Machine learning reveals strain-rate-dependent predictability of discrete dislocation plasticity

Marcin Mińkowski [ID],[*] David Kurunczi-Papp [ID], and Lasse Laurson [ID]
Computational Physics Laboratory, Tampere University, P.O. Box 692, FI-33014 Tampere, Finland

Predicting the behavior of complex systems is one of the main goals of science. An important example is plastic deformation of micron-scale crystals, a process mediated by collective dynamics of dislocations, manifested as broadly distributed strain bursts and significant sample-to-sample variations in the response to applied loading. Here, by combining large-scale discrete dislocation dynamics simulations and machine learning, we study the problem of predicting the fluctuating stress-strain curves of individual small single crystals subject to strain-controlled loading using features of the initial dislocation configurations as input. Our results reveal an intriguing rate dependence of deformation predictability: For small strains predictability improves with increasing strain rate, while for larger strains the predictability vs strain rate relation becomes nonmonotonic. We show that for small strains the rate dependence of deformation predictability can be captured by considering the fraction of dislocations moving against the direction imposed by the external stress, serving as a measure of strain-rate-dependent complexity of the dislocation dynamics. The nonmonotonic predictability vs strain rate relation for large strains is argued to be related to a transition from fluctuating to smooth plastic flow when strain rate is increased.

I. INTRODUCTION

Dislocation-mediated plasticity of crystalline solids is characterized by size effects implying that smaller systems tend to be "stronger" and exhibit larger fluctuations in their stress-strain response [1,2], as well as by rate effects which are typically manifested as larger average stresses needed to reach a given strain for higher strain rates [3], a phenomenon often referred to as strain rate sensitivity [4]. Understanding the combined effects of these key features of dislocation plasticity constitutes an important challenge in physics and materials science: For instance, the strain bursts exhibited by micropillars under compression [5] are interesting both from the perspective of fundamentals of nonequilibrium statistical physics [6,7] as well as in applications where the fluctuations of plastic deformation result in difficulties to control the resulting shape in plastic forming processes [8]. At the same time, the fact that in various applications the mechanical response of metals depends significantly on the loading rate (e.g., during a car crash) is of great importance and calls for novel approaches to understand the underlying fundamental physics and to control rate dependence of dislocation plasticity.

Recent years have witnessed a surge in application of artificial intelligence (AI) in general and machine learning (ML) in particular to gain novel insights on properties of materials and related problems in physics [9–19]. Broadly speaking, such developments fall under the umbrella of the emerging research field of *materials informatics* [20], where informatics methods—including ML—are used to search for novel

materials [21,22], establish novel structure-property relations [13,23], etc. A closely related problem is given by quantification of *deformation predictability* [24,25] of small solid samples where the response to applied stresses (i.e., the stress-strain curve) exhibits significant sample-to-sample variations. The general problem statement can be formulated as follows: Given a description of the initial microstructure (defined, e.g., by the arrangement of pre-existing dislocations and, if present, other defects within the crystal) of such a sample, how well can its subsequent sample-specific deformation dynamics be predicted when external loading is applied? Such questions are closely related to the more general issue of predictability of "complex systems," ranging from forecasting earthquakes [26–28] to predicting volcanic eruptions [29] and beyond [30].

Such predictions tend to be difficult as the mapping from the various features describing the initial state of the system to its subsequent behavior or dynamics is usually highly complicated. This complexity typically originates from the combination of nonlinear nature of the collective dynamics governing the behavior of the system and the high dimensionality of the feature set characterizing the system's state. On the other hand, while traditional forecasting methods struggle with such complexity, typical ML algorithms are well suited to address problems of this type. Our recent numerical work based on combining ML and discrete dislocation dynamics (DDD) simulations suggests that under quasistatic stress-controlled loading the degree of deformation predictability of plastically deforming crystals is controlled by the collective nature of dislocation dynamics, manifested as inherently hard to predict dislocation avalanches [24,25]. It is worth noting that even in situations where the dislocation motion is governed by deterministic equations of motion, the critical-like and possibly chaotic [31] dynamics exhibited by the system

---

*marcin.minkowski@tuni.fi

renders the prediction problem highly nontrivial [25]. These works highlight the usefulness of ML in providing a metric for deformation predictability, allowing us to also identify the key features of the initial microstructure that determine the mechanical properties of the sample. Yet, many outstanding questions related to deformation predictability remain to be addressed, including how rate dependence of plasticity is manifested in predictability of the deformation process.

Here, we combine large-scale strain-controlled DDD simulations [32] and ML to study how the imposed strain rate and the stiffness of the specimen-machine system affect deformation predictability of small single crystals. By employing ML models including linear regression, fully connected neural networks and convolutional neural networks, trained using data from our extensive 2D DDD simulations, we find that the ability of all these rather different ML models to predict the stress-strain curve using features of the initial dislocation configuration as input, quantified by the correlation between predicted and actual stress at a given strain, exhibits a similar dependence on the imposed strain rate. Hence, by employing the definition of deformation predictability used in Refs. [24,25], we conclude that deformation predictability of the 2D DDD model we study is rate dependent. For small strains, predictability improves with increasing strain rate, while for large strains the predictability vs strain rate relation becomes nonmonotonic. Interestingly, especially for small strains, we demonstrate a clear link between deformation predictability and the mean fraction of dislocations moving against the direction where the applied stress is driving them, such that deformation predictability improves when a smaller fraction of the dislocations move against the direction set by the applied stress. This fraction increases with decreasing strain rate, is a measure of how "interaction dominated" the dislocation dynamics is, and relates to the sizes and complexity of the multipolar dislocation structures moving collectively during the deformation process. Specifically, our results suggest that the more complex dislocation dynamics for low strain rates results in a larger prediction error of the ML models. We also discuss the effect of varying the stiffness of the specimen-machine system, as well as the role of the various features characterizing the initial dislocation configuration in determining the stress-strain response, including how these give rise to the nonmonotonic predictability vs strain rate relation for large strains. The latter feature is also demonstrated to be linked to a transition from fluctuating to smooth stress-strain curves upon increasing the imposed strain rate: At the transition point the stress-strain curve is typically dominated by an individual large intrinsically hard-to-predict dislocation avalanche, resulting in a larger prediction error of the ML models.

The paper is organized as follows: We start by going through the methods employed (Sec. II), including the DDD simulations in Sec. II A, input features used for the ML models in Sec II B, and the ML models considered (linear regression, fully connected neural networks and convolutional neural networks in Secs. II C, II D, and II E, respectively). This is then followed by presentation of our results in Sec. III, including details about dataset generation in Sec. III A, machine learning results in Sec. III B, and an account on how these are related to the rate-dependent complexity of dislocation

dynamics in Secs. III C and III D. The paper is finished with discussion and conclusions in Sec. IV.

II. METHODS

A. DDD simulations

The 2D DDD model we use is similar to the models studied in a number of previous works [6,7,32,33]. The dynamics of straight parallel edge dislocation lines oriented along the z direction (represented here by points in the xy plane) and moving within a single slip geometry along the x direction is taken to follow overdamped equations of motion where the dislocation velocities are proportional to the total Peach-Koehler force acting on them,

$$\frac{\dot{x}_i}{Mb} = s_i b \left[\sum_{j \neq i}^{N} s_j \sigma_{\mathrm{disl}}(\mathbf{r}_i - \mathbf{r}_j) + \sigma_{\mathrm{ext}} \right], \quad (1)$$

where N is the total number of dislocations in the system, $\mathbf{r}_i = (x_i, y_i)$ is the position of the dislocation i, M is the dislocation mobility, b is the length of the Burgers vector, and $s_i = \pm 1$ is its sign. Each dislocation moves in response to both the externally applied shear stress σ_{ext} and the internal shear stress produced by the other dislocations in the system, given by a sum over the other dislocations of $\sigma_{\mathrm{disl}}(\mathbf{r}) = \sigma_{\mathrm{disl}}(x, y) = Dbx(x^2 - y^2)/(x^2 + y^2)^2$, where $D = \mu/2\pi(1 - \nu)$, with μ and ν the shear modulus and Poisson ratio, respectively. Additionally, during the simulation two dislocations of opposite sign annihilate if the distance between them becomes smaller than b and periodic boundary conditions are implemented. In the simulations all lengths, times, and stresses are measured in units of b, $(MDb)^{-1}$ and D, respectively.

