

# Machine learning depinning of dislocation pileups

Cite as: APL Mater. **8**, 101109 (2020); <https://doi.org/10.1063/5.0020376>

Submitted: 30 June 2020 . Accepted: 02 October 2020 . Published Online: 20 October 2020

Mika Sarvilahti,  Audun Skaugen, and  Lasse Laurson



View Online



Export Citation



CrossMark

## ARTICLES YOU MAY BE INTERESTED IN

[Electrically driven transient and permanent phase transformations in highly strained epitaxial BiFeO<sub>3</sub> thin films](#)

APL Materials **8**, 101110 (2020); <https://doi.org/10.1063/5.0025673>

[In situ x-ray and electron scattering studies of oxide molecular beam epitaxial growth](#)

APL Materials **8**, 101107 (2020); <https://doi.org/10.1063/5.0025849>

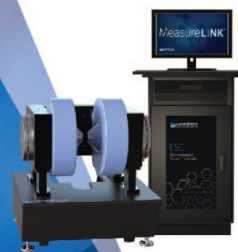
[Soft peripheral nerve interface made from carbon nanotubes embedded in silicone](#)

APL Materials **8**, 101111 (2020); <https://doi.org/10.1063/5.0021887>

 **Measure Ready**  
MCS-EMP Modular Characterization Systems

**NEW**

Multi-purpose platforms for  
automated variable-field experiments



 **Lake Shore**  
CRYOTRONICS

Find out more

**AIP**  
Publishing

# Machine learning depinning of dislocation pileups

Cite as: APL Mater. 8, 101109 (2020); doi: 10.1063/5.0020376

Submitted: 30 June 2020 • Accepted: 2 October 2020 •

Published Online: 20 October 2020



View Online



Export Citation



CrossMark

Mika Sarvilahti, Audun Skaugen,  and Lasse Laurson<sup>a)</sup> 

## AFFILIATIONS

Computational Physics Laboratory, Tampere University, P.O. Box 692, FI-33014 Tampere, Finland

**Note:** This paper is part of the Special Topic on Discovering Patterns in Disorder: Machine Learning for Fluctuating Mesoscopic Materials.

<sup>a)</sup> Author to whom correspondence should be addressed: [lasse.laurson@tuni.fi](mailto:lasse.laurson@tuni.fi)

## ABSTRACT

We study a one-dimensional model of a dislocation pileup driven by an external stress and interacting with random quenched disorder, focusing on the predictability of the plastic deformation process. Upon quasistatically ramping up the externally applied stress from zero, the system responds by exhibiting an irregular stress–strain curve consisting of a sequence of strain bursts, i.e., critical-like dislocation avalanches. The strain bursts are power-law distributed up to a cutoff scale that increases with the stress level up to a critical flow stress value. There, the system undergoes a depinning phase transition and the dislocations start moving indefinitely, i.e., the strain burst size diverges. Using sample-specific information about the pinning landscape as well as the initial dislocation configuration as input, we employ predictive models such as linear regression, simple neural networks, and convolutional neural networks to study the predictability of the simulated stress–strain curves of individual samples. Our results show that the response of the system—including the flow stress value—can be predicted quite well, with the correlation coefficient between the predicted and actual stress exhibiting a non-monotonic dependence on strain. We also discuss our attempts to predict the individual strain bursts.

© 2020 Author(s). All article content, except where otherwise noted, is licensed under a Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>). <https://doi.org/10.1063/5.0020376>

## I. INTRODUCTION

One of the key ideas of many disciplines including materials science in particular is that the structure and properties of materials tend to be related.<sup>1,2</sup> For instance, much of metallurgy is about tuning the materials' microstructure to design materials with desired mechanical properties.<sup>2</sup> Typically, the relation between, say, the precipitate content of a precipitation-hardened alloy<sup>3</sup> and its yield strength is understood as an average property, and indeed, in the case of macroscopic bulk samples, the sample-to-sample variations in the yield stress of identically prepared samples tend to be small. However, this changes when dealing with samples with sizes down to the micrometer range and below: Recent micropillar compression experiments have revealed the fluctuating, irregular nature of small-scale crystal plasticity,<sup>4,5</sup> originating from the collective, critical-like dynamics of interacting dislocations mediating the deformation process. The fact that the plasticity of small crystals proceeds via an irregular sequence of strain bursts with a broad size distribution

implies also that sample-to-sample variations of the plastic response or the stress–strain curves might be considerable<sup>6</sup> even if the samples have been prepared and the deformation experiments performed using the same protocol.

