dislocations in Body-Centered Cubic (BCC) metals, the complicated core structure of the BCC screw dislocations results in additional complexity. Screw dislocations in BCC are known to have high lattice resistance. This in turn results in a strong temperature dependence of the flow stress as the mobility of screw dislocations in BCC are controlled by kink-pair nucleation and migration enthalpy. In addition, screw dislocation

motion is strongly influenced by multiple stress components, in addition to the resolved shear stress in the direction of the Burgers vector, resulting in non-Schmid effects<sup>8–10</sup>. These complexities are typically accounted for in the mobility law by introducing functional dependencies on the local stress tensor  $\sigma$ , line orientation  $\xi$ , temperature and possibly other local internal variables, that is<sup>11</sup>

$$\mathbf{v} = \mathbf{g}_\alpha(\sigma, \xi, T, \dots). \quad (1)$$

The function  $\mathbf{g}_\alpha$  is known as the (glide) *mobility law* for slip system  $\alpha$ <sup>12–14</sup>. The functional dependencies of the mobility law are in turn informed by phenomenological approaches based on kink-pair nucleation and kink diffusion. A review of the phenomenological approach for deriving dislocation mobility laws can be found in Ref. 15. State-of-the-art methods to develop such models for dislocation mobility (e.g., in BCC materials)<sup>16–18</sup> have generally focused on screw and edge type dislocations and smooth interpolation between these for other line orientations. Researchers have explored calibrating an orientation-dependent phenomenological model using several intuition-based data points, e.g. Ref. 19 for BCC crystals<sup>20</sup>, for FCC Aluminum, and for superalloys<sup>21</sup>. However, this approach presents

<sup>1</sup>Information Sciences Group, Computer, Computational and Statistical Sciences Division (CCS-3), Los Alamos National Laboratory, Los Alamos, 87545 NM, USA.

<sup>2</sup>Physics and Chemistry of Materials, Theoretical Division (T-1), Los Alamos National Laboratory, Los Alamos, 87545 NM, USA. <sup>3</sup>Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, USA. <sup>4</sup>Department of Materials Science and Engineering, Clemson University, Clemson, 29623 SC, USA.

<sup>5</sup>Department of Mechanical and Aerospace Engineering, University of Miami, Miami, 33146 FL, USA. ✉e-mail: [yifengtian91@gmail.com](mailto:yifengtian91@gmail.com)

multiple pitfalls. For example, it has been shown previously that “singular” line orientations exist in BCC metals<sup>19,22</sup> which result in cusps in the mobility. Dislocations in these singular orientations, and those vicinal to them, move by motion and propagation of kink pairs. Therefore, their mobility cannot be accurately represented using a smooth interpolation between screw and edge dislocations. Moreover, the mobility becomes a complex function of local chemical environments in the case of conventional and high entropy alloys (HEAs). The atomically-rough energy landscape for dislocation motion, which is believed to be an important characteristic of HEAs<sup>23,24</sup>, results in Arrhenius-type mobility laws which are highly non-linear in stress and temperature<sup>15</sup>. It has been demonstrated in BCC-based HEAs that the mobility of mixed dislocations<sup>25</sup> and multiplicity of glide planes<sup>26</sup> can drive the overall plastic behavior, in contrast to BCC metals. The traditional phenomenological approach informed by a handful of intuition-driven data-points to derive the mobility law can quickly become error-prone and cumbersome in such scenarios.

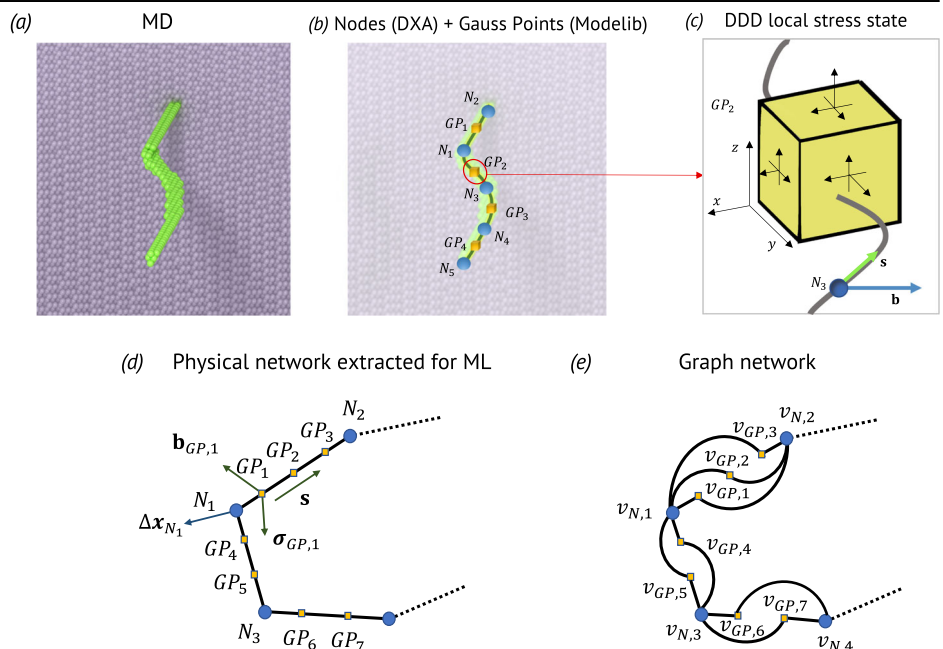
These challenges can in principle be addressed by using data-driven approaches, whereby active learning techniques are used to efficiently explore the parameter space for learning the dislocation mobility law with minimal human intervention. Such approaches are especially relevant with the advent of exascale computing platforms where hierarchical multi-scale methods, deployed at scale with automated model refinement, can potentially be used to drive materials design for targeted functionality. Moreover, the improved accuracy of Machine-Learning (ML) interatomic potentials<sup>27</sup> promises affordable near-quantum accurate predictions, which could enable upscaling of atomic-level defect dynamics to predictive mesoscale models such as Discrete Dislocation Dynamics (DDD)<sup>28–33</sup>. To date, the exploration of these ideas is in a relatively nascent stage. Previous studies have explored the possibility of using graph-theoretical techniques for learning specific aspects of dislocation glide. Moraes et al.<sup>34</sup> developed a graph-based surrogate model for glide of edge dislocations in BCC Fe, trained to molecular dynamics (MD) data. In this representation, sub-domains within the MD model are represented using nodes on a ring graph. The dislocations are then represented as random walkers which jump between neighboring nodes following a Poisson process. More recently, Bertin and Zhou<sup>35</sup> developed a Graph Neural Network (GNN)<sup>36</sup> approach to model DDD, with the aim of replacing its time-integration procedure. They demonstrated that this DDD-GNN approach can learn ground-truth (GT) DDD data of a dislocation gliding in a field of obstacles. While these studies

show the potential for accelerating materials modeling through machine learning, bridging mesoscale dynamics (i.e. DDD) with information from higher fidelity (i.e. atomistic) simulations involving more general defect characteristics (i.e. mixed orientations of dislocations) remains to be explored.

Our objective in this study is to develop a generalizable framework for data-driven modeling of dislocation dynamics mobility law. Our proposed framework integrates high-throughput MD simulation datasets, physics-inspired machine learning model structures, and uncertainty quantification, accelerating the development of mesoscale models of dislocation mobility. A schematic diagram with different components of our proposed modeling workflow is shown in Fig. 1. To demonstrate our approach, GT data was generated from classical MD simulations of dislocation motion using an Embedded Atom Method (EAM) type interatomic potential (Fig. 1a); the workflow would remain formally identical, albeit more expensive, if ML interatomic potentials were used instead. The MD simulation data consists of dislocation dipoles and shear loops moving under different values of resolved shear stress (RSS) at various temperatures. Note that, while Non-Schmid effects are important in BCC materials, our aim here is to demonstrate the ML workflow supported by high-throughput large-scale atomistic trajectory generation protocols for a relatively simple data set with only RSS. Coarse-grained representations of dislocations were obtained from this atomistic representation using the Dislocation Extraction Algorithm (DXA) as implemented in the Open Visualisation Tool (OVITO)<sup>37</sup> software. The DXA representation, which consists of nodes with line segments between them, was converted into a numerical representation with nodes, quadrature points or Gauss points (GPs), and line segments as shown in Fig. 1b. Relevant input features of these dislocation configurations, which will serve as input for an ML procedure, such as the Cauchy stress tensor ( $\sigma$ , shown schematically in Fig. 1c), segment tangent vector ( $s$ ), and the Burger’s vector ( $b$ ) were computed using a recent version of the open-source Mechanics of Defects Evolution Library (MoDELlib) supporting periodic boundary conditions<sup>38</sup>. We remark that the computation of stress tensor by MoDELlib is based on linear elasticity theory and the resulting stress can deviate from the true stress experienced by the dislocation. Nevertheless, we hypothesize that the stress tensor computed by MoDELlib contains sufficient information for us to make accurate predictions, and the error due to the linear assumption, if any, will be corrected by our proposed GNN-based ML procedure below.

**Fig. 1 | Schematic of ground truth data generation.**

**a** Atomistic representation of a dislocation segment (green atoms represent the defect core while the purple denotes pristine atoms). **b** Idealized nodal representation of the dislocation segment obtained from Dislocation Extraction Algorithm (DXA). Blue spheres represent dislocation nodes (denoted as  $N_1, N_2, \dots$ ) and yellow cubes represent the Cauchy stress tensor, one of the features, defined at Gauss Points (denoted as  $GP_1, GP_2, \dots$ ) using the linear elasticity kernel in MoDELlib Discrete Dislocation Dynamics (DDD) code. **c** Zoom-in of the local stress state ( $\sigma$ ) at a given Gauss point, with the definition of segment tangent vector ( $s$ ) and the Burger’s vector ( $b$ ). **d** General representation of the dislocation configuration with multiple Gauss points between nodes. **e** Graph representation of the dislocation configuration.