Creating a dataset proceeds by generating configurations containing a certain number of randomly distributed dislocations within a square simulation box, and subsequently relaxing them into a metastable state by running the DDD simulation (i.e., numerically integrating Eq. (1) for all the dislocations) at zero external shear stress, i.e., $\sigma_{\mathrm{ext}} = 0$. Because of the aforementioned annihilation processes the number of dislocations in the final state is always smaller than in the unrelaxed configuration. The relaxed state is then used as the initial state for the subsequent strain-controlled loading, during which σ_{ext} is applied according to the protocol

$$\sigma_{\mathrm{ext}} = k[\dot{\epsilon}_a t - \epsilon(t)], \quad (2)$$

where $\epsilon(t) = b/L^2 \sum_{i=1}^{N} s_i \Delta x_i(t)$ is the strain at time t, with $\Delta x_i(t)$ denoting the displacement of the ith dislocation, $\dot{\epsilon}_a$ is the applied strain rate, and k is the spring stiffness, modeling the combined stiffness of the specimen-machine system [32]. Varying k could mean that stiffness of the loading machine or the shear modulus of the sample changes, or both of them change simultaneously. Taking the liberty of adjusting k allows us to study its influence on deformation predictability. The above relation means that during the simulation the external stress σ_{ext} is adjusted in such a way that the true plastic strain rate $\dot{\epsilon}(t) = b/L^2 \sum_{i=1}^{N} s_i \dot{x}_i(t)$ approaches the applied one, $\dot{\epsilon}_a$, in the long-time limit, and thereafter fluctuates in its vicinity. The stiffness k measures how strongly σ_{ext} reacts to deviations of $\dot{\epsilon}(t)$ from the applied strain rate $\dot{\epsilon}_a$. Thus, higher values of k help to approach the applied strain rate sooner
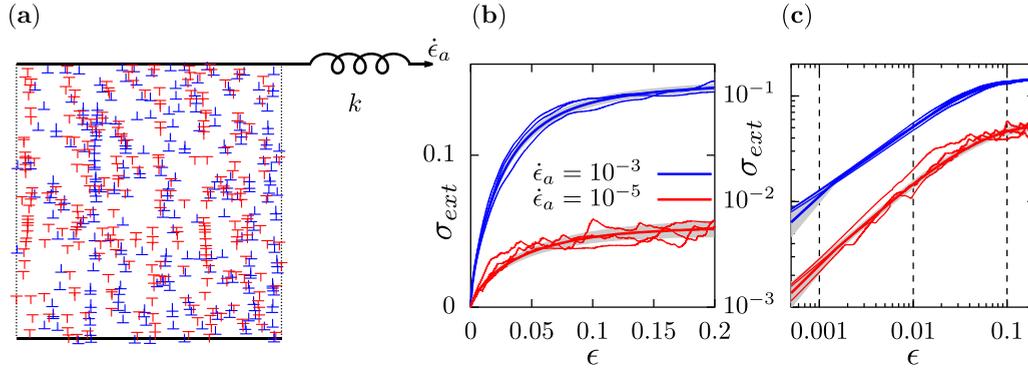
FIG. 1. Schematic of the DDD model with strain-controlled loading and examples of the resulting stress-strain curves. (a) Strain-controlled loading of a crystal containing dislocations with positive (red) and negative (blue) Burgers vectors, with the applied strain rate $\dot{\epsilon}_a$ and the stiffness of the specimen-machine system, modelled with a spring of stiffness k. (b) Stress-strain curves for a few samples for $k = 1$ and two different values of $\dot{\epsilon}_a$ plotted with thin lines in linear scale, with the thick line representing the average over all the samples and the grey area indicating their standard deviation. (c) The same curves in the logarithmic scale, with the vertical dashed lines representing the values of strain for which the predictive algorithms used in the work are trained.

and result in σ_{ext} exhibiting larger fluctuations. Concurrently, larger k values lead to smaller fluctuations of $\dot{\epsilon}$ around $\dot{\epsilon}_a$. Hence, the finite stiffness k of the specimen-machine system implies that we consider an intermediate case between "hard" (strain controlled with an infinitely stiff machine) and "soft" (stress-controlled) driving [32]. Figure 1(a) shows a schematic of the simulation setup, with examples of stress-strain curves obtained from the simulations in Figs. 1(b) and 1(c) for two different $\dot{\epsilon}_a$ values.

B. Input features for machine learning

In this work three different ML algorithms are used to predict the stress-strain curve based on the initial relaxed dislocation configuration and the parameters of the strain-controlled loading (strain rate $\dot{\epsilon}_a$ and the spring stiffness k). These algorithms are linear regression, fully connected neural network (NN), and convolutional neural network (CNN). The first two require manually defined descriptors that characterize the initial state. Those are chosen as the statistical features of fields defined on the configurations: density of dislocations $\rho(x, y)$, density of geometrically necessary dislocations (GND) $\rho_{\text{GND}}(x, y)$, defined as the difference between the density of positive and negative dislocations, and the internal stress field due to all the dislocations in the system, $\sigma_{\text{int}}(x, y) = \sum_i s_i \sigma_{\text{disl}}(\mathbf{r} - \mathbf{r_i})$. $\rho(x, y)$ and $\rho_{\text{GND}}(x, y)$ are sampled over 25 slices in both directions and $\sigma_{\text{int}}(x, y)$ over 300 slices. Because the number of positive and negative dislocations in the system is equal, the GND density is obviously zero globally, however, due to the dislocation density fluctuations, it has nonzero values on the local scale. Since the stress field diverges to infinity at the dislocation core, the limit $|\sigma_{\text{int}}| < 2.0$ is imposed for any $\mathbf{r}$. Fig. 2 shows examples of these fields. For the visualization purpose the limit $|\sigma_{\text{int}}| < 0.2$ was chosen because most of the values are within this range.

The following features are extracted from the fields: the first Fourier coefficient in the x and y directions (f_{x1} and

f_{y1}), and the mixed coefficient f_{xy} defined for the field $A(x, y)$ as

$$f_{x1} = \sum_{j=0}^{n} \sum_{k=0}^{m} A(j, k) \exp\left[-2\pi i \left(\frac{j}{n+1}\right)\right], \quad (3)$$

$$f_{y1} = \sum_{j=0}^{n} \sum_{k=0}^{m} A(j, k) \exp\left[-2\pi i \left(\frac{k}{m+1}\right)\right], \quad (4)$$

$$f_{xy} = \sum_{j=0}^{n} \sum_{k=0}^{m} A(j, k) \exp\left[-2\pi i \left(\frac{j}{n+1}\right)\right]$$
$$\times \exp\left[-2\pi i \left(\frac{k}{m+1}\right)\right], \quad (5)$$

where the summation is done over all the slices in both directions, as well as the average, kurtosis and skewness. Since the Fourier coefficients are in general complex numbers, their absolute values are used as the input. The other features are determined both on the original field and its absolute value with the exception of the dislocation density field $\rho(x, y)$, which is non-negative by definition. Moreover, since the averages of the GND density and stress field are by definition equal to 0, only the features defined on the absolute values of those fields are chosen. In total $N_{in} = 22$ descriptors are used, with all of them listed in Table I.

C. Linear regression

Linear regression is the simplest model employed in this work. It assumes that the external stress is a linear function of all the N_{in} descriptors of the values x_i, that is, $\sigma_{\text{ext}} = a_0 + \sum_{i=1}^{N_{in}} a_i x_i$, where the parameters a_i are optimized using the least squares method in such a way that the loss function is as small as possible. The loss function is defined as the sum of the squares of the differences between the predicted and the true external stress value. 80% of the configurations were used as the training set and 20% as the test set. Moreover, L2 regularization is utilized, in which a penalty term $\lambda \sum_{i=1}^{N_{in}} a_i^2$,
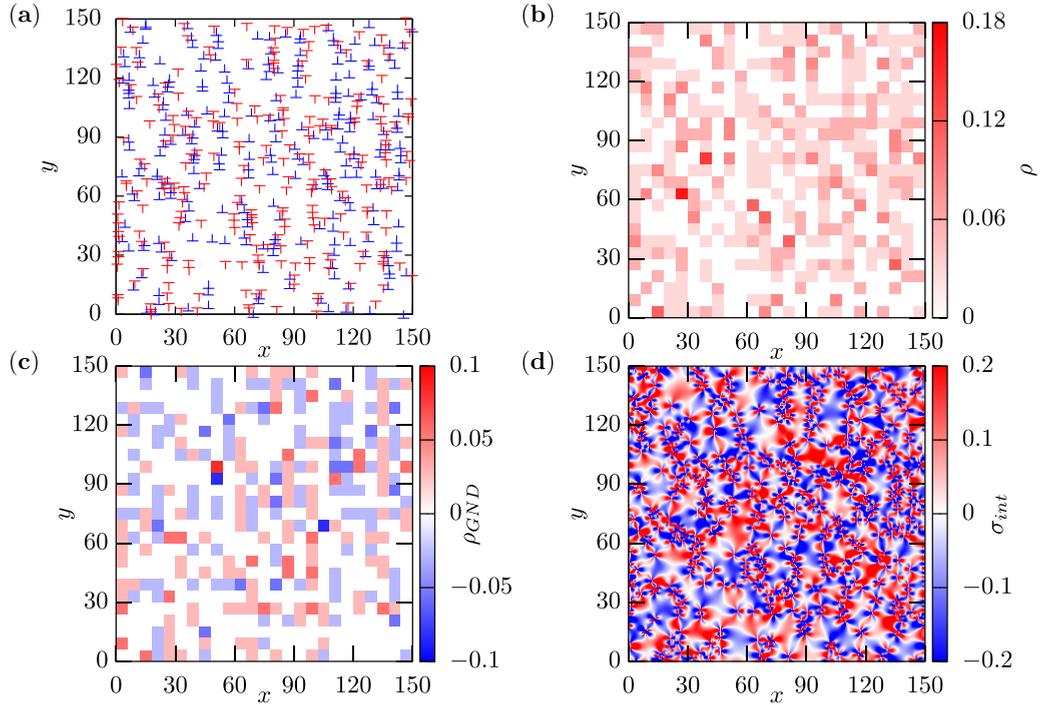
FIG. 2. Fields used as input for linear regression and fully connected neural networks. (a) An example initial configuration of the dislocations, (b) the corresponding density field of dislocations $\rho(x, y)$, (c) density of geometrically necessary dislocations $\rho_{\mathrm{GND}}(x, y)$, and (d) the internal stress field $\sigma_{\mathrm{int}}(x, y)$ produced by the dislocations. Various features derived from these fields (see Methods) are used as input for linear regression and fully connected neural networks.

where the factor $\lambda = 0.001$, is added to the loss function, however, in this work it does not change the final results.