This raises the important general question of what can be said about the relation between the initial structure and mechanical properties for individual samples of small crystals, when random variations in the initial (micro)structure become important. One way of framing the issue is to consider what we refer to as *deformation predictability*: Given a small crystalline sample with a specific arrangement of pre-existing dislocations and the possibility of some other defects interfering with dislocation motion, how well can one use such information to predict the plastic response of that sample. A key challenge here is the high dimensionality of any reasonable description of the disordered initial state, combined with the possibility of non-linearities in the mapping from the initial state properties to the ensuing response. This likely implies that simple empirical laws relating sample properties to its plastic response may

be difficult to formulate. Challenges of this type are a key factor behind the recent emergence of machine learning (ML) algorithms as an important part of the toolbox of scientists in a wide range of fields including also physics and materials science.<sup>7,8</sup> ML has been demonstrated to be an efficient approach to address a wide spectrum of problems including materials' property prediction, discovery of novel materials, etc.<sup>9–15</sup> These developments have led to the emergence of a novel research field referred to as *materials informatics*,<sup>16</sup> where informatics methods—including ML—are used to determine material properties that are hard to measure or compute using traditional methods.

In general, many supervised ML algorithms including neural networks are capable of learning non-linear mappings from a high-dimensional feature vector to a desired output, and hence, these methods should be useful tools when addressing the question of deformation predictability. An important opening in this direction was recently achieved by Salmenjoki *et al.*,<sup>17</sup> who applied ML to study the deformation predictability in the case of simple two-dimensional discrete dislocation dynamics (DDD) simulations. The key idea of Ref. 17 was that exploiting ML algorithms provides a useful set of methods to quantify the predictability of complex systems such as plastically deforming crystals. This was achieved by training ML models including neural networks and support vector machines to infer the mapping from the initial dislocation microstructure to the response of the system to applied stresses, characterized by the stress–strain curve. Following Ref. 17, ML has been further applied to predicting stress–strain curves of plastically deforming crystals<sup>18</sup> and learning the interaction kernel of dislocations.<sup>19</sup> Recently, ML has also been applied to the closely related problems of, e.g., predicting the local yielding dynamics of dry foams,<sup>20</sup> the creep failure time of disordered materials,<sup>21</sup> and the occurrence times of “laboratory earthquakes.”<sup>22</sup>

By its nature, the problem of predicting the plastic deformation process of crystalline samples depends on details such as whether the crystal only contains pre-existing glissile dislocations (this was the case in Ref. 17) or if other defect populations interacting with the dislocations are present. The latter could include solute atoms, precipitates, voids, or even grain boundaries in the case of polycrystals. If present, a description of these additional defects needs to be included in the initial state of the system used as input for the predictive ML models. The presence of such static defects within the crystal may also change the response of the crystal to applied stresses and, thus, the process to be predicted in a fundamental manner. It has been shown that the deformation dynamics of “pure” dislocation systems are governed by dislocation jamming,<sup>23,24</sup> resulting in glassy dislocation dynamics characterized by “extended criticality,” with the cutoff scale of the size distribution of dislocation avalanches diverging with the system size at any stress level.<sup>25–27</sup> This can be contrasted with systems where significant pinning of dislocations due to other defects within the crystal may instead induce a depinning phase transition of the dislocation assembly, resulting in critical-like dislocation dynamics only in the immediate proximity of the critical point (stress) of the depinning transition.<sup>28,29</sup>

In this paper, we study deformation predictability in a one-dimensional periodic model of a dislocation pileup, interacting with a quenched (frozen) random pinning landscape. The model is perhaps the simplest possible system including both interacting mobile

dislocations and a quenched random pinning field interfering with dislocation motion, and therefore, it serves as a useful playground to explore the ideas of ML-based deformation predictability discussed above. It is known to exhibit a depinning phase transition at a critical flow stress value  $\sigma_{\text{flow}}$ , separating pinned and moving phases of the dislocation system.<sup>30</sup> While the general topic of applying ML in the context of disordered materials has recently gained some momentum,<sup>17,18,20–22,31</sup> to our knowledge, no study addressing the question of deformation predictability in a dislocation system interacting with quenched disorder and exhibiting a depinning phase transition has been performed before. To this end, we generate a large database of stress–strain curves, each corresponding to a unique randomly generated initial microstructure, by simulating the model with a quasistatically increasing applied stress for different system sizes (or different numbers of dislocations,  $N$ ). First, the statistical properties of the stress–strain curves and the strain bursts are analyzed, followed by training and testing of various predictive models ranging from linear regression to convolutional neural networks (CNNs; considering these different ML models allows us to assess the possible model dependence of deformation predictability) to establish mappings from the initial random microstructure (defined by both the static pinning landscape and the initial configuration of the dislocations) to the response of the system to applied stresses (i.e., the stress–strain curve). Notice that the deformation predictability as defined above is in general dependent on the ML model considered, but by studying the problem using multiple different ML models allows one to use the results from the one with the best predictive power as measures of the lower limit of the more general, ML model independent deformation predictability.