We also note that the physical features of the dislocation lines, such as the Burger’s vector, segment tangent vector, and local stress, are defined on both the GPs and nodes, while the displacements ( $\Delta\mathbf{x}$ ) are defined on the nodes alone, with several GPs being present between any given pair of nodes in general. Dislocation network, thus represented as a set of nodes and Gauss points (GPs) in DDD can be naturally encoded as a graph. In Fig. 1d and e, we present the visualization of the dislocation network, showcasing both its physical representation extracted from DDD for ML and its graph representation. The physical network serves as an abstract depiction of the dislocation network, which consists of nodes ( $N_i$ ) and GPs ( $GP_i$ ), connected by the dislocation segments. Figure 1e illustrates the corresponding heterogeneous graph, a type of graph containing vertices with different features, for learning purposes. Within this heterogeneous graph, we introduce two types of vertices,  $v_N$  representing the nodes and  $v_{GP}$  representing the GPs in the physical network. In the graph representation, instead of connecting  $v_{GP}$ ’s sequentially into a segment between nodes  $v_N$ , we directly link GPs to their nearest nodes. This heterogeneous graph enables the direct exchange of information between GPs and their neighboring nodes, facilitating efficient information transfer.

To achieve the goal of the proposed work, we present a general physics-informed data-driven approach for modeling dislocation mobility law. Our approach incorporates a physics-inspired mathematical structure based on the theoretical descriptions of dislocation mobility into the proposed GNN model. Features are designed to respect physical constraints, e.g. rotational and translational invariance. The proposed physics-informed GNN mobility law (PI-GNN) is trained using a coarse-grained MD dataset, with a loss function defined in “Modeling mobility law using Physics-informed Graph Neural Network (PI-GNN)” in “Methods”. To quantify the predictive uncertainty and facilitate accelerated learning, we employ an uncertainty quantification-driven active learning framework<sup>39</sup> that is capable of providing on-the-fly uncertainty measures and automatically query new MD simulations to improve the model’s predictability and uncertainty. By studying the mobility of {110} dislocations in BCC Fe, we demonstrate the effectiveness of this PI-GNN framework to accurately capture the relevant physics of dislocations systematically.

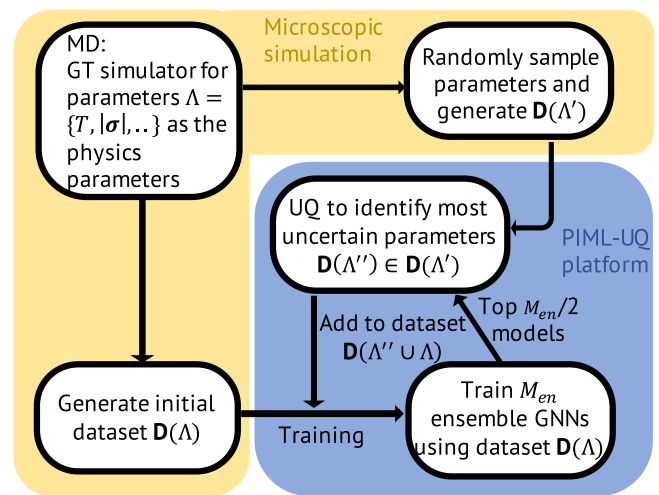
## Results

### Learning the mobility function using Graph Neural Networks in the Physics-informed Machine Learning (PIML) framework

A GNN is a type of neural network designed to act on graph-structured data, where nodes represent entities (e.g., the nodes and GPs on the dislocation) and edges represent relationships (e.g., the spatial proximity of the nodes) between them. GNNs learn powerful representations of graph data by propagating information between neighboring nodes. Unlike convolutional neural networks on grids, GNNs respect rotational and translational invariance, making them well-suited for learning on irregular graph domains. As such, GNNs serve as an ideal neural architecture for the coarse-grained representation of dislocation structures generated by DXA. Furthermore, the design of the PI-GNN architecture was informed by established physics in linear elasticity theory pertaining to dislocation dynamics in continuum media<sup>15</sup>, i.e.,

$$\frac{d\mathbf{x}}{dt} = \frac{\mathbf{f}(\boldsymbol{\sigma}, \mathbf{b}, \mathbf{s})}{B(T, \mathbf{b}, \mathbf{s})} \exp\left(-\frac{\epsilon(\boldsymbol{\sigma}, \mathbf{b}, \mathbf{s}, T)}{k_B T}\right), \quad (2)$$

where  $\mathbf{f}$ ,  $B$ , and  $\epsilon$  are separate networks modeling the contribution of the Peach-Koehler force, drag, and thermally activated processes in dislocation dynamics. These networks, referred to as F-net, B-net, and E-net, are jointly trained on time-series data of the dislocation dynamics, generated by MD and processed through the DXA-MoDELlib pipeline. Through message passing<sup>40</sup>, the PI-GNN not only uses the local information but also information on nearby nodes to learn higher-order corrections to linear theories. More details of the PI-GNN model we adopted can be found in “Methods” and in the Supplementary Note 1.



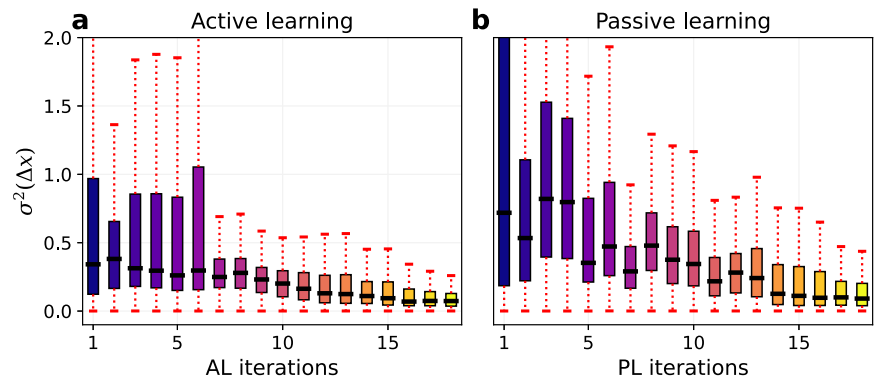
**Fig. 2 | Flow chart of UQ-driven active learning framework.** The UQ-AL framework is composed of two computational platforms, the high-throughput microscopic MD simulations, and PIML-UQ platform. The UQ-AL workflow begins by generating and learning from an initial dataset, and then progressively queries new datasets where the trained models exhibit the most uncertainty. In each iteration, we independently train an ensemble of identical GNN models with randomized initialization of weights and biases. These models are then evaluated on newly generated datasets with randomly sampled simulation parameters, and the uncertainties of the models’ single-step predictions are quantified. We integrate the top 50% of new configurations that induce the highest predictive uncertainty into the existing training dataset. The PI-GNN ensemble is retrained on this augmented dataset, and the process repeats.

### Uncertainty quantification-driven active learning (UQ-AL) framework to facilitate efficient learning

We propose a workflow, outlined in Fig. 2, for training the PI-GNN while aiming to minimize its predictive uncertainty. Rather than generating one large dataset all at once, the workflow begins by learning from an initial dataset, and then progressively queries new datasets where the trained models exhibit the most uncertainty. In each iteration, we independently train  $M_{en}$  identical GNN models with randomized initialization of weights and biases. These models are then evaluated along the trajectories of a set (8) of newly generated datasets with randomly sampled simulation parameters  $\Lambda = \{T, |\boldsymbol{\sigma}|, \dots\}$ , and the uncertainties of the models’ single-step predictions are quantified. This can be considered as an ensemble learning<sup>39</sup> method. We integrate the top 50% of new configurations that induce the highest predictive uncertainty into the existing training dataset. The PI-GNN ensemble is retrained on this augmented dataset, and the process repeats. We note that this method is an approximation of more rigorous but computationally expensive UQ-AL frameworks (e.g., Bayesian Optimization<sup>41</sup>) for two reasons. First, the variability of the  $M_{en}$  model predictions is only a lower bound to the true uncertainty<sup>39</sup>. Second, we select the most uncertain candidates from randomly sampled simulation parameters, rather than directly querying the single most uncertain point across the entire simulation parameter space. More details of the proposed UQ-AL framework can be found in “Methods”.

We illustrate the advantages of the proposed UQ-AL framework in learning the mobility law using the dipole dataset generated from MD simulations as described in “Automated on-the-fly generation of dislocation mobility data in MD” in “Methods”. In this numerical experiment, the PI-GNN models are continuously retrained as more data are acquired and integrated into the training set, and we compare the performance of the model when the new datasets are either actively selected by the UQ-AL or passively queried by randomization. We fix the total number of ensemble models  $M_{en}$  at 20. The initial dataset  $D(\Lambda)$  are generated on the set of simulation parameters: temperature  $T \in \{300 \text{ K}, 500 \text{ K}\}$ , RSS  $|\boldsymbol{\sigma}| \in \{0.25 \text{ GPa}, 0.5 \text{ GPa}\}$ , and the line orientation (i.e., the angle between the

**Fig. 3 | Comparison of the proposed UQ-AL framework with a passive learning method.** The uncertainty measure of predictions for nodal displacement as a function of the (a) active (AL) and (b) passive learning (PL) iterations.



tangent vector  $\mathbf{s}$  and the Burger's vector  $\mathbf{b}$  angle  $(\mathbf{s}, \mathbf{b}) \in \{45^\circ, 135^\circ\}$ . In each iteration, we maintain a candidate pool of 8 new trajectories by sampling the simulation parameters,  $T \sim \text{Unif}(100 \text{ K}, 1000 \text{ K})$ ,  $|\boldsymbol{\sigma}| \sim \text{Unif}(0.1 \text{ GPa}, 1.0 \text{ GPa})$ , and angle  $(\mathbf{s}, \mathbf{b}) \in (0^\circ, 180^\circ)$ , to create new trajectories of dislocation dynamics. With UQ-AL, the trained PI-GNN ensemble is evaluated along the trajectories to quantify their averaged single-step predictive uncertainty; we select and integrate the top 50% (4) trajectories with the highest predictive uncertainties (i.e. the PI-GNN ensemble models disagree with the most) into the existing training dataset. With passive learning, we randomly select 4 new trajectories from the pool and integrate them into the training dataset. For both methods, the selected trajectories are removed from the candidate pool, which is then replenished by 4 newly sampled parameter settings to generate the trajectories for the next iteration.