D. Fully connected neural network

A fully connected neural network (NN) employs a series of hidden layers inserted between the input and the output layer. The value of the mth node in the nth layer, denoted as y_m^n, is in general a nonlinear function of values of all the nodes in the $(n - 1)$th layer in the following way:

$$y_m^n = f_a^{n-1}\left(w_{0m}^{n-1} + \sum_{i=1}^{N_{n-1}} w_{im}^{n-1} y_i^{n-1}\right), \qquad (6)$$

where f_a^{n-1} is the activation function between the layers $n - 1$ and n, N_{n-1} is the number of nodes in the layer $n - 1$, and w_{im}^{n-1} are parameters of the NN. For $i \neq 0$ they are called weights and they multiply the values of the nodes in the previous layer, and for $i = 0$ they are called biases, which are added to the argument of the activation function.

TABLE I. List of all the features used for training linear regression and NN.

Field	Features
ρ_{GND}	$\|f_{x1}\|$, $\|f_{y1}\|$, $\|f_{xy}\|$, kurtosis, skewness
$\|\rho_{\mathrm{GND}}\|$	average, kurtosis, skewness
ρ	$\|f_{x1}\|$, $\|f_{y1}\|$, $\|f_{xy}\|$, average, kurtosis, skewness
σ_{int}	$\|f_{x1}\|$, $\|f_{y1}\|$, $\|f_{xy}\|$, kurtosis, skewness
$\|\sigma_{\mathrm{int}}\|$	average, kurtosis, skewness

The schematic representation of the NN used in this work is shown in Fig. 3. It consists of three hidden layers contains five neurons each. The rectifier function is chosen as the activation function, except for the output layer, where the linear function is used.

The NN training is performed in Python using the Keras library with the Adam optimizer with the learning rate set to 0.001. During each epoch the parameters of the NN, weights and biases, are updated. 80% of the configurations constitute the training set, and the rest is divided equally into the test and validation set. The validation set is used as the stopping criterion for the training. The value of the loss function is
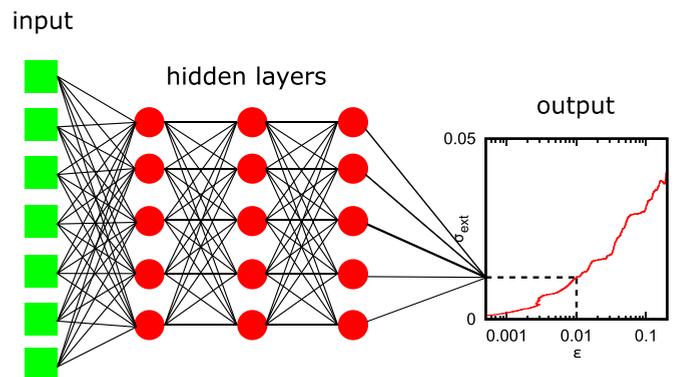
FIG. 3. Schematic representation of the fully connected neural network. The input values are passed to the neurons in the hidden layers through the activation function to get the external stress value as the final result.
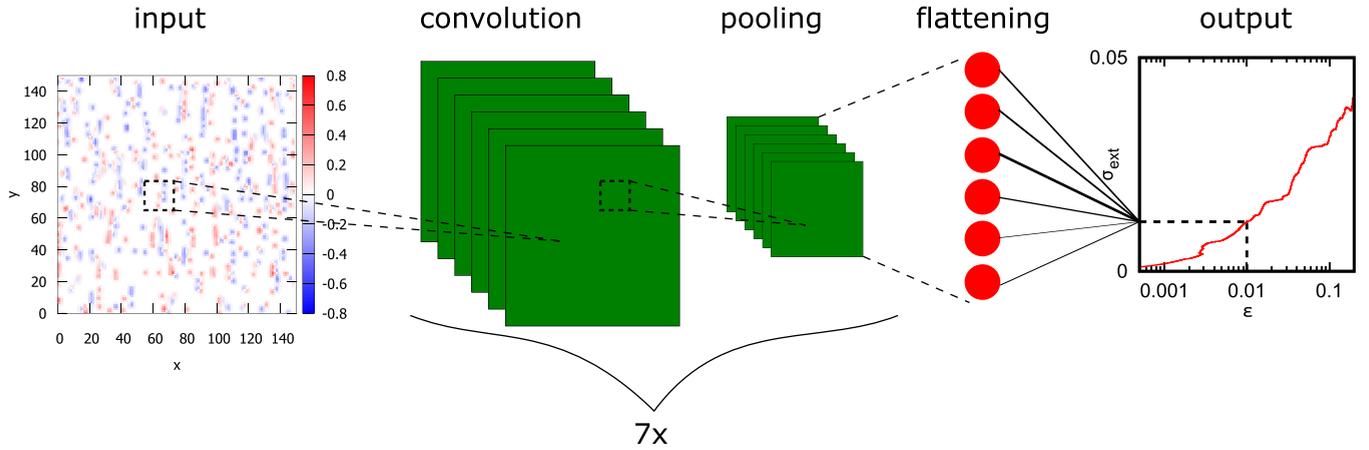
FIG. 4. Schematic representation of the convolutional neural network. On the left the image representing the dislocation configuration is shown. It is subsequently transformed by the convolutional and the pooling layer. Those two operations are performed seven times, after which the array is flattened and passed to the fully connected neural network on the right. The final result is the external stress value.

determined on that set at each epoch and if it increases continuously for over 500 epochs, the training is stopped and the parameters for the lowest loss so far are used. Again, L2 regularization is applied on every layer with $\lambda = 0.001$.

E. Convolutional neural network

On the other hand, CNN takes an image of the relaxed configuration as the input, which is later processed by the filters in the convolutional layers. The image is prepared from the field J defined as a sum of Gaussian functions centered at the dislocations [25],

$$J(\mathbf{r}) = \sum_{i=1}^{N} \frac{s_i}{\sqrt{2\pi}} \exp\left(-\frac{|\mathbf{r} - \mathbf{r}_i|^2}{2}\right). \qquad (7)$$

The architecture of the CNN implemented in this work within the Keras framework as the sequential model is shown in Fig. 4. First the input is passed in a form of an array of the size 128×128, corresponding to the chosen resolution of the J field. Then it is converted using a convolution layer with 16 filters of the size 3×3 and stride equal to 1. In order to take into account the periodicity of the system, periodic padding is chosen. Since that padding is not one of the available types of padding in Keras, in this work it is implemented manually. A function which extends the array periodically in all directions is created and then a separate Keras layer that executes that function is defined and added before each convolutional layer in the CNN. In that way the output of the convolutional layer has the same size as the input. The convoluted array is subsequently passed into a maximum pooling layer with a kernel 2×2 and stride 2, which reduces the size of the array by 2. There are in total seven sequences of convolutional and pooling layers, after which the size of the array becomes $1 \times 1 \times 16$. Finally, after the flattening procedure a fully connected layer with the linear activation function is inserted to obtain a single value as the output. The detailed structure of the CNN with all the layers listed is shown in Table II. As with the previous methods, the CNN training is also performed with L2 regularization with $\lambda = 0.001$.

III. RESULTS

A. Dataset generation: Strain-controlled DDD simulations

In order to generate the training data for the machine learning models, we fix the initial number of dislocations to be 900 within a square simulation box of linear size of $L = 150$ simulation units and consider a wide range of driving parameter values, i.e., $k = 0.1$, 1, and 10, and $\dot{\epsilon}_a$ varying from 10^{-5} to

TABLE II. Detailed structure of the CNN with all the layers listed together with the shape of their output and the number of the parameters that are updated during the training.