Our results show that the different predictive models employed are capable of learning the relation between the input and the stress–strain curves quite well, as measured by the correlation coefficient between the predicted and simulated stress values at a given strain. In particular, the sample flow stress (i.e., the sample-dependent finite size critical point of the depinning transition) can be predicted surprisingly well with the correlation coefficient reaching values as high as 0.89 for a regularized CNN. We also explore the predictability of individual strain bursts taking place during the deformation process. Critical avalanches are expected to be inherently unpredictable, and indeed, we find approximately zero predictability for most of the strain bursts belonging to the power-law scaling regime of their size distribution. Interestingly, we also find that strain bursts belonging to the cutoffs of the distributions as well as those taking place very close to the flow stress exhibit non-vanishing predictability.

## II. THE MODEL: EDGE DISLOCATION PILEUP WITH QUENCHED DISORDER

The dislocation pileup model we study here is similar to the one in Refs. 30 and 32. It describes a system of  $N$  straight, parallel edge dislocations subject to an external shear stress  $\sigma$  gliding along the direction set by their Burgers vector (here, the  $x$  direction) within a given glide plane. By neglecting the roughness of the dislocation lines, the system may be described by a set of point dislocations with coordinates  $x_i$  representing the cross sections of the dislocation lines

moving in one dimension. The dislocations interact repulsively with each other via long-range stress fields with the stress field magnitude inversely proportional to their mutual distance.<sup>30</sup> In addition, the dislocations are taken to interact via short-range forces with a set of  $N_p$  randomly positioned Gaussian pinning centers, playing here the role of quenched disorder. In real crystals, these pinning centers could be solute atoms or immobile forest dislocations threading the plane of the pileup. The overdamped equations of motion of the  $N$  point dislocations thus read

$$\chi \frac{dx_i}{dt} = \mu b^2 \sum_{j \neq i} \frac{1}{x_i - x_j} + b\sigma + F(x_i), \quad (1)$$

where  $\chi$  is the effective viscosity,  $x_i$  is the position of the  $i$ th dislocation,  $\mu$  is the shear modulus,  $b$  is the Burgers vector magnitude, and  $F$  is the pinning force landscape. We choose a pinning force consisting of Gaussian potential wells centered at pinning sites  $\{x_p\}$ , given by

$$F(x) = -\frac{dE}{dx}, \quad E(x) = -\sum_{p=1}^{N_p} E_p e^{-\frac{1}{2} \left(\frac{x-x_p}{s_p}\right)^2}, \quad (2)$$

where  $x_p$  is the position of the  $p$ th pinning point,  $E_p$  is a pinning energy scale, and  $s_p$  is the standard deviation of the Gaussian. For simplicity, we set  $\chi = \mu = b = 1$  and  $E_p = s_p^2 = 0.5$ . The periodic boundary conditions within a system of size  $L$  are employed by performing an infinite sum over the stress fields of the periodic images of each dislocation,<sup>30</sup>

$$\sum_{k=-\infty}^{\infty} \frac{1}{x+kL} = \frac{\pi}{L \tan(\pi x/L)}. \quad (3)$$

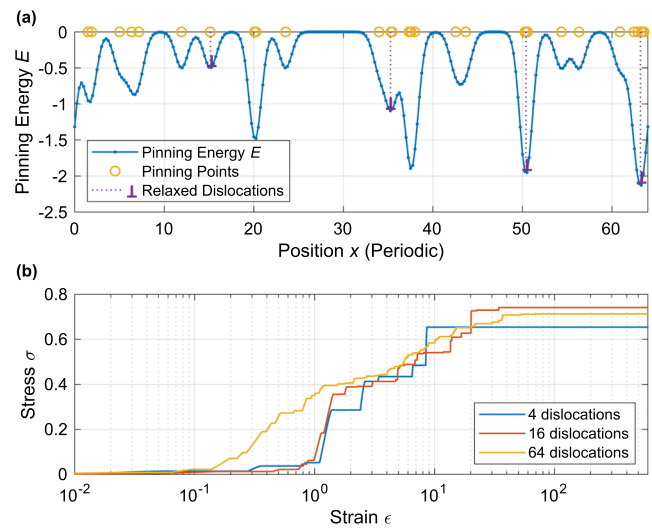
The average spacing between dislocations is set to  $L/N = 16$  and that between pinning sites is set to  $L/N_p = 2$ . The sum over pinning sites is truncated so that only pinning sites that are closer than a cutoff distance of 8 are included in the sum.