Figure 3 illustrates how the model performance improves as more data are acquired into the training set, for both active and passive selection of the datasets. Evidently, with the proposed UQ-AL framework, the prediction uncertainty is consistently reduced through the iterations. In comparison, the predictive uncertainty with passive selection is higher and fluctuates more between iterations, even though it eventually decreases in the large-data limit. Suppose we set the uncertainty threshold to be 25% of the mean nodal displacement,  $\sigma(\Delta \mathbf{x}) = 25\% \mu(\Delta \mathbf{x})$ , the active and passive acquisition of the new training dataset requires 8 and 15 iterations, or equivalently 36 and 64 queries (initial 8 fixed, then 8 randomly sampled ones to populate the candidate pool in the first iteration, and 4 randomly sampled ones in each of the subsequent iterations) of new trajectories, respectively. As such, UQ-AL saves more than 40% of computational resources for generating training dataset in this experiment. Because the proposed UQ-AL can efficiently query computationally expensive high-fidelity datasets to learn mesoscale models and thus accelerate the development of data-driven mesoscopic models for DDD, we deploy this framework in the following analysis as part of the data acquisition pipeline.

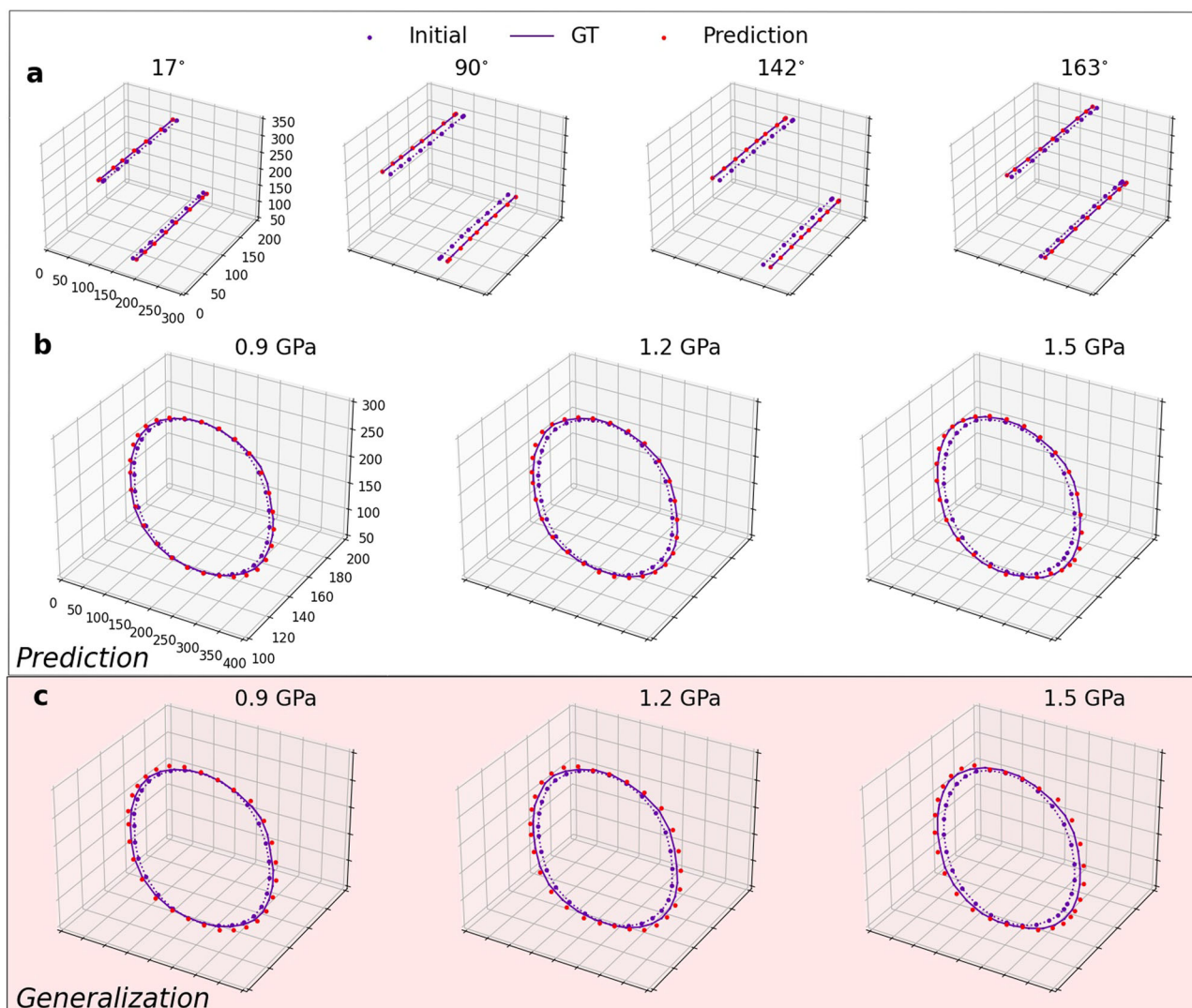
### Predictability and generalizability of the PI-GNN mobility law

In the following analysis, the proposed UQ-AL is used to train the PI-GNN model until the predicted uncertainty is reduced to 25% of the mean nodal displacement. The 50% top-performing models in the ensemble are selected for the following validation analysis, in which we compare the single-step ( $\Delta t \approx 1 \text{ ps}$ ) prediction of the trained models to a validation dataset. The model used for prediction was trained using dipole dataset with  $\text{RSS} \in \{0.1 \text{ GPa}, 0.25 \text{ GPa}, 0.5 \text{ GPa}, 0.75 \text{ GPa}, 1.0 \text{ GPa}\}$ ,  $T \in \{100 \text{ K}, 300 \text{ K}, 500 \text{ K}, 700 \text{ K}, 1000 \text{ K}\}$ , and angle  $(\mathbf{s}, \mathbf{b}) \in (0^\circ, 180^\circ)$ . A separate validation dataset was used to demonstrate the prediction capability of the PI-GNN model. See “Methods” for details on the train/test/validation dataset. We designed three different experiments: (a) the PI-GNNs ensemble is trained on a dataset comprised of nominally straight dislocation dipoles and compared against the dipole validation dataset, (b) the PI-GNNs ensemble is trained on a dataset comprised of expanding shear loops and tested on a loop validation dataset, and (c) the PI-GNNs ensemble is trained on the

dipole dataset and tested on the loop dataset. Figure 4 shows selected examples of the single-step prediction; more results can be found in Supplementary Note 2.

We observe that the PI-GNN is capable of learning an effective mobility law for various dislocation dipole configurations. Experiment (a) (Fig. 4a) shows that the PI-GNN model trained on various conditions of temperature, RSS, and line orientation of a dipole configuration, can learn an effective mobility law to predict the evolution of the dislocation with an out-of-sample dipole configuration. In contrast to experiment (a), where we explicitly sampled the line orientation of the dislocation for training, the PI-GNN learns the dependence of the line orientation directly from the angular distribution on the evolving loops in experiment (b), and accurately predicts the evolution of out-of-sample loop configurations. Experiment (c) shows that the PI-GNN is able to learn the mobility law from the dipole dataset and generalize to the out-of-distribution prediction task for loop configurations. Such generalizability indicates that the PI-GNN is learning the physics of the dislocation dynamics, instead of memorizing the training data.

Quantitative error analyses of the learned PI-GNN mobility law are reported in Fig. 5, where the panels' layout is analogous to Fig. 4. The prediction error is quantified by both the percentage and absolute errors, the computation of which is based on the proposed loss function (See “Modeling mobility law using Physics-informed Graph Neural Network (PI-GNN)” in “Methods” for more information). For the dislocation dipole, the median prediction error is in general very small for different orientations, temperatures and stress levels (500 K and 0.9 GPa is shown in the figure). For screw-like dislocation where the angle  $(\mathbf{b}, \mathbf{s})$  is close to  $0^\circ$  and  $180^\circ$ , the percentage error is comparatively larger. This is mainly caused by an imbalanced training dataset, due to the timescale separation between the dynamics of screw dislocations and mixed/edge dislocations. At the timescales of the dataset (several picoseconds) used in this study, the screw dislocation dynamics are dominated by thermal fluctuations. Therefore, the signal-to-noise ratio of the nodal displacement, which is the prediction goal of the PI-GNN mobility law, is much smaller than that of the edge. The small signal-to-noise ratio can only be seen for configurations near the pure screw configuration (i.e., angle  $(\mathbf{b}, \mathbf{s}) = 0^\circ$  and  $180^\circ$ ), while the majority of the data points do not have this signature. Consequently, the trained PI-GNN model predicts the edge and mixed configurations better than the pure screw configuration. Future work to rectify the training data imbalance will be discussed in “Discussion”. When we compare the absolute error in Fig. 5b, the absolute errors for different angles are at approximately the same range. For dislocation loops, the single-step prediction percentage error is also very small. The errors in generalization are plotted in Fig. 5d. As explained above, we train the model in this case using dislocation dipoles and make predictions on shear loops. Compared to Fig. 5c, even though the percentage error is larger, it still falls in an acceptable range, with the median at  $\sim 20\%$ . This shows that with the PIML structure and features, the PI-GNN model can still capture roughly 80% of the total motion even without any additional training on the complex dataset. These numerical experiments demonstrate the potential generalizability of the proposed UQ-AL PI-GNN framework for modeling complex dislocation networks using simple datasets.