Layer	Shape of the output	Number of trainable parameters
periodic_padding_layer_1	$130 \times 130 \times 1$	0
conv2d_1	$128 \times 128 \times 16$	160
max_pooling_1	$64 \times 64 \times 16$	0
periodic_padding_layer_2	$66 \times 66 \times 16$	0
conv2d_2	$64 \times 64 \times 16$	2320
max_pooling_2	$32 \times 32 \times 16$	0
periodic_padding_layer_3	$34 \times 34 \times 16$	0
conv2d_3	$32 \times 32 \times 16$	2320
max_pooling_3	$16 \times 16 \times 16$	0
periodic_padding_layer_4	$18 \times 18 \times 16$	0
conv2d_4	$16 \times 16 \times 16$	2320
max_pooling_4	$8 \times 8 \times 16$	0
periodic_padding_layer_5	$10 \times 10 \times 16$	0
conv2d_5	$8 \times 8 \times 16$	2320
max_pooling_5	$4 \times 4 \times 16$	0
periodic_padding_layer_6	$6 \times 6 \times 16$	0
conv2d_6	$4 \times 4 \times 16$	2320
max_pooling_6	$2 \times 2 \times 16$	0
periodic_padding_layer_7	$4 \times 4 \times 16$	0
conv2d_7	$2 \times 2 \times 16$	2320
max_pooling_7	$1 \times 1 \times 16$	0
flatten_1	16	0
dense_1	1	17

10^{-2}. For all combinations of values of k and $\dot{\epsilon}_a$ we study deformation predictability for three values of strain ϵ, i.e., $\epsilon = 0.001, 0.01$, and 0.1. A dataset consisting of 10 000 relaxed initial dislocation configurations was created by initializing the systems with 900 randomly positioned dislocations with an equal number of dislocations with positive and negative Burgers vectors, which were then allowed to relax in zero external stress to a metastable state. During the initial relaxation and subsequent loading stage a significant fraction of the dislocations gets annihilated such that after deformation a typical sample contains roughly half of the initial number of dislocations.

Stress-strain curves were then generated from all of these initial configurations for each combination of the driving parameters $\dot{\epsilon}_a$ and k by simulating the DDD model. See Fig. 1 for an example of a dislocation configuration, and ensemble-averaged as well as individual stress-strain curves corresponding to different initial states and strain rates. One may note that each sample has its own unique fluctuating stress-strain curve, consisting of a sequence of stress drops (due to the instantaneous strain rate $\dot{\epsilon}$ exceeding the externally applied strain rate $\dot{\epsilon}_a$) separated by segments where σ_{ext} increases which happens when $\dot{\epsilon} < \dot{\epsilon}_a$ [see Fig. 1(b)]. At the beginning of the deformation the dislocations are still very close to their local minima in the relaxed configuration, therefore the harmonic approximation is valid, which results in the linear dependence of σ_{ext} on ϵ. The system is then deformed elastically and returns to the initial state if the stress is removed. For slightly higher strains the system enters the plastic regime, which is characterized by the collective dislocation dynamics. A key feature of that regime is that for small strains the average of the applied stress σ_{ext} exhibits a power-law dependence on ϵ [Fig. 1(c)]; Supplemental Fig. S1 [34] shows that the related power-law exponent depends on k and $\dot{\epsilon}_a$. Notice also the clear rate effect in the stress-strain curves (Fig. 1): Higher $\dot{\epsilon}_a$ implies a larger average stress σ_{ext} at a given strain ϵ [32]. The resulting dataset, consisting of the relaxed initial dislocation configurations and the corresponding stress-strain curves for different values of $\dot{\epsilon}_a$ and k, was then used to train the ML algorithms to infer the link between the features of the initial relaxed dislocation configurations and the stress-strain curves for all combinations of the $\dot{\epsilon}_a$ and k values considered.

B. Machine learning rate-dependent dislocation dynamics

Predicting the shape of the stress-strain curve by means of ML algorithms with and without user-defined features allows us to assess how much the results depend on the choice of a specific ML model and if it is possible to improve the predictions by extending the feature set beyond the obvious one motivated by physical intuition. As described in Sec. II B, in this work those features were extracted from three fields defined on a configuration. They should in principle provide a full description of the system, however, in practice we consider mainly large-scale statistical quantities derived from these fields, leaving some room for improvement if local details and correlations turn out to be important. The coefficients of the linear correlation between the selected features of the fields and the corresponding values of the external stress σ_{ext}

at a given strain ϵ are defined as

$$r^2 = 1 - \frac{\sum_i \left[\sigma_{\text{ext}}(i) - \sigma_{\text{ext}}^{fit}(i) \right]^2}{\sum_i [\sigma_{\text{ext}}(i) - \langle \sigma_{\text{ext}}(i) \rangle]^2}, \tag{8}$$

where the sums are performed over all the configurations and $\sigma_{\text{ext}}^{fit}(i)$ is the value of the fitted linear function of the given feature. They are shown in Fig. 5 for the GND density field and in Supplemental Figs. S2 and S3 [34] for the other two fields. Depending on the driving parameters ($\dot{\epsilon}_a$ and k) and the value of ϵ considered, different features were found to be important. Generally, at small ϵ and high $\dot{\epsilon}_a$ the average of the absolute value of the GND density, $\langle |\rho_{\text{GND}}| \rangle$, correlates strongly with σ_{ext}, while at large ϵ and low $\dot{\epsilon}_a$ the first Fourier coefficient of ρ_{GND} in the y direction (i.e., $f_{y1}(\rho_{\text{GND}})$, see also Ref. [24]) is more important. As seen in the insets, both of those features correlate negatively with σ_{ext}, however, for other features, for instance the skewness of $|\rho_{\text{GND}}|$, the correlation can be positive.

In contrast to linear regression and NN, CNN does not require any manually defined features. Instead, a pixelated image of the dislocation configuration is passed as the input. In principle, all the relevant features to the given problem can be extracted from those images by the CNN without user input, and hence CNN has the potential to do better at predicting the deformation dynamics than the other two models using user-defined features as input. Here images of the resolution of 128×128 were used as input for the CNN (see an example input image in Fig. 4), chosen in order to be able to resolve individual dislocations while at the same time minimizing the size of the image.

The ML models were trained with five different random seeds, which determine how the dataset is split into the training, test, and validation set at their fixed ratio of 80:10:10% (80:20:0% for linear regression). Our main result, i.e., the predictability scores defined analogously to the linear correlation coefficient, Eq. (8), determined for the test set of all the ML models and averaged over all the seeds are plotted in Fig. 6. Scores for all sets (training, test, and validation) are shown in Supplemental Fig. S4 for linear regression and in Supplemental Fig. S5 [34] for NN. For CNN, we consider also the dependence of the results on the dataset size: The CNN scores for different sizes of the dataset are shown in Supplemental Figs. S6–S11, with the average gap between the training and test set as a function of the dataset size shown in Supplemental Fig. S12 [34]. These results show that the training of linear regression and NN have converged well for the size of the dataset at hand (10 000 samples) in that the scores for the training and test sets are almost equal (no overfitting). On the other hand, CNN could use more data to fully close the gap between the training and test set: Linear extrapolation in Supplemental Fig. S12 [34] suggests that a dataset size of approximately 17 000 would be needed for that.