The simulations are carried out by first choosing the positions of the  $N_p$  pinning sites at random from a uniform distribution along the one-dimensional system of length  $L$ .  $N$  dislocations are then placed in the system with the constant spacing of  $L/N = 16$ ; this equally spaced dislocation configuration would be the minimum energy state of the pileup in the absence of pinning sites. The dislocations are initially allowed to relax to a metastable configuration with  $\sigma = 0$ ; see Fig. 1(a) for an example of a relaxed configuration with  $N = 4$ , showing also the pinning energy landscape. Then,  $\sigma$  is increased quasistatically from zero with the rate

$$\frac{d\sigma}{dt} = \frac{m_\sigma}{e^{\frac{r[\tilde{v}(t)-v_{th}]}{v_{th}}} + 1}}, \quad (4)$$

where  $m_\sigma = 10^{-4}$  is the maximum stress rate,  $v_{th} = 2 \cdot 10^{-4}$  is a threshold velocity value,  $\tilde{v}(t) = (1/N) \sum_{i=1}^N v_i$  (with  $v_i = dx_i/dt$ ) is the spatially averaged dislocation velocity, and  $r = 100$  is a shape parameter. In practice, this amounts to a continuous, smooth approximation of the step function, with  $d\sigma/dt \approx m_\sigma$  for  $\tilde{v} < v_{th}$  and  $d\sigma/dt \approx 0$  for  $\tilde{v} > v_{th}$ , such that  $\sigma$  is increased only in between strain bursts and kept constant during them. The strain  $\epsilon$  is accumulated by dislocation motion,

$$\epsilon(t) = \frac{1}{N} \sum_{i=1}^N [x_i(t) - x_i(0)]. \quad (5)$$



**FIG. 1.** (a) An example of a relaxed dislocation configuration in the simulation model with  $N = 4$  (and hence of linear size  $L = 16N = 64$ ). The randomly positioned pinning sites (there are  $N_p = L/2 = 32$  of them) are shown as circles, and the resulting pinning energy landscape  $E(x)$  is shown with the solid line. (b) Example stress–strain curves  $\sigma(\epsilon)$  for three different system sizes ( $N = 4, 16$ , and  $64$ ), illustrating the characteristic irregular staircase shape of the curves as well as the size effect where smaller systems exhibit fewer but larger strain bursts.

The simulations are run until the dislocations start to flow indefinitely, i.e., the strain burst size diverges, which happens at a sample-dependent flow stress  $\sigma_{flow}$ . The stress–strain curves  $\sigma(\epsilon)$  are recorded by storing  $\sigma(t)$  and  $\epsilon(t)$  every 20 dimensionless time unit during the simulations. Example stress–strain curves are shown in Fig. 1(b) for three different system sizes. Notice the characteristic irregular staircase-like structure of the curves as well as the larger size of strain bursts for smaller systems.

The resulting stress–strain curves, each corresponding to a different realization of the random pinning landscape and relaxed initial dislocation positions, are stored in a database consisting of 10 000 stress–strain curves for each system size to be used as training and testing data for the predictive models. For the largest systems ( $N = 64, \dots, 512$ ), we also generate a separate dataset of 1000 stress–strain curves stored at a finer time resolution of  $\Delta t = 2$  in order to better detect individual strain bursts. This latter dataset is used to analyze the statistical properties of the strain bursts as well as to train and test ML models for predicting strain bursts.

### III. PREDICTIVE MODELS: FROM LINEAR REGRESSION TO CONVOLUTIONAL NEURAL NETWORKS

In what follows, we will introduce the predictive models used in this study. These encompass both linear and non-linear models using hand-picked features as input as well as a CNN that requires less feature engineering as it uses a complete description of the system’s initial state as input. Using these different models allows us to assess the dependence of our results on any particular model.

Before training, each input feature (each input channel in the CNN case) is standardized by subtracting the mean and dividing by the standard deviation, and the dataset is divided randomly into a training set containing 80% of the data points and a testing set containing the remaining 20%. Predictability is measured by calculating the correlation coefficient between the predicted and target outputs of the testing set and averaging the result over five training instances. Predictive models are implemented using the Python libraries *scikit-learn*<sup>33</sup> for linear absolute shrinkage and selection operator (LASSO) and *Keras*<sup>34</sup> for simple neural networks and CNNs.