**Fig. 4 | The predicted single-step evolution for various dislocation configurations and learning method.** **a** PI-GNN mobility law is trained on coarse-grained dipole dataset with different RSS values, temperatures, and angles between Burger's vector and tangent vector. The predictions are also made on dipole configurations with

different line orientations at 500 K and 0.9 GPa. **b** PI-GNN mobility law trained on dislocation loop expansion data and prediction on dislocation loops at 500 K and 0.9, 1.2, and 1.5 GPa. **c** Testing the generalizability of the PI-GNN mobility law by training on dipole dataset and predicting on loop expansion.

### Integrating PI-GNN mobility law with DDD for iterative multi-step predictions

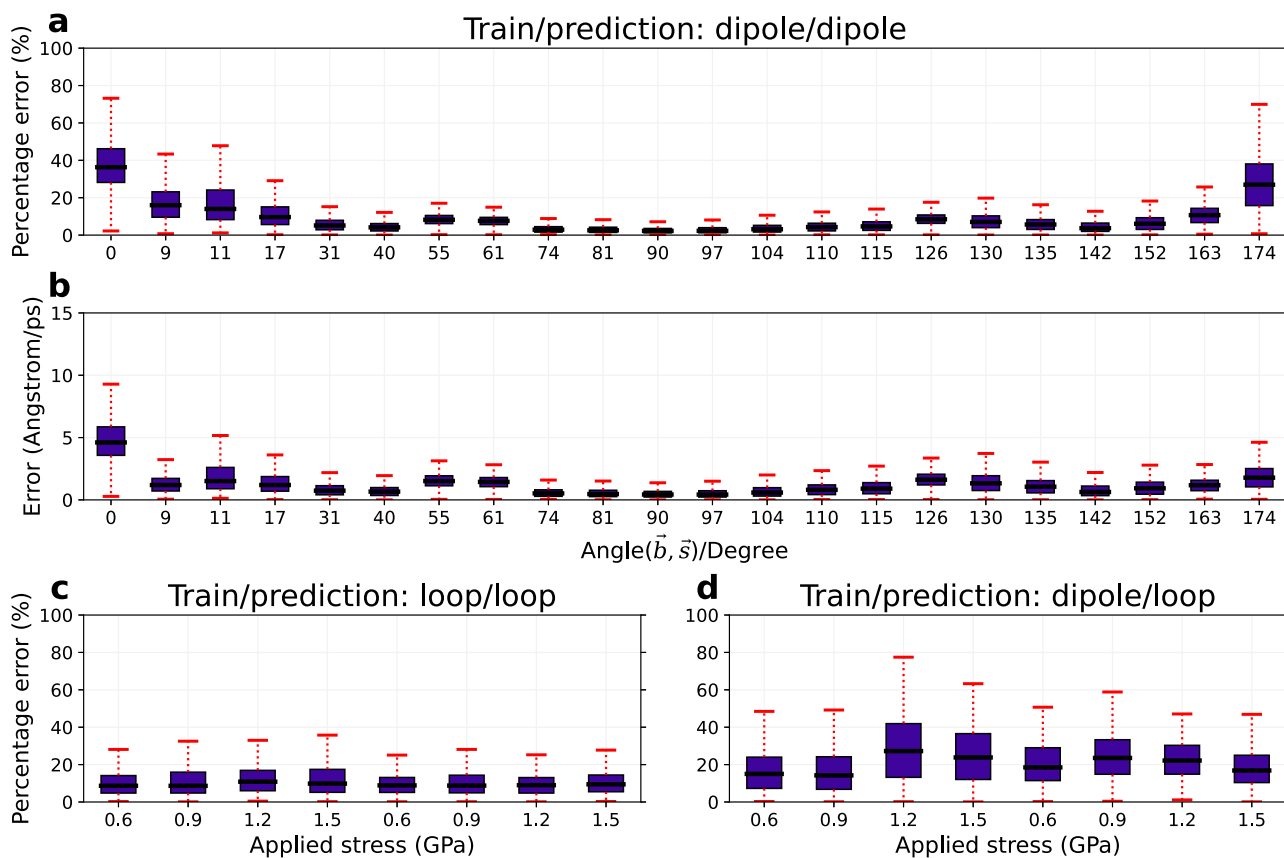
In Fig. 6, the predicted iterative multi-step evolution of both dipole and loop configurations are visualized at different consecutive time steps. We show results at 0.9 GPa and 500 K, and angle  $(\mathbf{b}, \mathbf{s}) \approx 80^\circ$  for the dislocation dipole. For each row in Fig. 6, the panels from left to right show the forward evolution in time, where configurations from preceding time steps are also visualized as translucent markers. As observed in Fig. 6, the PI-GNN mobility law enables us to achieve qualitatively accurate long-time predictions. The coarse-grained model prediction will eventually deviate from the GT, but within the visualized prediction horizon, the relative error remains  $<10\%$ . Quantitative analysis of the errors in multi-step prediction is provided in Supplementary Fig. 13.

### Interpretability of the PI-GNN mobility law

The phenomenological model, represented by Eq. (2), involves three distinct physical aspects of dislocation dynamics, namely the force experienced by the dislocation given a stress state, dissipative mechanisms resulting in drag, and thermal activation, each of these making a unique contribution. We model each aspect using a separate GNN as shown in Eq. (2) and train them altogether with a single loss function. Even though quantitatively isolating

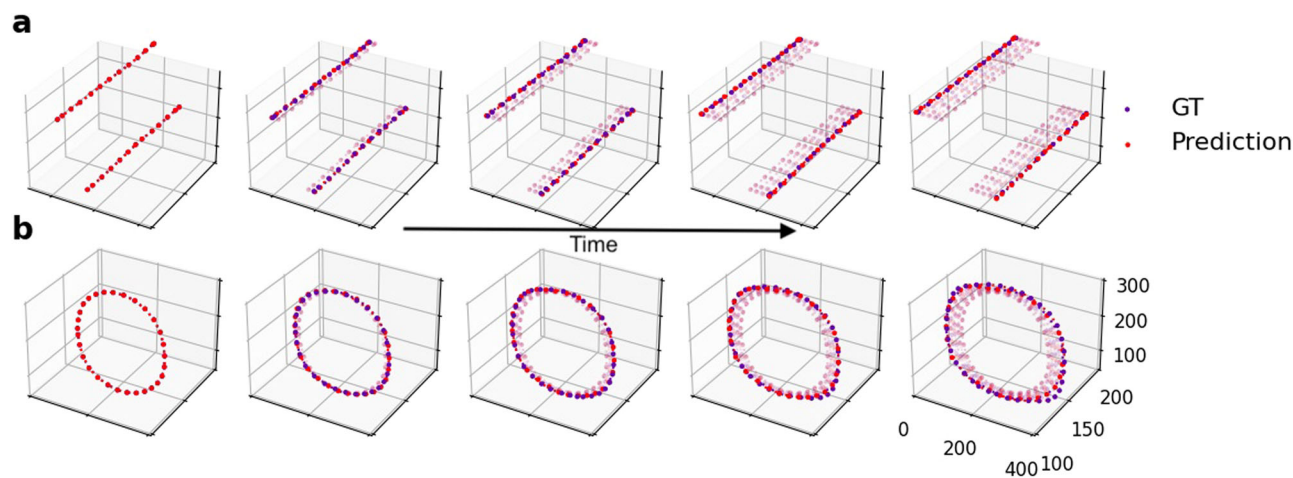
each process from the trained PI-GNN mobility law is challenging, the proposed framework still offers qualitative insights into their sensitivity and dependence on the input features. For instance, scaling both F-net and B-net by a constant factor does not alter the output. To resolve the identifiability issue arising from the linear constant factor, we compute the ratio of F-net over B-net, as well as other components of the PI-GNN mobility law, against various input features.

In Fig. 7a, we plot the output of F-net/B-net at different angles and applied stress levels. Despite the identifiability challenge, combining the outputs of both networks allows us to study their collective contributions to the prediction. With the increase of applied stress level, the predicted contributions to nodal displacement become larger as expected. For different angles representing edge, screw or mixed dislocations, the magnitude of the output is larger around  $90^\circ$  (i.e., edge dislocation), and small at  $0^\circ$  and  $180^\circ$  (i.e., screw dislocation), accurately capturing the well-known disparities in the mobility of screw and edge dislocations in BCC metals. An interesting observation is that the output is not symmetric with respect to  $90^\circ$  and is skewed toward larger angles. In Fig. 7b, the dependence of the combined F-net/B-net on temperatures is plotted. We observe that the increase in phonon drag with increasing temperature is also captured by the GNNs. It is important to note that the temperature-dependence of phonon drag is



**Fig. 5 | The predicted single-step error for various dislocation configurations and learning method.** PI-GNN mobility law is trained on coarse-grained dipole dataset with different applied stress levels, temperatures, and angles between Burger’s vector and tangent vector. The predictions are also made on dipole configurations with different orientations at 500 K and 0.9 GPa. Both the (a) percentage error and (b)

absolute unnormalized error are shown. **c** PI-GNN mobility law trained on dislocation loop expansion data and prediction on dislocation loops at 500 K and 700 K and various applied stress levels. **d** Testing the generalizability of the PI-GNN mobility law by training on dipole dataset and predicting on loop expansion.