The scores for linear regression and NN are always very close to each other, even though the latter method allows for nonlinear relations between the input and output data, suggesting that their performance reflects the set of input features provided rather than differences in the ML models. On the other hand, even if its training did not fully converge using the available dataset, CNN almost always outperforms
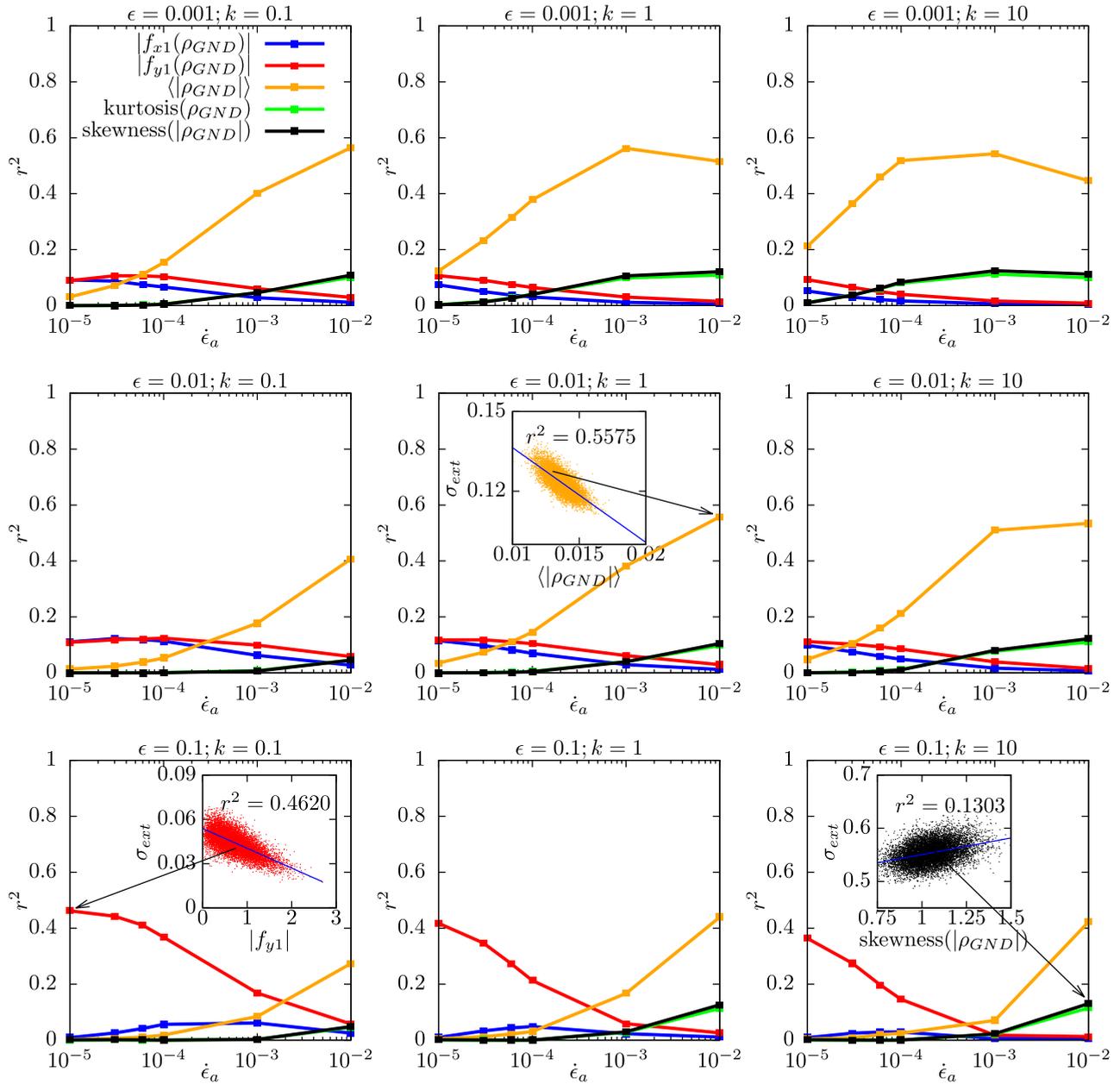
FIG. 5. Correlation of $\rho_{GND}(x, y)$ with the stress-strain response. Coefficients of linear correlation r^2 between various properties of the geometrically necessary dislocations field $\rho_{GND}(x, y)$ and the values of the external stress σ_{ext} obtained at given strain ϵ, stiffness k, and applied strain rate $\dot{\epsilon}$. Only the features exhibiting significant correlation are shown. The insets show example fits of the chosen properties at the points indicated by the arrow. Corresponding figures showing the r^2's for the dislocation density field $\rho(x, y)$ and internal stress field σ_{int} are included as Supplemental Figs. S2 and S3 [34], respectively.

the other two methods, indicating that it is able to extract additional relevant features not included in the manually engineered features used for linear regression and NN. Hence, in what follows, we take the CNN score to be the best available estimate of how well the stress-strain curves can be predicted using information of the initial dislocation configurations as input. Notably, in most cases the resulting deformation predictability increases with the strain rate, especially at lower strains. However, at the highest strain studied, i.e., for $\epsilon = 0.1$, this relation becomes nonmonotonic. Especially for the largest k considered, the predictability score r^2 starts at a relatively

high value followed by a decrease, reaching the minimum at an intermediate strain rate, and finally r^2 increases again for the highest strain rate considered. It is worth noting that for small strains and high strain rates, deformation predictability is almost perfect (r^2 close to 1), and the large-rate r^2 exceeds 0.8 even for the largest ϵ of 0.1 largely independently of the ML model. Decreasing $\dot{\epsilon}_a$ results in general in a significant reduction of r^2 and CNN outperforming the other predictive models. These observations are generally consistent with the values of linear correlation coefficients for the selected features (see Fig. 5 and Supplemental Figs. S2 and S3) [34]:
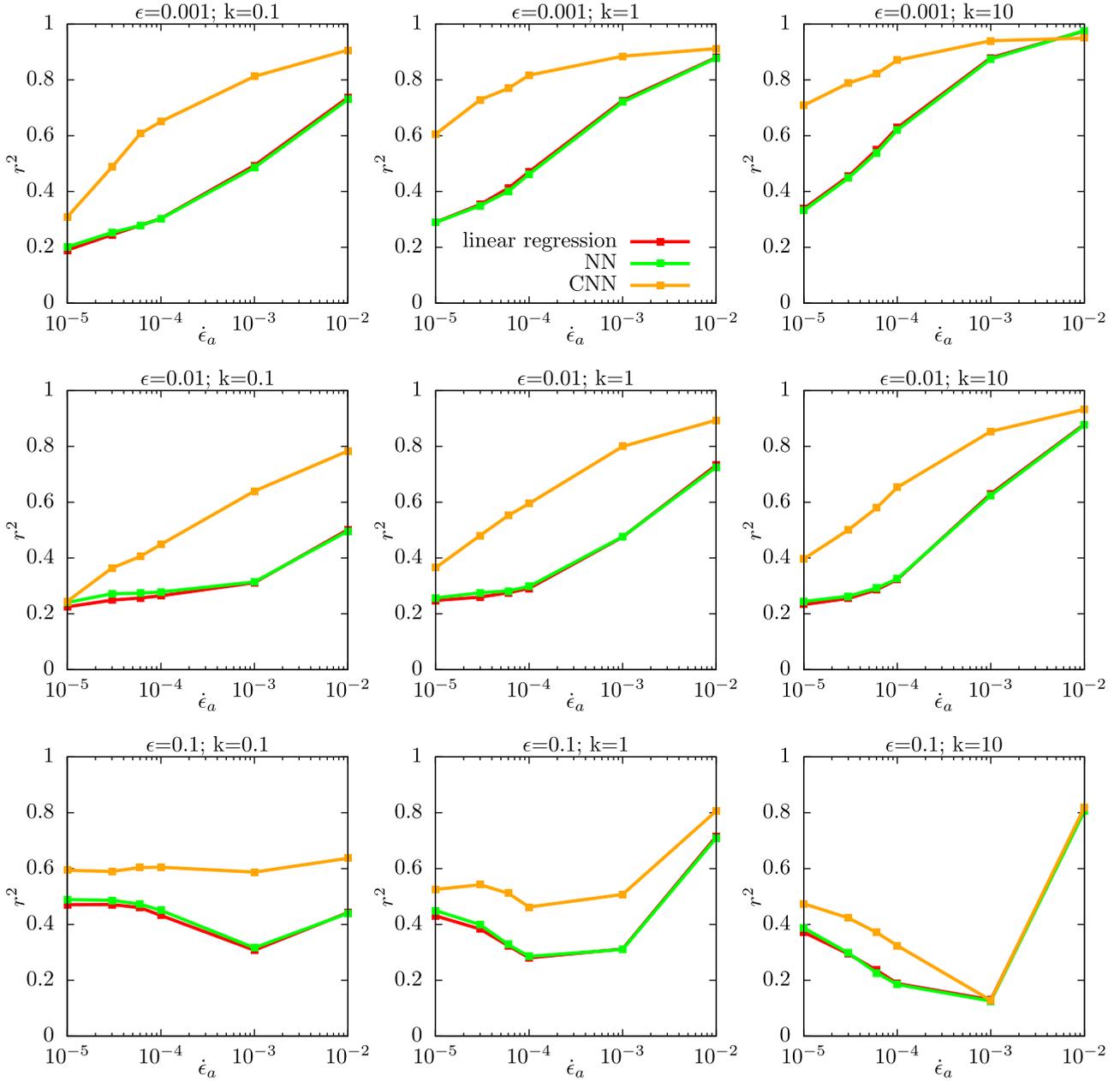
FIG. 6. Rate-dependent deformation predictability revealed by machine learning. Comparison of test set predictability scores averaged over five random seeds for the three different ML methods considered. Notice how predictability improves with increasing $\dot{\epsilon}_a$ for small ϵ and exhibits nonmonotonic dependence on $\dot{\epsilon}_a$ for large ϵ. Note also how linear regression and NN results are almost indistinguishable, while CNN generally outperforms the other two predictive models.

For most of those features the correlation increases with the strain rate. The ML models, combining information from all of these features into a single prediction, are in most cases able to do significantly better than just considering the correlation of $\sigma_{\text{ext}}(\epsilon)$ with the individual features.