The prediction models work by minimizing the *loss*, which mainly consists of the mean squared error. In addition, most models are  $L^1$ -regularized, which adds a penalty term to the loss to prevent overfitting.  $L^1$  regularization often improves the result and encourages the model to focus on the most useful/promising input features while discarding the unimportant, and as a byproduct, it allows us to determine which input features are used by the model.

### A. Linear regression: LASSO

A linear regression model performs an affine transformation, where input features are first linearly combined by multiplication with a (weight) matrix and then biased by adding a translation vector. If the model is optimized using  $L^1$ -regularization for the weight matrix, it is commonly called LASSO (linear absolute shrinkage and selection operator<sup>35</sup>). In this study, this model is applied by using `sklearn.linear_model.Lasso`, which has a built-in optimizer and only requires the user to choose the regularization parameter  $\alpha$ . We found that values of  $\alpha$  in the range  $10^{-3} \lesssim \alpha \lesssim 10^{-2}$  gave good results and chose  $\alpha = 10^{-2}$ .

### B. Simple neural networks

When two linear regression models are stacked in such a way that the intermediate result is *activated* by applying a non-linear activation function, such as the *rectified linear unit (ReLU)*, the combined model becomes a simple neural network that can, to some extent, learn general non-linear mappings. An intermediate result of such a model is called a *hidden layer*, and elements of the intermediate result are referred to as *hidden units*.  $L^1$ -regularization is applied to each layer with a parameter  $\alpha_n$  describing the strength of regularization at each layer.

Tested variants for stress–strain curve prediction differ by the number of hidden units, hidden layers, and regularization. Most have one hidden layer with 64 hidden units and use  $\alpha_1 = 10^{-2}$  for the first weight matrix and  $\alpha_n = 10^{-5}$  for the rest. All models use ReLU-activations for the hidden layers. One variant has less regularization, meaning that  $\alpha_1 = 10^{-3}$  instead of  $10^{-2}$ . Theoretically, one hidden layer is enough for reproducing any continuous mapping given enough hidden units, but a model with three hidden layers is employed here as well for comparison to other studies.<sup>17,36</sup> The models are trained using the *Adam* optimizer<sup>37</sup> with learning rate  $10^{-3}$  for 100 epochs.

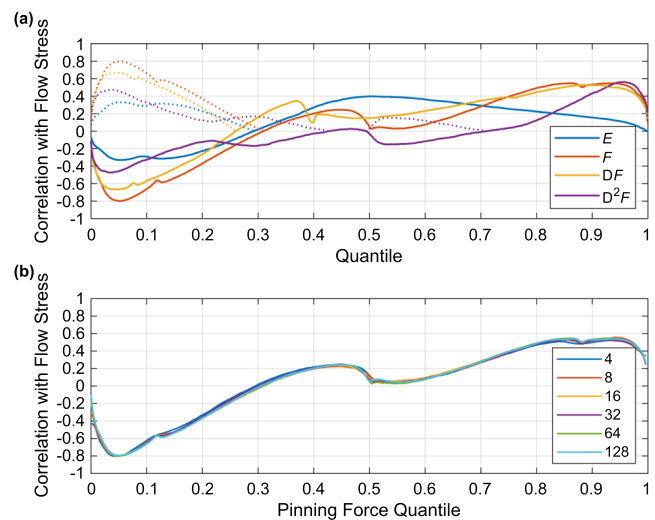
The predictability of strain bursts is studied using a simple neural network having one hidden layer with 64 hidden units. The model output is in this case very big, consisting of a vectorized (one-dimensional) version of an avalanche map that is originally two-dimensional (see Sec. IV for details). We found that all parameters

collapsed to 0 if  $\alpha$  is too high, so a value of  $10^{-5}$  is used for all weight matrices. A lower  $\alpha$  parameter makes the model susceptible to overfitting. Therefore, models are trained only for 10 epochs but with a higher learning rate of  $10^{-2}$ .

### 1. Selecting features for prediction

The LASSO and simple neural network models require hand-picked features for predicting stress–strain curves and strain bursts. It turns out that good, information-rich features for flow stress prediction are given by quantiles of distributions derived from the sample-specific pinning landscapes. Figure 2(a) shows the correlation of the flow stress with quantiles of the pinning energy  $E$ , the pinning force  $F$ , and  $DF = \frac{dF}{dx}$  and  $D^2F = \frac{d^2F}{dx^2}$ . In a non-interacting system, the flow stress would be controlled by the most negative pinning force  $F$ , which provides the strongest obstacle for dislocation motion. In an interacting system, it is possible for the dislocations to push on each other and overcome the strongest pinning force at a smaller applied stress. As a result, the correlation between the most negative value of  $F$  [given by the 0% quantile of  $F$  in Fig. 2(a)] and the flow stress is quite small. Interestingly, the 5% quantile of  $F$  exhibits a much larger negative correlation of  $\sim -0.8$  with  $\sigma_{\text{flow}}$ , independent of the system size [Fig. 2(b)]. Other quantiles of the pinning force as well as of the other quantities listed in Fig. 2(a) also appear to contain useful information for flow stress prediction. We, therefore, give the entire quantile curve (sampled on 64 uniformly chosen points) of each field as input features.