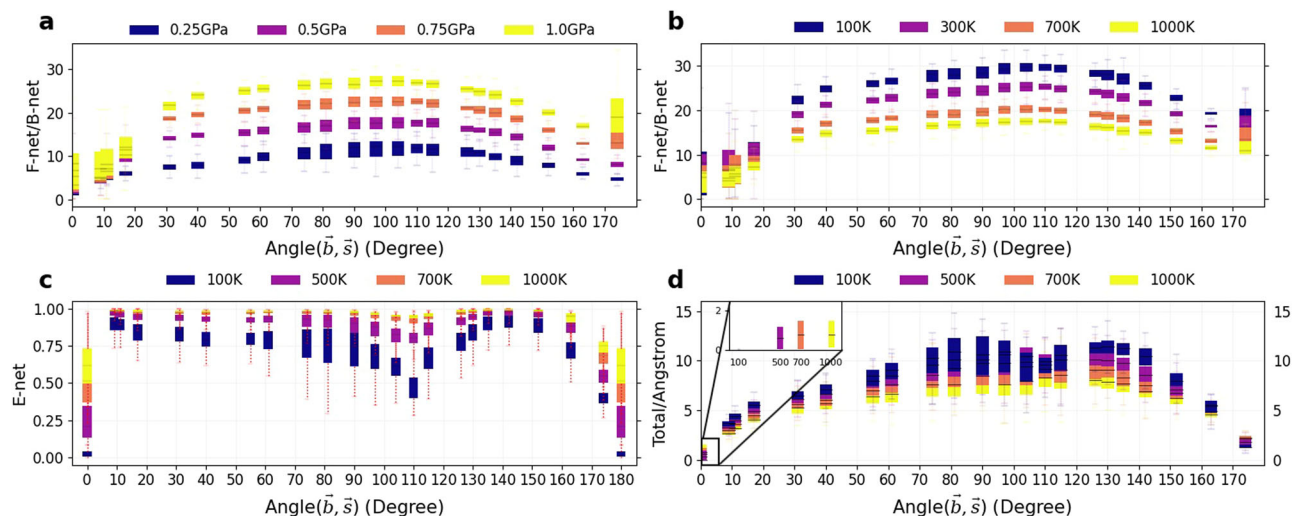


**Fig. 6 | The predicted multi-step evolution for various dislocation configurations and learning method.** For a given snapshot, the GT is represented using blue-filled circles, the PI-GNN prediction using red-filled circles with transparency applied to configurations from all preceding time steps. **a** A dislocation dipole (angle ( $\vec{b}, \vec{s}$ )  $\approx$

$80^\circ$ ) at 0.9 GPa and 500 K predicted using PI-GNN mobility law trained on dislocation dipole data. **b** An expanding dislocation loop at 500 K and 1.2 GPa, predicted using PI-GNN mobility law trained on dislocation loop expansion data.

known to be stronger for edge-like dislocations<sup>17</sup> and the GNN learns this behavior; as shown in Fig. 7b that the temperature-dependence is stronger close to  $90^\circ$  or edge dislocation as it is apparent from the GT MD data (see Supplementary Fig. 11 for dislocation velocities as a function of stress and temperature for screw, edge, and mixed configurations).

Figure 7c illustrates the dependence of the contribution from E-net output on the dislocation orientations. We note that for screw dislocations, with angles ( $\vec{b}, \vec{s}$ )  $\approx 0^\circ, 180^\circ$ , the PI-GNN mobility law qualitatively captures the comparatively higher activation energy, leading to lower output from the E-net. Conversely, at the mixed configurations, the E-net output becomes



**Fig. 7 | Physics-informed interpretation of the three components of the PI-GNN mobility law. a** Box plots of the  $\frac{F}{B}$  for different orientations, applied stress levels at 500 K. **b** Box plots of the  $\frac{F}{B}$  for different orientations, and different temperatures at

0.5 GPa. **c** Box plots of the E for different orientations, and different temperatures at 0.5 GPa. **d** Box plots of the total predicted displacement for different orientations, and different temperatures at 0.25 GPa.

larger. In addition, we observe a cusp at  $110^\circ$ , predicting a reduced contribution from the E-net for that singular orientation. The combined output from all three networks in Fig. 7d shows that the PI-GNN mobility law is capable of predicting expected dependencies on temperatures and angles. For instance, the increased phonon drag from high temperatures causes the dislocation lines to move slower at some orientations. In the case of screw dislocations with a low applied stress level, the nodal displacement is dominated by thermal activation, which has a reversed order compared to other angles (see the inset in Fig. 7d). This is also apparent for the singular orientation at  $110^\circ$ , where thermal activation dominates at low temperatures. Such cusps in dislocation mobility is known to be present in BCC metals<sup>19,22</sup> due to singular dislocation orientations which move by kink nucleation and propagation. The predicted trends are aligned very well with what is observed in MD simulations (see Supplementary Fig. 10), demonstrating the accuracy of the PI-GNN mobility law and the ability of GNNs to learn complex dependence of dislocation mobility on different parameters.

## Discussion

The central objective of this work is to provide a general, physics-informed data-driven framework for modeling the mobility law in discrete dislocation dynamics simulations autonomously from higher fidelity molecular dynamics data. When constructing the modeling framework, we focused on the following aspects: generalizability, interpretability, and uncertainty quantification. To achieve generalizability for the model, we adopt the flexible graph-based data structure to represent the dislocation network, which can be easily extended to dislocation configurations beyond those studied in this paper. We demonstrate that this framework is capable of learning many complex aspects of dislocation motion in BCC metals, such as highly anisotropic dislocation mobility between screw and edge configurations, cusps in dislocation mobility at singular line orientations, and temperature dependence of phonon drag. The interpretability of the proposed framework results from the physics-inspired structure of the PI-GNN mobility law and physics-informed feature engineering. We illustrate the interpretability of the proposed framework by analyzing learned components of the PI-GNN, which showed qualitative agreement with various physical aspects of dislocation mobility. Uncertainty quantification is built into the proposed framework to inform data selection from MD simulations for active learning, which is shown to significantly accelerate the development of PI-GNN mobility laws. We acknowledge that with the limited dataset used for training, this specific model may not perform well on large DDD simulations with complex configurations. However, the proposed uncertainty-driven AL, along with PI-GNN architecture will help accelerate

the model development and refinement for complex configurations. We also expect the proposed framework to be beneficial when the dimensionality of the simulation parameters is even larger, e.g. for learning Non-Schmid effects in BCC.

Further work focusing on improving the PI-GNN mobility law using the proposed UQ-AL framework include (1) mixed modeling of the screw and mixed/edge dislocations, trained by MD data generated at different timescales for capturing both the fast-evolving mixed/edge dislocations and slow-evolving screw dislocations, (2) time scale-aware feature engineering for mobility law input, for enabling variable time-step integration, and (3) stochastic modeling of screw-like dislocations. In addition to the mobility law, an important component of DDD simulations is the description of discrete events resulting in topological re-arrangements of the dislocation network. These consist of cross-slip of screw dislocations, junction formation and annihilation, and dislocation nucleation. Our PI-GNN mobility law is capable of predicting a 3-D displacement and thus, it is well set up to predict cross-slip events. Extension of the PI-GNN mobility law to describe dislocation junctions and nucleation is possible using advanced methodologies like dynamic graphs<sup>42,43</sup>, for example. We also note that the extension of our framework to more complex systems like alloys is possible by using additional features for local environment dependence. These advancements will be considered in future studies.

We are aware of a pre-print by Bertin et al.<sup>44</sup>, independently advocating a GNN-based method for learning dislocation mobilities. While the basic concept is similar, namely representing dislocations as graphs, there are multiple distinctions between the two approaches as follows: (1) Training protocols: we tested multiple, UQ-driven, training protocols, with relatively simple MD simulations of straight dislocations and expanding shear loops in broad parameter settings. In contrast, Bertin et al. used large-scale MD simulations of high strain-rate deformation, using a narrow set of loading conditions; (2) Our PI-GNN approach allows the development of a physically interpretable non-linear mobility law, while such an interpretation is not straightforward in the approach by Bertin et al.; (3) Features for our PI-GNN model are composed of local stress tensor and line geometry (defined by Burgers vector and line tangent) as opposed to nodal forces in Bertin et al.; (4) Loss function in our approach is defined using a distance metric between nodal positions, as opposed to a Nye tensor based approach in Bertin et al. While there are pros and cons to the two approaches, a detailed evaluation of their relative merits will be the subject of a future study.

**Table 1 | List of a few sample atomistic supercell orientations out of a total of 25 cases**

Line angles w.r.t. Burger's vector ( $\vec{b}$ )	x-axis orientation	y-axis orientation (line direction)
0°	[12 $\bar{1}$ ]	[111]
40.32°	[114 $\bar{1}\bar{1}$ ]	[2112]
81.07°	[107 $\bar{1}\bar{0}$ ]	[7207]
110.41°	[2858 $\bar{2}\bar{8}$ ]	[292829]
163.14°	[8348]	[17817]

Glide-plane normal [101] is along z-axis for all the dipole configurations.

## Methods

In this section, we describe the generation of GT data from MD simulations, calculation of relevant features using the linear elasticity kernel in MoDELlib Discrete Dislocation Dynamics (DDD) code, physics-informed graph neural networks (PI-GNN) framework for modeling mobility law of discrete dislocation dynamics, including the GNN architecture, physics-informed feature engineering, generalized loss function, and uncertainty-driven active learning (AL).

### Automated *on-the-fly* generation of dislocation mobility data in MD

Generation of dislocation motion data from molecular dynamics simulations has been widely addressed in past<sup>16–18,45</sup>. It is a common practice to use free surface boundary conditions in the direction normal to the glide plane while building atomistic models with specific dislocation configurations moving under externally applied stress and temperature. Evaluation of the effect of free surfaces, on the stress experienced by the dislocation, requires solution of a boundary value problem coupled with the DDD framework. To circumvent this, we used periodic boundary conditions and devised an automated model generation protocol where straight dislocation dipoles on a slip system [11 $\bar{1}$ ](101) are created for an arbitrary line orientation (defined by angle  $\theta$ ) with the Burger's vector ( $\pm\vec{b}$ ). The simulation box is aligned as  $b \sin \theta$  (x),  $b \cos \theta$  (y) and [101] (z) with approximate lengths of 30 nm, 18 nm and 40 nm in these directions, adjusted based on the crystallography of Fe BCC unit cell. The dislocation line is oriented parallel to y-axis. Extra half-plane of atoms are removed along x with a thickness of  $b \sin \theta$ . Note, for screw dipoles ( $\theta = 0^\circ$ ) this results in no deletion, while for an edge ( $\theta = 90^\circ$ ) a complete half-plane with a thickness of  $b$  is removed. We use the periodic image correction superimposed on the Volterra displacement field with 4 periodic images on both sides along Cartesian directions x and z of the box, i.e.,  $63 [(4 \times 2)^2 - 1]$  images in total following Cai et al.<sup>46</sup>. We report a few sample box-orientations in Table 1, chosen out of a total 25 cases ranging from 0° to 180°.