C. Rate-dependent complexity of dislocation dynamics controls deformation predictability for small strains

The evolution of deformation predictability with $\dot{\epsilon}_a$ should be related to changes in the fundamental nature of dislocation dynamics as the strain rate is changed. One can expect that at higher strain rates the dynamics becomes simpler—and

hence easier to predict—because it is dominated by the strong external stress, due to the rate effect discussed above implying higher average σ_{ext} at a given ϵ as $\dot{\epsilon}_a$ is increased. Therefore the internal stresses σ_{int} due to the other dislocations are in relative terms less important compared to σ_{ext} for high $\dot{\epsilon}_a$. This implies that for high $\dot{\epsilon}_a$ most dislocations move in the direction set by σ_{ext} largely independently of the other dislocations. On the other hand, for low $\dot{\epsilon}_a$, σ_{ext} is typically smaller and dislocation-dislocation interaction stresses are in relative terms more important, implying complex (and hence likely harder-to-predict) dislocation dynamics where large multipolar dislocation structures slowly drift in the direction set by their net Burgers vector [33]. Hence, for low $\dot{\epsilon}_a$, a significant
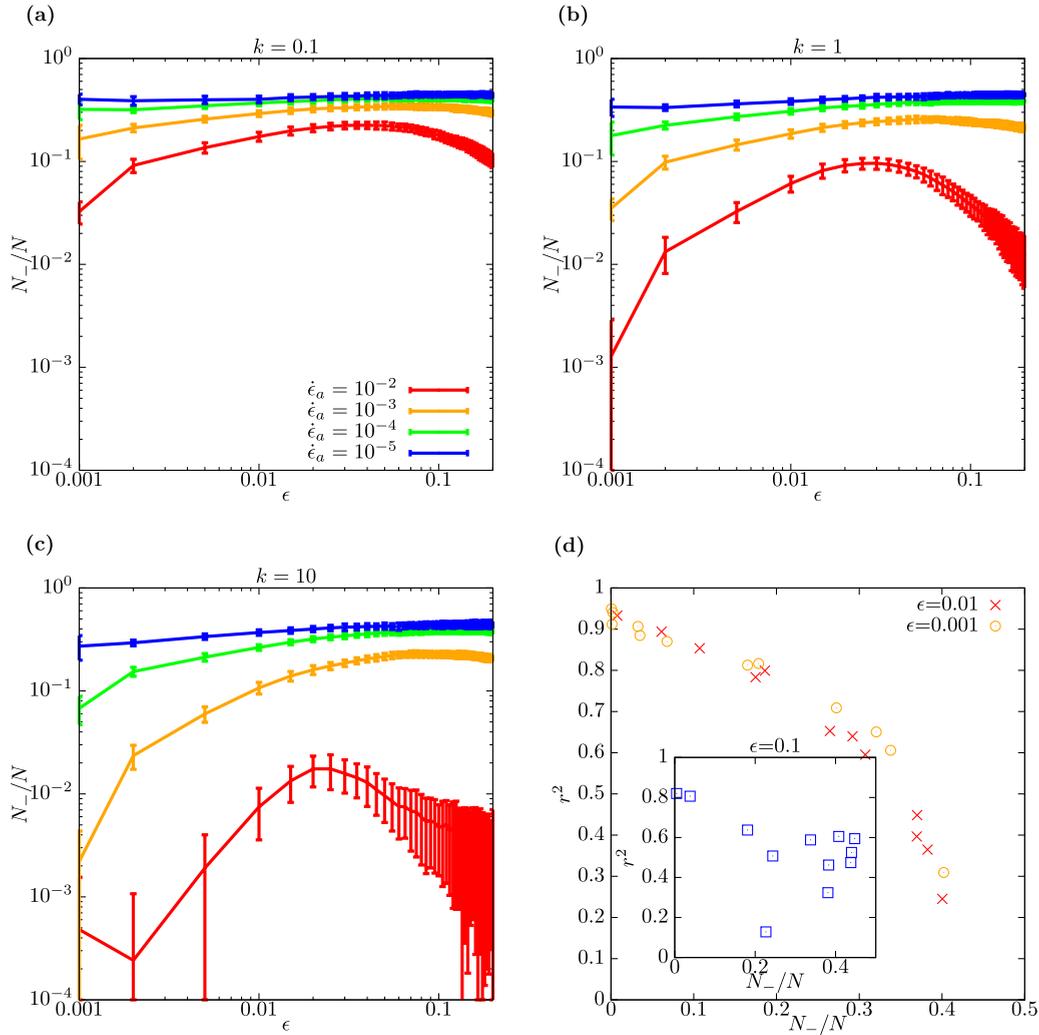
FIG. 7. Rate-dependent complexity of dislocation dynamics controls deformation predictability. Fraction of dislocations moving in the direction opposite to the one set by the applied external stress, N_-/N, as a function of ϵ for (a) $k = 0.1$, (b) $k = 1$, and (c) $k = 10$, averaged over 100 configurations. (d) shows the correlation between the r^2 of the CNN test set and N_-/N at the values of ϵ equal to 0.001, 0.01, and additionally 0.1 in the inset. Error bars are standard error of the mean (SEM).

fraction of the dislocations, being part of such collectively moving complex multipolar dislocation structures, is expected to move against the direction set by σ_{ext}, with strain accumulating due to only a small majority of the dislocations moving in the direction set by external loading. To quantify the rate dependence of this effect, we consider in Figs. 7(a)–7(c) the ratio of the dislocations moving against the external stress N_- and the total number of dislocations N as a function of $\dot{\epsilon}_a$ for different values of k. All the curves start from zero at $\epsilon = 0$ (linear response of dislocations in a metastable configuration to a small σ_{ext}) and plateau at a certain value that is higher for lower $\dot{\epsilon}_a$, implying gradual buildup of dislocation structures of increasing size moving collectively in a direction set by their net Burgers vector. The exception is the curve for the highest strain rate $\dot{\epsilon}_a = 0.01$ considered, which after reaching a maximum starts to drop significantly, implying that the collectively moving dislocation structures formed during the early stages of loading break up as σ_{ext} increases with increasing ϵ. The effect becomes stronger for larger values of k, for which the system is stiffer and the strain rate is more

accurately controlled by larger fluctuations of σ_{ext} as indicated by Eq. (2). As a consequence, for higher k fewer dislocations move against the external stress, especially for larger values of $\dot{\epsilon}_a$.

Remarkably, Fig. 7(d) shows that the CNN predictability score r^2 exhibits a very clear negative correlation with N_-/N for low strains. Notice how the data points corresponding to different values of $\dot{\epsilon}_a$ and k fall on the same curve independently of whether $\epsilon = 0.001$ or 0.01 is considered. This supports the idea that dislocation dynamics becomes more complex and hence harder to predict when smaller $\dot{\epsilon}_a$ and k values are considered. This correlation is weaker for the largest strain value considered, i.e., $\epsilon = 0.1$ [inset of Fig. 7(d)], and hence we need to consider other measures of complexity of the dislocation dynamics in that case.

D. Nonmonotonic predictability for large strains

In order to account for the nonmonotonic dependence of r^2 on $\dot{\epsilon}_a$ for large ϵ, we start by considering the role of different
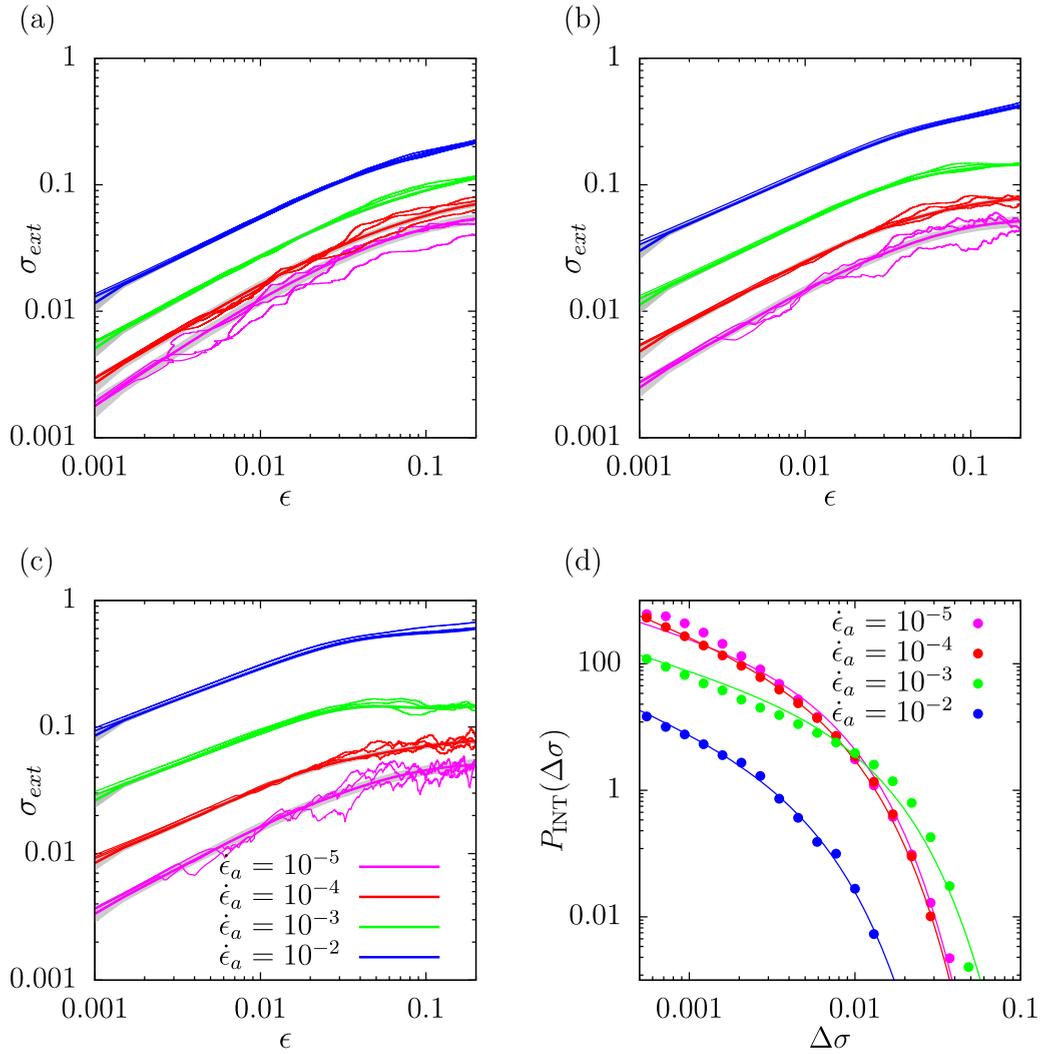
FIG. 8. Stress-strain curves and distribution of stress drops as a function of their magnitude. (a) $k = 0.1$, (b) $k = 1$, and (c) $k = 10$. (d) shows the integrated over the whole stress-strain curves of all the samples distribution of stress drops during the loading as a function of their magnitude for $k = 10$. The curves represent the power law distribution with an exponential cutoff as expressed by Eq. (9) fitted to the data.