In addition to these input features describing the pinning landscape, we also give input features that describe the initial relaxed configuration of dislocations  $\{x_j\}_{j=1}^N$ . In particular, we use the



**FIG. 2.** (a) Correlation between the flow stress and quantiles of pinning potential energy  $E$ , pinning force  $F$ , derivative of the pinning force  $D(F)$ , and second derivative of the pinning force  $D^2(F)$  for a system of  $N = 64$  dislocations. The dotted lines show the absolute value of the correlation for easier comparison of magnitudes. (b) demonstrates that the correlations of the pinning force quantiles are not dependent on the system size  $N$  (legend). These quantiles are used as hand-picked features for linear regression and simple neural networks. Notice that the 5% quantile of the pinning force exhibits a high (negative) correlation of  $\approx 0.8$  with the flow stress.

values  $E(x_j)$ ,  $F(x_j)$ ,  $DF(x_j)$ , and  $D^2F(x_j)$  as well as the distances between successive dislocations,  $Dx_j = x_{j+1} - x_j$  (with  $x_{N+1} = x_1 + L$ ). These features are given as sorted lists, so they can be interpreted as quantiles.

Finally, we choose a set of features given by local extrema of  $E$  and  $F$ , as well as the successive differences in energy extrema  $D(E$  extrema), also given as 64 uniformly spaced quantiles.

### C. Convolutional neural networks

These models take discretized fields as input and have the ability to find out if the target output depends on the spatial variations within the input fields. The CNN structure used in this study is adapted for periodic input signals by applying a convolution-activation-pooling scheme that continues to down sample the input array down to a single spatial element. Initially, an input field has  $1024 = 2^{10}$  spatial elements, so there are ten convolutional layers (each layer halves the number of spatial elements). This makes the result effectively independent of the particular choice of the origin of the input fields. Therefore, by contrast to more standard CNNs using fewer convolutional layers, this model's final spatial element's receptive field covers the whole system, and translational invariance is complete because the ten pooling operations leave no information left about the absolute locations of where the previous convolutional layers have activated. This is desirable in the case of periodic inputs, where circular shifting does not change the properties (such as the plastic response) of the system.

All convolutional layers employed here use a window size of 3, periodic boundary conditions, ReLU-activation, and non-overlapping max-pooling of size 2 (taking the maximum value within non-overlapping bins) and have 16 hidden units. The result from the final pooling layer is fed into a linear regression layer to obtain the model output. When predicting a single output feature from two input fields (channels), the CNN has 7185 learnable parameters.

The CNN is trained using the same *Adam* optimizer as for other neural network models, with a learning rate of  $10^{-3}$ , a learning rate decay of  $10^{-5}$ , and 50 training epochs. We found that even weak  $L^1$ -regularization caused the parameters to collapse to 0 in most cases, so we do not regularize these models. However, in the case of predicting flow stress only, a kernel regularization value of  $5 \cdot 10^{-4}$  for the convolutional layers is found to work.

The input to the CNNs consists of the pinning landscape fields  $E$ ,  $F$ ,  $DF$ , and  $D^2F$  (although we found that giving one pinning field, such as  $E$  or  $F$ , suffices because the CNN can compute the higher derivatives by differentiation) as well as a field  $J$  defined as a sum of Gaussians centered on the relaxed dislocation positions  $x_i$ ,

$$J(x) = \sum_{i=1}^N \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-x_i)^2}{2}}. \quad (6)$$

This field  $J$  serves to describe the dislocation positions in terms of an input field with a similar structure as the pinning landscape, which is independent of the spatial resolution [see Fig. 7(a) for examples of these input fields]. Notice that by contrast to the case of hand-picked features, these fields are given in order rather than sorted.

## IV. RESULTS

In this section, we first consider the statistical properties of our dataset of stress-strain curves. Then, we report our results for training the predictive models and employ them to explore the predictability of the deformation dynamics of the pileup model. Finally, we also discuss our results on the problem of predicting individual strain bursts.