For simulating the motion of a shear loop, an initial loop of radius 15 nm is created on a [11 $\bar{1}$ ](101) system. The simulation box is then affinely deformed according to the linear elastic solution corresponding to the applied stress to prevent the loop from shrinking due to self interactions. To generate high-throughput dislocation trajectories, we subject both the loop and dipole configurations to a resolved shear stress range of  $\sigma \in \{0.1, 0.25, 0.5, 0.75, 1.0\}$  GPa and a temperature range of  $T \in \{100, 300, 500, 700, 1000\}$  K. All MD simulations were performed using the EAM potential for iron (Fe) developed by Mendeleev et al.<sup>47</sup>.

To efficiently perform active learning, *on-the-fly* processing of million atom MD trajectories is crucial. Our dipole models have 3–4 million atoms and MD simulations were performed using the kokkos-implementation of LAMMPS<sup>48</sup>. We apply a predefined global stress tensor and temperatures of the simulation box by running simulations in the NPT ensemble for 40 ps with a timestep of 1 fs. On NVIDIA A100 GPUs, 1 million EAM atoms/GPU can be simulated at a rate of up to 10 ns/day, excluding any input/output (I/O) overhead. To avoid reducing the simulation rate due to I/O as

well as the storage of large volumes of MD data, we implement an *in memory* and *on-the-fly* analysis of dislocation trajectories via the Dislocation Extraction Algorithm (DXA) from Ovito-API<sup>49</sup>. Simulations were analyzed every 0.1 ps, with nodal spacings of 1.5 nm. This is schematically shown in Fig. 1a, b.

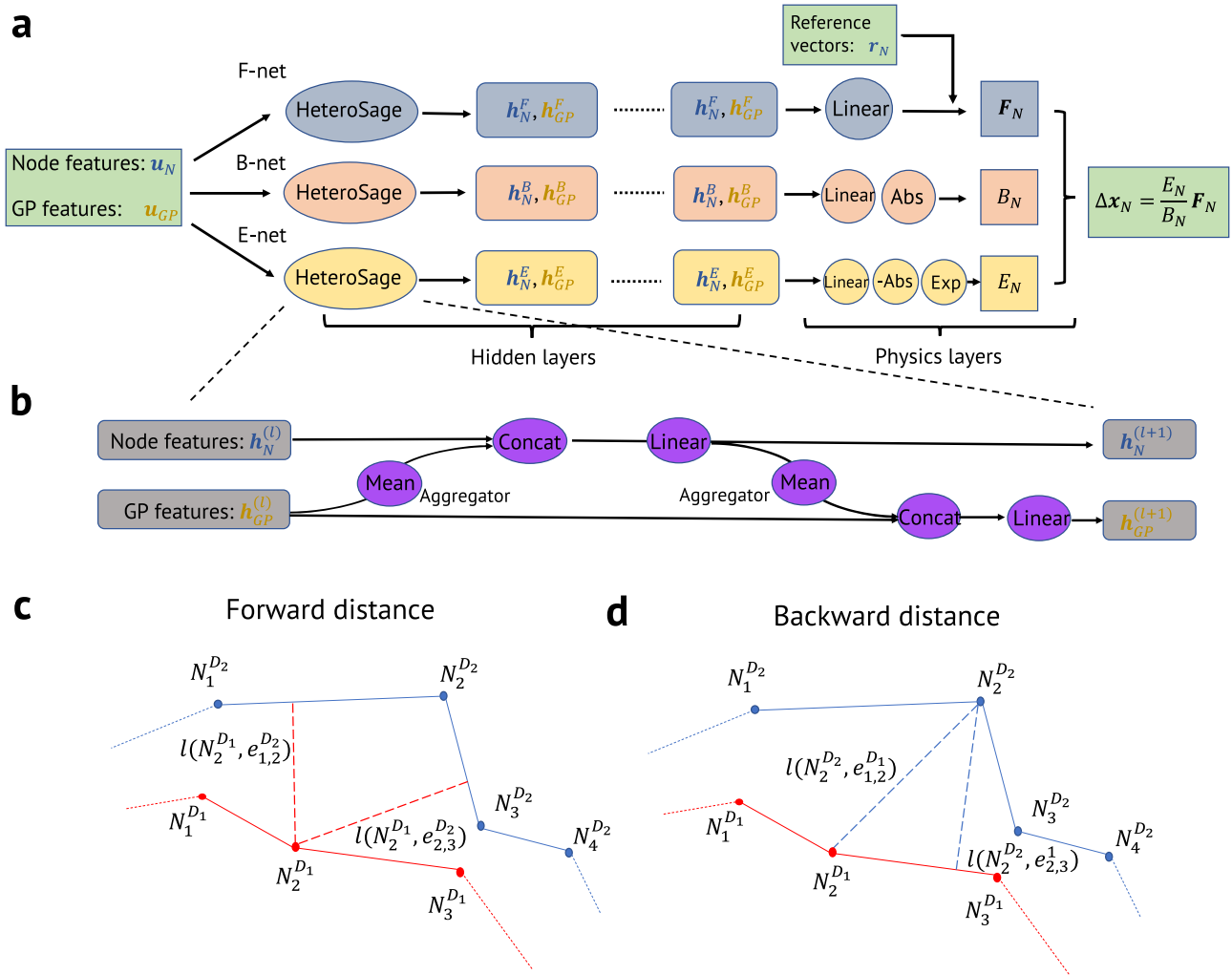
### Augmenting coarse-grained MD data with local features

Coarse-grained representations of dislocations present in the MD simulations are generated using the DXA algorithm. Using the same nodes identified by DXA and connecting the nodes via segments that are further discretized using Gauss points (GPs) (as shown in Fig. 1b), an exact geometrical representation of the MD dislocation can be replicated in MoDELlib, which is used to compute the local stress state at the GPs (Fig. 1c). The stress state along the dislocation network is computed using a non-singular isotropic linear elastic theory of dislocations<sup>33</sup>, and it includes dislocation-dislocation interactions as well as the external stress<sup>50–59</sup>. MoDELlib supports a periodic dislocation topology<sup>38</sup>, and the stress field of the periodic dislocation network is computed using an Ewald summation strategy<sup>60</sup>, which tapers the real-space long-range stress field by a user-defined characteristic length. Convergence of the stress calculation is typically realized using eight periodic images in each direction. Augmented coarse-grained MD snapshots are finally obtained, which consist of discretized dislocation lines using nodes, connectivities, and stress information at the GPs including both dislocation-dislocation interactions and external contributions. It is from these snapshots, containing both the geometry from MD and the local stress state computed by DDD that we move forward to construct the physical network and train the PI-GNN. More details about the methodology can be found in Supplemental Information.

### Modeling mobility law using Physics-informed Graph Neural Network (PI-GNN)

**Representing dislocations using heterogeneous graphs.** Graphs are a type of data structure that describes a set of objects, often referred to as vertices, and their relationships, or edges. Graphs are expressive and powerful at representing complex non-Euclidean data structures in social science, engineering, and physical science applications<sup>40,61</sup>. In the context of dislocation dynamics, we demonstrate the mapping between the physical and computational representation of a dislocation network in Fig. 1d, e. Physical representation consists of nodes and GPs, connected by the dislocation segments, illustrated by Fig. 1d. The physical features of the dislocation lines are defined on the GPs between two nodes, such as the Burger's vector, segment tangent vector, or local stress, while the features of the nodes consist only of positions. The data structure of a dislocation network can be easily mapped to a heterogeneous graph, where we introduce two types of vertices:  $v_N$  representing nodes and  $v_{GP}$  representing GPs in the physical network.

**Physics-informed graph neural network.** Figure 8a presents the structure of the graph neural network inspired by the physics-based phenomenological model of dislocation dynamics<sup>15</sup> as shown in Eq. (2). We decompose the PI-GNN mobility law into three components: F-net, B-net, and E-net, each with distinct input features and network structures, depending on the physics-based mobility function. As shown in Fig. 8a, we model the three network components with two types of layers, hidden GNN layers, which consist of a series of black-box layers for learning feature embedding from the graph data structure, and physics layers, in which physical constraints for each component of the mobility function can be injected. With these two types of layers, we aim to combine the functional expressiveness of GNNs and the interpretability of physical structure to construct a powerful, robust, and predictive model for dislocation dynamics. For the hidden layers, we extend the widely used GraphSAGE<sup>62</sup>, which were designed to operate on homogeneous graphs to sample and aggregate features from large graphs, to one that can operate on heterogeneous graphs. We refer to this extended model as HeteroSAGE in this work. Figure 8b illustrates the message-



**Fig. 8 | Overview of the PI-GNN architecture and training loss function.**

**a** Physics-informed graph neural network structure for learning the mobility law in Discrete Dislocation Dynamics (DDD), which consists of three different networks: F-net, B-net, and E-net. Each network consists of hidden layers, which are stacked HeteroSage GNNs, and physical layers, which are physically inspired functions

adopted from a phenomenological description of dislocation mobility. **b** The proposed HeteroSage GNN structure and message passing diagram for heterogeneous graphs. **c, d** Illustrations of the proposed generalized loss function for measuring the distance between two dislocation configurations with different total number of nodes resulting from different discretization.

passing and feature operations of each HeteroSAGE layer: the features defined on the neighboring vertices are aggregated using an aggregator function, which is then concatenated with local features and passed through a fully-connected network. This process is then repeated in the opposite direction to generate the output of one HeteroSAGE layer. By stacking multiple HeteroSAGE layers, the model can learn complex feature embeddings from the graph. This enables the PI-GNN model to learn non-local configuration-dependent corrections to the stress computations using linear theory. For problems tested in this work, we used two of HeteroSAGE layers, each with 20 nodes for each network. For the physics layers, we encode the mathematical structure of physics-based mobility law using differentiable functions. This design ensures that the output of each component adheres to the corresponding constraints dictated by the physics-based model. More details on the structure can be found in the Supplementary Information.