features to understand the origin of the nonmonotonic evolution of deformation predictability with $\dot{\epsilon}_a$ in that case. As seen in Fig. 5, the first Fourier coefficient of ρ_{GND} in the y direction, $f_{y1}(\rho_{\mathrm{GND}})$, exhibits for $\epsilon = 0.1$ a large r^2 for small $\dot{\epsilon}_a$, which then decreases towards zero with increasing $\dot{\epsilon}_a$. $f_{y1}(\rho_{\mathrm{GND}})$ is conserved during the dynamics, and hence it is natural that it becomes relevant for large ϵ [24]. At the same time, for many other features [notably for $\langle |\rho_{\mathrm{GND}}| \rangle$ (see Fig. 5), $\langle \rho \rangle$ (Supplemental Fig. S2), and several features of the σ_{int} field (Supplemental Fig. S3)] [34], r^2 increases with $\dot{\epsilon}_a$, such that they become important when $f_{y1}(\rho_{\mathrm{GND}})$ ceases to be relevant for large $\dot{\epsilon}_a$, naturally resulting in nonmonotonic dependence of the overall r^2 on $\dot{\epsilon}_a$ for the ML models using these features as input.

In order to understand the nonmonotonic dependence of r^2 on $\dot{\epsilon}_a$ at high ϵ more qualitatively one can analyze the evolution of the corresponding stress-strain curves, keeping in mind that individual critical-like dislocation avalanches are typically understood to be inherently hard-to-predict [25]. In Figs. 8(a)–8(c) example stress-strain curves are shown for

$k = 0.1, 1,$ and 10, respectively, for the four largest strain rates studied. Generally, the curves become more regular for larger values of $\dot{\epsilon}_a$. Those for $\dot{\epsilon}_a = 10^{-2}$ contain no significant stress drops, therefore, they are the easiest to predict by the ML algorithm since the stress increases monotonically with the strain without any hard-to-predict stress drops. On the other hand, the stress-strain curves for the lowest strain rates tend to be very irregular and exhibit a large number of stress drops. In such cases predicting the stress value requires estimating the number and magnitude of those events occurring prior to the given strain value. However, it is easier to perform when those events are more frequent because their occurrences are then closer to some averaged distribution described by a certain function. It is also expected that recent observations of correlations between subsequent avalanche events driving the individual curves towards the mean ones [35] might play a role here. In Fig. 8(d) the distributions of stress drops as a function of their magnitude are shown for different $\dot{\epsilon}_a$ in the case of $k = 10$, i.e., the k value for which the nonmonotonic behavior of r^2 is most pronounced. The fitted curves (shown as

lines) represent the power law distribution with an exponential cutoff [32],

$$P_{\text{INT}}(\Delta\sigma) = a(\Delta\sigma)^{-\tau_{\sigma,\text{INT}}} \exp(-\Delta\sigma/\Delta\sigma_0), \qquad (9)$$

where a, $\tau_{\sigma,\text{INT}}$, and $\Delta\sigma_0$ are fitting parameters. The largest drops occur for $\dot{\epsilon}_a = 10^{-3}$, for which r^2 has the minimum. However, there are also fewer small drops than for $\dot{\epsilon}_a = 10^{-4}$. Generally, one can expect that strain drops that occur rarely, that is, much less often than once per configuration, do not have a significant detrimental effect on the predictability. On the other hand, the number of events that occur very frequently can also be predicted relatively easily because the estimation error becomes smaller. The hardest to predict are stress-strain curves with events that occur on average once per stress-strain curve: For accurate estimation of the stress at a given strain which is in that case largely controlled by the occurrence of an individual stress drop, the algorithm must predict whether the event will occur before or after a certain value of strain and what the magnitude of that event will be. Since for $\dot{\epsilon}_a = 10^{-3}$ the stress drops are rare, and as seen in Fig. 8(d) their distribution has a larger cutoff than for the other strain rates, the corresponding stress-strain curves are also the most difficult to predict. Interestingly, the minimum of r^2 (or the maximum of the prediction error) thus appears to coincide with a transition from fluctuating to smooth plastic flow upon increasing $\dot{\epsilon}_a$.

IV. DISCUSSION AND CONCLUSIONS

To conclude, we have established that predictability of strain-controlled deformation of plastically deforming single crystals is rate dependent. While the predictability scores of CNN, using images of the configurations as input without explicit feature engineering by the user, are generally higher than those of linear regression and NN using manually engineered input features, all the algorithms exhibit very similar trends of r^2 as a function of $\dot{\epsilon}_a$, ϵ, and k: Better deformation predictability (i.e., lower prediction error of the ML models) with increasing $\dot{\epsilon}_a$ at low ϵ and nonmonotonic relation between predictability and $\dot{\epsilon}_a$ for high ϵ, especially for large driving spring stiffness k. These results suggest that deformation predictability as measured by the ML models is related at least partially to some intrinsic aspects of the dislocation system and the deformation process studied. Indeed, we showed that at lower strains a linear correlation between the predictability and the fraction of dislocations moving against the external stress, which we chose as a measure of the complexity of the deformation process, exists. Since for high $\dot{\epsilon}_a$ σ_{ext} is also high, most dislocations move in the direction of the applied stress, and the dislocation dynamics is "simple," and as our results suggest, easier to predict. On the other hand, for lower strain rates the dislocations are more free to follow the system's internal dynamics, governed mostly by the interactions between the dislocations. Since $f_{y1}(\rho_{\text{GND}})$ does not change during the dynamics and hence encodes some constant property of the large-scale distribution of the dislocations, it may be relevant for describing those interactions. Indeed, we found that $f_{y1}(\rho_{\text{GND}})$ is strongly correlated with $\sigma_{\text{ext}}(\epsilon)$ for large ϵ and small $\dot{\epsilon}_a$, something that together with the correlations exhibited by the other features (which generally increase with

increasing $\dot{\epsilon}_a$) allows us to understand the nonmonotonic dependence of deformation predictability on $\dot{\epsilon}_a$ for large ϵ. On the other hand, this nonmonotonic dependence of r^2 on $\dot{\epsilon}_a$ for large ϵ was also argued to be linked to a transition from fluctuating to smooth plastic flow, such that at the transition point the large-strain part of the stress-strain curve is largely controlled by a single hard-to-predict stress drop. In addition to deformation predictability, the degree of complexity of the deformation process evolving with the strain rate outlined above could have more general implications in understanding rate effects in plasticity.

We emphasize that even if the DDD model we have studied obeys deterministic equations of motion, the complex and possibly chaotic nature of the dislocation dynamics implies that perfect predictability (i.e., $r^2 = 1$) is not reachable in practice, i.e., the prediction error of any real ML model remains finite. One could of course in theory envisage a hypothetical ML model that would be exactly equivalent to a numerical integration algorithm of the equations of motion of the dislocations and would use the initial dislocation positions given with infinite precision as input, and obviously such a model would result in $r^2 = 1$. However, as soon as there is any deviation from perfect accuracy, e.g., in the initial positions of the dislocations, or if the model does not know the equations of motion perfectly, the perfect predictability is expected to break down. When considering the prospect of studying deformation predictability experimentally, e.g., by extracting descriptors of the initial dislocation microstructure by means of x-ray measurement techniques [36–38], the interesting question is to what extent a characterization of the initial state with a finite accuracy is suitable for making predictions of the subsequent deformation dynamics. Our results suggest that the answer to this question depends on the strain rate, and it would be interesting to verify this prediction experimentally.