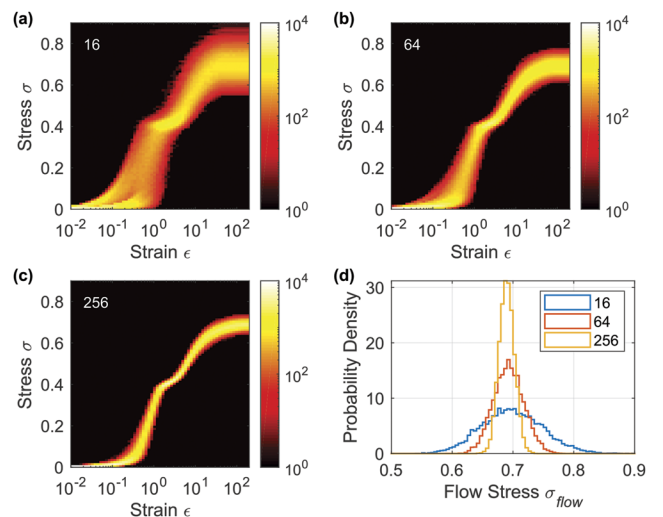
### A. Statistical properties of the stress-strain curves

We start by quantifying the system size dependent sample-to-sample variations in the plastic response of the system. Figures 3(a)–3(c) show the bivariate histograms of the stress-strain curves for three system sizes (with  $N = 16$ , 64, and 256, respectively), illustrating that the sample-to-sample variability in the plastic response is more pronounced in smaller systems. The same applies to the probability densities of the sample flow stresses  $\sigma_{\text{flow}}$ , as shown in Fig. 3(d): The distributions are quite broad for small systems and get narrower upon increasing the system size, presumably approaching a delta peak in the thermodynamic limit. The average magnitude of the  $\sigma_{\text{flow}}$ -values is in agreement with previous results.<sup>30</sup>

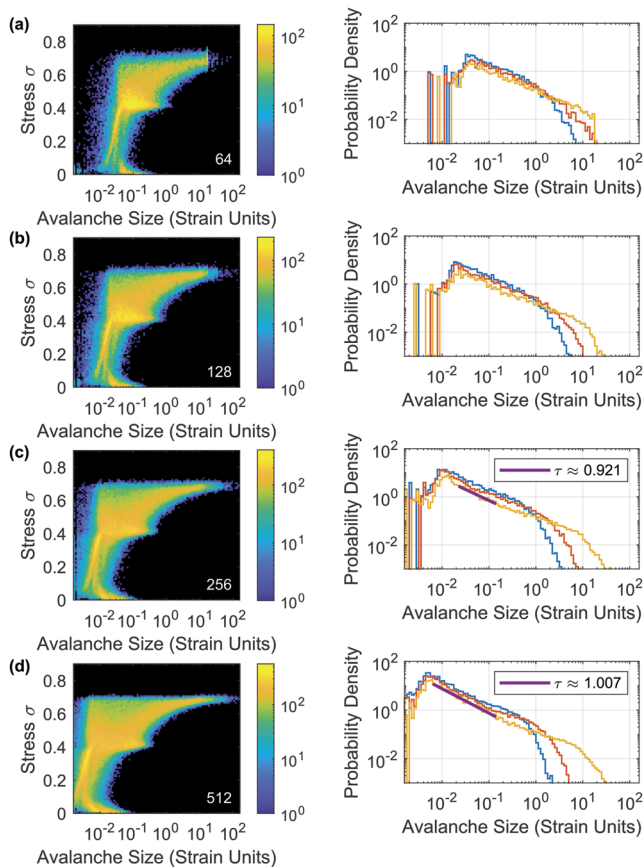
As the stress-strain curves consist of a sequence of discrete strain bursts or dislocation avalanches  $\Delta\epsilon$ , we also consider their stress-dependent size distributions  $P(\Delta\epsilon, \sigma)$ . Quite generally, for systems exhibiting a depinning phase transition, one expects them to obey the scaling form

$$P(\Delta\epsilon, \sigma) = (\Delta\epsilon)^{-\tau} f\left(\frac{\Delta\epsilon}{\Delta\epsilon^*(\sigma)}\right), \quad (7)$$

where  $\tau$  is the critical exponent of the avalanche size distribution,  $f(x)$  is a scaling function that obeys  $f(x) = \text{const.}$  for  $x \ll 1$  and  $f(x) \rightarrow 0$  for  $x \gg 1$ , and  $\Delta\epsilon^*(\sigma)$  is a stress-dependent cutoff avalanche size. The left panels of Figs. 4(a)–4(d) show the bivariate histograms



**FIG. 3.** (a)–(c) show the bivariate histograms of the stress-strain curves for different system sizes [ $N = 16$  in (a),  $N = 64$  in (b), and  $N = 256$  in (c)]. (d) displays the probability distributions of the sample flow stresses  $\sigma_{\text{flow}}$ , showing how the distributions get increasingly narrow upon increasing the system size.