**Feature engineering.** In modern ML applications, deep neural networks, coupled with large datasets are commonly used to discover informative feature embeddings for effective predictions. However, in scientific applications, a set of carefully designed features can inject physics into the model, leading to improved prediction robustness and reduced dataset size requirements. In this work, we use a set of specially designed features from

nodes and GPs as the input of the PI-GNN mobility law. We leverage translational invariance and rotational invariance, which are important in physical systems, for guiding the feature engineering processes. The input features of the PI-GNN are designed such that the aforementioned constraints are automatically satisfied by the model, thus injecting crucial physics-based induction biases into the model. Details on the selected features can be found in the Supplementary Note 1.

**Generalized loss function.** After constructing the GNN model, we define a loss function that quantifies the differences between the model-predicted evolution of the dislocation and the ground truth, whose minimization will constrain the parameters of the PI-GNN during learning. However, a challenge arises when formulating the loss function from coarse-grained MD simulation snapshots, as the total number of nodes in dislocations separated by the time step  $\Delta t$  may differ and there is no clear one-to-one mapping between nodes. This hinders a straightforward computation of their differences using the commonly used Mean-Squared-Error (MSE) in traditional ML tasks. To resolve this issue, we propose a generalized loss function to measure the distances between two dislocations. As illustrated in Fig. 8c, d, the dislocations can be described as a collection of nodes and segments/edges connecting them. Instead of using the traditional point-wise positional difference to

construct the loss function, we define a node-to-segment distance based on the geometric projection as:

$$l(N_i^{D_1}, e_{j,k}^{D_2}) = \begin{cases} |\mathbf{x}_{N_i}^{D_1} - \mathbf{x}_{N_j}^{D_2}|, & \mathbf{x}_{N_i}^{D_1} \in e_{j,k}^{D_2}, \\ \min\{|\mathbf{x}_{N_i}^{D_1} - \mathbf{x}_{N_j}^{D_2}|, |\mathbf{x}_{N_i}^{D_1} - \mathbf{x}_{N_k}^{D_2}|\}, & \mathbf{x}_{N_i}^{D_1} \notin e_{j,k}^{D_2}. \end{cases} \quad (3)$$

With the node-to-segment distance  $l(N_i^{D_1}, e_{j,k}^{D_2})$  defined, we use the minimum value of the node  $N_i^{D_1}$  to all the segments  $e_{j,k}^{D_2} \in D_2$  as the node-to-dislocation distance  $l(N_i^{D_1}, D_2)$ , which can then be used to compute the dislocation-to-dislocation distance as the sum of the node-to-dislocation distances  $l(D_1, D_2) = \sqrt{\sum_i l(N_i^{D_1}, D_2)^2}$ . We notice that the proposed distance measure based on equations (3), is not symmetric, i.e.  $l(D_1, D_2) \neq l(D_2, D_1)$ . In this work, we adopt a loss function based on a symmetric distance measure, which is achieved by averaging the distances computed in both directions to avoid the potential issue due to the smoothing effects (see Supplementary Note 1 for more details):

$$L(D_1, D_2) = \frac{1}{2} [l(D_1, D_2) + l(D_2, D_1)], \quad (4)$$

$$Loss = \sum_i L(D_i^{GT}, D_i^{Pred}). \quad (5)$$

### Uncertainty quantification-driven active learning

**Ensemble learning for uncertainty quantification.** In the context of machine learning for scientific applications, particularly in cases where fully-resolved simulation data is sparse, the neural network-based model demonstrates competitive prediction accuracy. However, this may become problematic when a single, over-parameterized NN provides predictions for new datasets with unwarranted confidence. To address this issue, we adopt the ensemble learning approach to properly quantify the prediction uncertainty for modeling the mobility law using PI-GNN.

Ensemble learning is a machine learning training approach aimed at achieving improved prediction performance by combining the predictions from multiple independently-learned models. The objective of the ensemble learning approach lies in generating multiple instances of model parameters that span the local minima within the parameter-loss landscape of the NN-based model. We then approximate the PDF of the predictions using a Gaussian distribution, whose mean and variance are computed using the predictions from the top-performing ensemble models:

$$\hat{dx} \approx \bar{\mu} \left( \frac{dx}{dt} \right) = \frac{2}{M_{en}} \sum_{i=1}^{M_{en}/2} \left. \frac{dx}{dt} \right|_{\theta_i}, \quad (6)$$

$$\sigma^2 \left( \frac{dx_j}{dt} \right) \approx \frac{2}{M_{en}} \sum_{i=1}^{M_{en}/2} \left( \left. \frac{dx_j}{dt} \right|_{\theta_i} - \hat{dx}_j \right)^2, \quad (7)$$

where  $dx/dt|_{\theta_i}$  refer to the PI-GNN prediction of the nodal displacement using the parameter  $\theta_i$ . Through the ensemble learning procedure, we can obtain an estimate of the distribution of PI-GNN outputs, which can inform the UQ of predictions.

**Active learning.** Active learning is a special procedure in ML that can interactively query a new dataset based on a pre-determined strategy. The active learning approach is commonly used in supervised learning setups where it is very expensive to manually enumerate new data points for the query. When a proper strategy for selecting new data points is used, the total number of data points to learn a model can be dramatically reduced compared to a passive learning strategy with a randomly queried dataset.

Figure 2 shows the workflow for the UQ-based active learning loop, which includes data generation, model training, UQ and data point selection. The details of the active learning procedure are described below:

1. Generate initial MD dataset with a set of randomly selected simulation input features  $\Lambda = \{\lambda_i\}$ , where  $\lambda_i$  can be different combinations of temperature, applied stress, orientations, Burger's vector, etc. The simulated MD trajectories are then coarse-grained and post-processed to obtain snapshots of the dislocations nodes/GPs and their features. We denote the MD dataset with parameters  $\Lambda$  as  $\mathbf{D}(\Lambda)$ .
2. Train an ensemble of  $M_{en}$  PI-GNNs on the dataset  $\mathbf{D}(\Lambda)$  each with random initialization of model parameters. We employ a train/test/validation data split ratio of 70%/20%/10%.
3. Randomly sample a new set of MD simulation input features  $\Lambda'$  and run new MD simulations with them and regroup with existing test dataset to generate a test dataset  $\mathbf{D}_t(\Lambda')$ .
4. Take the top 50% ( $M_{en}/2$ ) performing trained PI-GNN models and use the ensemble variance of the nodal displacement prediction as a measure of prediction uncertainty.
5. If the highest uncertainty level is lower than a preset threshold, stop training. Otherwise, combine the most uncertain set of the test data with the training data to form a new set of training data, i.e.  $\mathbf{D}(\Lambda \cup \Lambda')$ , which is then used to re-train (using previously converged model parameters as the initial weights) the ensemble models.
6. Move to step 3.

### Data availability

The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

### Code availability

The codes we developed to perform the experiments are available from the corresponding author on reasonable request.

Received: 26 March 2024; Accepted: 15 August 2024;

Published online: 14 September 2024

### References

1. Orowan, E. Zur kristallplastizität. iii Über den mechanismus des gleitvorganges. *Z. Phys.* **89**, 634–659 (1934).
2. Polanyi, M. Über eine art gitterstörung, die einen kristall plastisch machen könnte. *Z. Phys.* **89**, 660–664 (1934).
3. Taylor, G. I. The mechanism of plastic deformation of crystals. part i.—theoretical. *Proc. R. Soc. A. Math. Phys. Char.* **145**, 362–387 (1934).
4. Hull, D. & Bacon, D. Movement of dislocations. In *Introduction to Dislocations*. (eds Hull, D. & Bacon, D.) Ch. 3 (Butterworth-Heinemann, Oxford, 2011).
5. Brailsford, A. Anharmonicity contributions to dislocation drag. *J. Appl. Phys.* **43**, 1380–1393 (1972).
6. Swinburne, T., Dudarev, S. & Sutton, A. Classical mobility of highly mobile crystal defects. *Phys. Rev. Lett.* **113**, 215501 (2014).
7. Swinburne, T. & Dudarev, S. Phonon drag force acting on a mobile crystal defect: Full treatment of discreteness and nonlinearity. *Phys. Rev. B Condens. Matter Mater. Phys.* **92**, 134302 (2015).
8. Christian, J. W. Some surprising features of the plastic deformation of body-centered cubic metals and alloys. *Metall. Trans. A* **14**, 1237–1256 (1983).
9. Gröger, R., Bailey, A. & Vitek, V. Multiscale modeling of plastic deformation of molybdenum and tungsten: I. atomistic studies of the core structure and glide of 1/2(111) screw dislocations at 0k. *Acta Mater.* **56**, 5401–5411 (2008).
10. Hale, L. M., Lim, H., Zimmerman, J. A., Battaile, C. C. & Weinberger, C. R. Insights on activation enthalpy for non-schmid slip in body-centered cubic metals. *Scr. Mater.* **99**, 89–92 (2015).
11. Anderson, P., Hirth, J. & Lothe, J. *Theory of Dislocations* 3rd edn, Vol. 718 (Cambridge University Press, 2017).
12. Lothe, J. Theory of dislocation mobility in pure slip. *J. Appl. Phys.* **33**, 2116–2125 (1962).