These results could be generalized to three-dimensional systems with flexible, curved dislocation lines with arbitrary Burgers vectors [39]. While the dynamics would be more complex and dislocations would have more degrees of freedom, one can expect that the same effect of predictability increasing with strain rate could be observed. That is because (as in the 2D system considered here) the complexity of the dislocation dynamics would be expected to decrease at higher strain rates such that the dislocations are increasingly forced by the external stress to move in a given direction. On the other hand, it would be interesting to study what happens at low strain rates in such systems and if the nonmonotonic relation between strain rate and predictability at high strain still exists. While $f_{y1}(\rho_{\text{GND}})$ would no longer be conserved during the dynamics in a multislip system, there may exist other features that would control the dynamics at low strain rates and large strains. In particular, one could consider dynamics in the presence of pinning centers [40], whose distribution would be an intrinsic feature of the system, which may increase the predictability at larger strains [25]. The above-described extension to 3D DDD would of course be computationally quite demanding, due to increased computational cost of generating the training data set as well as increased complexity of the required ML models. As dislocations would be able to form more complex structures, in the ML models one would need to

incorporate additional features related to existence of dislocation junctions, locks, etc. Analogously to the two-dimensional model, one could prepare a three-dimensional voxelized representation of the dislocation configuration to train a CNN model. Another natural extension of the study would be to consider polycrystals containing several crystal grains. This would imply incorporating additional fundamental features to the ML models such as crystal structure, presence of grain boundaries, and point defects, etc. Finally, it would also be of interest to consider possible rate effects in the predictability of deformation of amorphous solids [41,42].

[1] M. D. Uchic, P. A. Shade, and D. M. Dimiduk, Plasticity of micrometer-scale single crystals in compression, Annu. Rev. Mater. Res. **39**, 361 (2009).

[2] D. Dimiduk, M. Uchic, and T. Parthasarathy, Size-affected single-slip behavior of pure nickel microcrystals, Acta Mater. **53**, 4065 (2005).

[3] H. Fan, Q. Wang, J. A. El-Awady, D. Raabe, and M. Zaiser, Strain rate dependency of dislocation plasticity, Nat. Commun. **12**, 1845 (2021).

[4] R. Schwaiger, B. Moser, M. Dao, N. Chollacoop, and S. Suresh, Some critical experiments on the strain-rate sensitivity of nanocrystalline nickel, Acta Mater. **51**, 5159 (2003).

[5] D. M. Dimiduk, C. Woodward, R. LeSar, and M. D. Uchic, Scale-free intermittent flow in crystal plasticity, Science **312**, 1188 (2006).

[6] P. D. Ispánovity, L. Laurson, M. Zaiser, I. Groma, S. Zapperi, and M. J. Alava, Avalanches in 2D Dislocation Systems: Plastic Yielding is Not Depinning, Phys. Rev. Lett. **112**, 235501 (2014).

[7] M. Ovaska, L. Laurson, and M. J. Alava, Quenched pinning and collective dislocation dynamics, Sci. Rep. **5**, 10580 (2015).

[8] F. F. Csikor, C. Motz, D. Weygand, M. Zaiser, and S. Zapperi, Dislocation avalanches, strain bursts, and the problem of plastic forming at the micrometer scale, Science **318**, 251 (2007).

[9] S. Papanikolaou, M. Tzimas, A. C. E. Reid, and S. A. Langer, Spatial strain correlations, machine learning, and deformation history in crystal plasticity, Phys. Rev. E **99**, 053003 (2019).

[10] M. Raissi and G. E. Karniadakis, Hidden physics models: Machine learning of nonlinear partial differential equations, J. Comput. Phys. **357**, 125 (2018).

[11] P. Baldi, P. Sadowski, and D. Whiteson, Searching for exotic particles in high-energy physics with deep learning, Nat. Commun. **5**, 4308 (2014).

[12] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Machine learning for molecular and materials science, Nature (London) **559**, 547 (2018).

[13] G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, Accelerating materials property predictions using machine learning, Sci. Rep. **3**, 2810 (2013).

[14] L. Viitanen, J. R. Mac Intyre, J. Koivisto, A. Puisto, and M. Alava, Machine learning and predicting the time-dependent dynamics of local yielding in dry foams, Phys. Rev. Res. **2**, 023338 (2020).

[15] S. Biswas, D. F. Castellanos, and M. Zaiser, Prediction of creep failure time using machine learning, Sci. Rep. **10**, 16910 (2020).

[16] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, Nat. Rev. Phys. **3**, 422 (2021).

[17] H. Salmenjoki, M. J. Alava, and L. Laurson, Mimicking complex dislocation dynamics by interaction networks, Eur. Phys. J. B **91**, 275 (2018).

[18] H. Salmenjoki, L. Laurson, and M. J. Alava, Probing the transition from dislocation jamming to pinning by machine learning, Mater. Theory **4**, 5 (2020).

[19] F. E. Bock, R. C. Aydin, C. J. Cyron, N. Huber, S. R. Kalidindi, and B. Klusemann, A review of the application of machine learning and data mining approaches in continuum materials mechanics, Front. Mater. **6**, 110 (2019).

[20] R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, Machine learning in materials informatics: Recent applications and prospects, npj Comput. Mater. **3**, 54 (2017).

[21] Y. Liu, T. Zhao, W. Ju, and S. Shi, Materials discovery and design using machine learning, J. Materiomics **3**, 159 (2017).

[22] K. Guo, Z. Yang, C.-H. Yu, and M. J. Buehler, Artificial intelligence and machine learning in design of mechanical materials, Mater. Horiz. **8**, 1153 (2021).

[23] Z. Yu, S. Ye, Y. Sun, H. Zhao, and X.-Q. Feng, Deep learning method for predicting the mechanical properties of aluminum alloys with small data sets, Mater. Today Commun. **28**, 102570 (2021).

[24] H. Salmenjoki, M. J. Alava, and L. Laurson, Machine learning plastic deformation of crystals, Nat. Commun. **9**, 5307 (2018).

[25] M. Sarvilahti, A. Skaugen, and L. Laurson, Machine learning depinning of dislocation pileups, APL Mater. **8**, 101109 (2020).

[26] B. Rouet-Leduc, C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson, Machine learning predicts laboratory earthquakes, Geophys. Res. Lett. **44**, 9276 (2017).

[27] Y. Y. Kagan and D. D. Jackson, Probabilistic forecasting of earthquakes, Geophys. J. Int. **143**, 438 (2000).

[28] P. M. Anderson, J. P. Hirth, and J. Lothe, *Theory of Dislocations* (Cambridge University Press, Cambridge, 2017).

[29] B. Voight, A method for prediction of volcanic eruptions, Nature (London) **332**, 125 (1988).

[30] N. Israeli and N. Goldenfeld, Coarse-graining of cellular automata, emergence, and the predictability of complex systems, Phys. Rev. E **73**, 026203 (2006).

[31] V. Deshpande, A. Needleman, and E. Van der Giessen, Dislocation dynamics is chaotic, Scr. Mater. **45**, 1047 (2001).

[32] D. Kurunczi-Papp and L. Laurson, Dislocation avalanches from strain-controlled loading: A discrete dislocation dynamics study, Phys. Rev. E **104**, 025008 (2021).

[33] M. Carmen Miguel, A. Vespignani, M. Zaiser, and S. Zapperi, Dislocation Jamming and Andrade Creep, Phys. Rev. Lett. **89**, 165501 (2002).

[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevMaterials.6.023602 for additional figures supporting the results of the main paper.

[35] H. Salmenjoki, L. Laurson, and M. J. Alava, Avalanche correlations and stress-strain curves in discrete dislocation plasticity, Phys. Rev. Mater. **5**, 073601 (2021).

[36] W. Ludwig, P. Cloetens, J. Härtwig, J. Baruchel, B. Hamelin, and P. Bastie, Three-dimensional imaging of crystal defects by 'topo-tomography', J. Appl. Crystallogr. **34**, 602 (2001).

[37] L. E. Levine, B. C. Larson, W. Yang, M. E. Kassner, J. Z. Tischler, M. A. Delos-Reyes, R. J. Fields, and W. Liu, X-ray microbeam measurements of individual dislocation cell elastic strains in deformed single-crystal copper, Nat. Mater. **5**, 619 (2006).

[38] I. Groma, D. Tüzes, and P. Ispánovity, Asymmetric x-ray line broadening caused by dislocation polarization induced by external load, Scr. Mater. **68**, 755 (2013).

[39] A. Lehtinen, G. Costantini, M. J. Alava, S. Zapperi, and L. Laurson, Glassy features of crystal plasticity, Phys. Rev. B **94**, 064101 (2016).

[40] H. Salmenjoki, A. Lehtinen, L. Laurson, and M. J. Alava, Plastic yielding and deformation bursts in the presence of disorder from coherent precipitates, Phys. Rev. Mater. **4**, 083602 (2020).

[41] S. Karmakar, A. Lemaître, E. Lerner, and I. Procaccia, Predicting Plastic Flow Events in Athermal Shear-Strained Amorphous Solids, Phys. Rev. Lett. **104**, 215502 (2010).

[42] D. Richard, M. Ozawa, S. Patinet, E. Stanifer, B. Shang, S. A. Ridout, B. Xu, G. Zhang, P. K. Morse, J.-L. Barrat, L. Berthier, M. L. Falk, P. Guan, A. J. Liu, K. Martens, S. Sastry, D. Vandembroucq, E. Lerner, and M. L. Manning, Predicting plasticity in disordered solids from structural indicators, Phys. Rev. Mater. **4**, 113609 (2020).