13. Christian, J. & Vitek, V. Dislocations and stacking faults. *Rep. Prog. Phys.* **33**, 307–411 (1970).
14. Nabarro, F. *Theory of Crystal Dislocations. International Series of Monographs on Physics* 1st edn (Clarendon Press, 1967).
15. Po, G. et al. A phenomenological dislocation mobility law for bcc metals. *Acta Mater.* **119**, 123–135 (2016).
16. Gilbert, M., Queyreau, S. & Marian, J. Stress and temperature dependence of screw dislocation mobility in  $\alpha$ -Fe by molecular dynamics. *Phys. Rev. B* **84**, 174103 (2011).
17. Queyreau, S., Marian, J., Gilbert, M. & Wirth, B. Edge dislocation mobilities in bcc Fe obtained by molecular dynamics. *Phys. Rev. B* **84**, 064106 (2011).
18. Maresca, F., Dragoni, D., Csányi, G., Marzari, N. & Curtin, W. A. Screw dislocation structure and mobility in body centered cubic Fe predicted by a Gaussian approximation potential. *npj Comput. Mater.* **4**, 69 (2018).
19. Kang, K., Bulatov, V. V. & Cai, W. Singular orientations and faceted motion of dislocations in body-centered cubic crystals. *Proc. Natl Acad. Sci. USA.* **109**, 15174–15178 (2012).
20. Cho, J., Molinari, J.-F. & Anciaux, G. Mobility law of dislocations with several character angles and temperatures in fcc aluminum. *Int. J. Plast.* **90**, 66–75 (2017).
21. Wu, B. et al. <110> superdislocation mobility with different character angles in ni3Al. *Intermetallics* **166**, 108161 (2024).
22. Yamaguchi, M. & Vitek, V. Core structure of nonscrew 1/2(111) dislocations on (110) planes in b.c.c. crystals. i. core structure in an unstressed crystal. *J. Phys. F Metal Phys.* **3**, 523 (1973).
23. Varvenne, C., Luque, A. & Curtin, W. A. Theory of strengthening in fcc high entropy alloys. *Acta Mater.* **118**, 164–176 (2016).
24. Rida, A., Martinez, E., Rodney, D. & Geslin, P.-A. Influence of stress correlations on dislocation glide in random alloys. *Phys. Rev. Mater.* **6**, 033605 (2022).
25. Chen, B. et al. Unusual activated processes controlling dislocation motion in body-centered-cubic high-entropy alloys. *Proc. Natl Acad. Sci.* **117**, 16199–16206 (2020).
26. Wang, F. et al. Multiplicity of dislocation pathways in a refractory multiprincipal element alloy. *Science* **370**, 95–101 (2020).
27. Montes de Oca Zapiain, D. et al. Training data selection for accuracy and transferability of interatomic potentials. *npj Comput. Mater.* **8**, 189 (2022).
28. LeSar, R. & Capolungo, L. Advances in discrete dislocation dynamics simulations. In *Handbook of Materials Modeling*. (eds. Andreoni, W., Yip, S.) 1–10 (Springer International Publishing, Cham, 2020).
29. Po, G. et al. Recent progress in discrete dislocation dynamics and its applications to micro plasticity. *JOM* **66**, 2108–2120 (2014).
30. Zhou, C., Biner, S. B. & LeSar, R. Discrete dislocation dynamics simulations of plasticity at small scales. *Acta Mater.* **58**, 1565–1577 (2010).
31. Groh, S. & Zbib, H. M. Advances in discrete dislocations dynamics and multiscale modeling. *J. Eng. Mater. Technol.* **131**, 041209 (2009).
32. Mordehai, D., Clouet, E., Fivel, M. & Verdier, M. Introducing dislocation climb by bulk diffusion in discrete dislocation dynamics. *Philos. Mag.* **88**, 899–925 (2008).
33. Cai, W., Arsenlis, A., Weinberger, C. R. & Bulatov, V. V. A non-singular continuum theory of dislocations. *J. Mech. Phys. Solids* **54**, 561–587 (2006).
34. Barros de Moraes, E. A., Suzuki, J. L. & Zayernouri, M. Atomistic-to-meso multi-scale data-driven graph surrogate modeling of dislocation glide. *Comput. Mater. Sci.* **197**, 110569 (2021).
35. Bertin, N. & Zhou, F. Accelerating discrete dislocation dynamics simulations with graph neural networks. *J. Comput. Phys.* **487**, 112180 (2023).
36. Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A. & Vandergheynst, P. Geometric deep learning: going beyond euclidean data. *IEEE Signal Process. Mag.* **34**, 18–42 (2017).
37. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO—the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
38. Pachaury, Y., Po, G. & El-Azab, A. Discrete dislocation dynamics for crystal rves. part 1: periodic network kinematics. *J. Mech. Phys. Solids* **163**, 104861 (2022).
39. Krogh, A. & Vedelsby, J. *Neural Network Ensembles, Cross Validation, and Active Learning*. [https://proceedings.neurips.cc/paper\\_files/paper/1994/file/b8c37e33defde51cf91e1e03e51657da-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/1994/file/b8c37e33defde51cf91e1e03e51657da-Paper.pdf) (1994).
40. Bronstein, M. M., Bruna, J., Cohen, T. & Veličković, P. Geometric deep learning: grids, groups, graphs, geodesics, and gauges. *arXiv* <https://doi.org/10.48550/arXiv.2104.13478> (2021).
41. Garnett, R. *Bayesian Optimization* 1st edn (Cambridge University Press, 2023).
42. Kazemi, S. M. et al. Representation learning for dynamic graphs: a survey. *J. Mach. Learn. Res.* **21**, 70:1–70:73 (2020).
43. Zhang, M., Wu, S., Yu, X., Liu, Q. & Wang, L. Dynamic graph neural networks for sequential recommendation. *IEEE Trans. Knowl. Data Eng.* **35**, 4741–4753 (2023).
44. Bertin, N., Bulatov, V. V. & Zhou, F. Learning dislocation dynamics mobility laws from large-scale MD simulations. *arXiv* <https://doi.org/10.48550/arXiv.2309.14450> (2023).
45. Cereceda, D., Perlado, J. M. & Marian, J. Techniques to accelerate convergence of stress-controlled molecular dynamics simulations of dislocation motion. *Comput. Mater. Sci.* **62**, 272–275 (2012).
46. Cai, W., Bulatov, V. V., Chang, J., Li, J. & Yip, S. Periodic image effects in dislocation modelling. *Philos. Mag.* **83**, 539–567 (2003).
47. Mendeleev, M. et al. Development of new interatomic potentials appropriate for crystalline and liquid iron. *Philos. Mag.* **83**, 3977–3994 (2003).
48. Thompson, A. P. et al. LAMMPS—a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Comput. Phys. Commun.* **271**, 108171 (2022).
49. Stukowski, A., Bulatov, V. V. & Arsenlis, A. Automated identification and indexing of dislocations in crystal interfaces. *Model. Simul. Mater. Sci. Eng.* **20**, 085007 (2012).
50. Lepinoux, J. & Kubin, L. P. The dynamic organization of dislocation-structures: a simulation. *Scr. Metall. Mater.* **21**, 833–838 (1987).
51. Ghoniem, N. M. & Amodeo, R. Computer simulation of dislocation pattern formation. *Sol. St. Phen.* **3,4**, 377–388 (1988).
52. Gulluoglu, A. N., Srolovitz, D. J., LeSar, R. & Lomdahl, P. S. Dislocation distributions in two dimensions. *Scr. Metall.* **23**, 1347–1352 (1989).
53. Kubin, L. P. et al. Dislocation microstructures and plastic flow: a 3D simulation. *Sol. St. Phen.* **23**, 455–472 (1992).
54. Schwarz, K. W. Interaction of dislocations on crossed glide planes in a strained epitaxial layer. *Phys. Rev. Lett.* **78**, 4785–4788 (1997).
55. Zbib, H., Rhee, M. & Hirth, J. On plastic deformation and the dynamics of 3D dislocations. *Int. J. Mech. Sci.* **40**, 113–127 (1998).
56. Ghoniem, N. M., Tong, S. & Sun, L. Parametric dislocation dynamics: a thermodynamics-based approach to investigations of mesoscopic plastic deformation. *Phys. Rev. B* **61**, 913–927 (2000).
57. Weygand, D., Friedman, L., Van der Giessen, E. & Needleman, A. Aspects of boundary-value problem solutions with three-dimensional dislocation dynamics. *Model. Simul. Mater. Sci.* **10**, 437–468 (2002).
58. Bulatov, V. V., Cai, W., Fier, J. & Hiratani, M. Scalable line dynamics in ParaDiS. In Proc. 2004 ACM/IEEE Conference on Supercomputing. 19 (Pittsburgh, PA, 2004).
59. Po, G. & Ghoniem, N. M. A variational formulation of constrained dislocation dynamics coupled with heat and vacancy diffusion. *J. Mech. Phys. Solids* **66**, 103–116 (2014).
60. Wells, B. A. & Chaffee, A. L. Ewald summation for molecular simulations. *J. Chem. Theory Comput.* **11**, 3684–3695 (2015).
61. Zhou, J. et al. Graph neural networks: a review of methods and applications. *AI Open* **1**, 57–81 (2020).

62. Hamilton, W., Ying, Z. & Leskovec, J. Inductive representation learning on large graphs. *Adv. Neural Inform. Proc. Sys.* <https://doi.org/10.48550/arXiv.1706.02216> (2017).

## Acknowledgements

The team is supported by the Laboratory Directed Research and Development (LDRD) Project “Accelerated Dynamics Across Computational and Physical Scales” (20220063DR). The study has been approved for unlimited release with LA-UR-23-26562. The authors acknowledge significant support from the Darwin test bed at Los Alamos National Laboratory (LANL), which is funded by the Computational Systems and Software Environments subprogram of LANL’s Advanced Simulation and Computing program (NNSA/DOE), and Institutional High Performance Computing.

## Author contributions

Y.T.—ML methodology, software, initial draft, S.B.—framework/software for MD data generation, generation of ground truth MD data, initial draft, L.M.—software to featurize MD data for ML, G.P.—MoDELlib developments, E.M.—conceptualization, MD data validation, Y.T.L.—ML methodology, design of study, supervision, initial draft, N.M.—MD methodology, design of study, supervision, initial draft, D.P.—conceptualization, funding acquisition, supervision. All authors participated in discussing the results and writing of the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41524-024-01394-4>.

**Correspondence** and requests for materials should be addressed to Yifeng Tian.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024