**FIG. 4.** Left panels: Bivariate avalanche histograms (avalanche counts) as a function of the stress  $\sigma$  and avalanche size  $\Delta\epsilon$ . Right panels: The stress-resolved probability distributions of the strain bursts sizes  $P(\Delta\epsilon, \sigma)$ , considering three stress bins: 0.59–0.60 (blue), 0.63–0.64 (red), and 0.67–0.68 (yellow). The average flow stress is around 0.69. (a)  $N = 64$ , (b)  $N = 128$ , (c)  $N = 256$ , and (d)  $N = 512$ . The right panel of (d) shows that the highest stress bin data in the largest system size exhibit a power-law part consistent with an exponent  $\tau \approx 1.0$  [see Eq. (7)].

of the avalanche count as a function of  $\sigma$  and avalanche size  $\Delta\epsilon$  for different system sizes (with  $N = 64, 128, 256$ , and  $512$ , respectively). These illustrate that, indeed, the maximum strain burst size increases upon approaching the flow stress from below, and the increase close to the flow stress is more clear-cut for larger systems. Notice also that these histograms show an excess of strain bursts around  $\sigma \approx 0.4$ . This feature is likely related to the characteristic pinning force magnitude due to an isolated pinning point given by Eq. (2). The right panels of Figs. 4(a)–4(d) show the corresponding stress-resolved strain burst size distributions  $P(\Delta\epsilon, \sigma)$ , considering three different stress bins below the flow stress for each system size. For the largest system sizes considered, the data appear to be consistent with a power-law part characterized by an exponent  $\tau \approx 1.0$  [see the right panels of Figs. 4(c) and 4(d)]. This value is somewhat lower than the expectation of  $\tau \approx 1.25$  based on the scaling of the elastic energy of the pileup in the Fourier space,<sup>30</sup> suggesting that the pileup model would be in the same universality class with depinning of contact lines<sup>38</sup>

and planar cracks.<sup>39,40</sup> As we focus here on studying rather small systems, this difference is likely due to finite size effects.

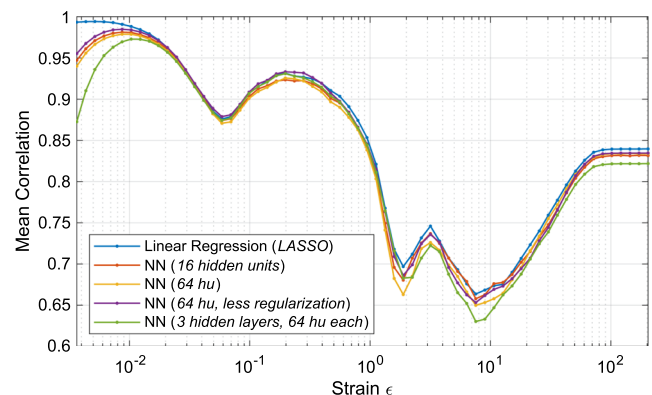
Overall, the statistical properties of the plastic response of the pileup model show two key features relevant for deformation predictability: (i) there are sample-to-sample variations in the response (that we will try to predict using machine learning) and (ii) the distribution of strain bursts  $\Delta\epsilon$  displays the characteristics of a critical system, suggesting that individual strain bursts may be inherently unpredictable. In what follows, we will use our trained predictive models to explore how well (i) the sample stress–strain curves and (ii) individual strain bursts can be predicted.

## B. Predictability of the sample stress–strain curves

### 1. LASSO and simple neural networks

Figure 5 shows the mean correlation coefficient between the predicted and actual stress values as a function of strain for  $N = 64$ , using linear regression (LASSO) and different variants of simple neural networks. First, we note that the flow stress prediction is only slightly better than the magnitude of the correlation with the 5% quantile of the pinning force landscape found above ( $\sim 0.8$ , see Fig. 2), with the correlation coefficient assuming values in the range of 0.825–0.84 depending on the model used (the large-strain plateau in Fig. 5). Second, the correlation clearly depends on strain, exhibiting a rather pronounced dip reaching values as low as 0.65 or so for intermediate strains, along with two additional, lesser dips for smaller strains, before reaching a value close to 1 as the strain goes to zero. The latter is true especially for LASSO, while the neural network models seem to struggle a bit in capturing the deformation dynamics for very small strains. The reduced predictability for intermediate strains could be related to the importance of inherently unpredictable dislocation avalanches in analogy with the results of Ref. 17.

It is interesting to analyze which features the predictive models end up using for prediction and how that might depend on strain.

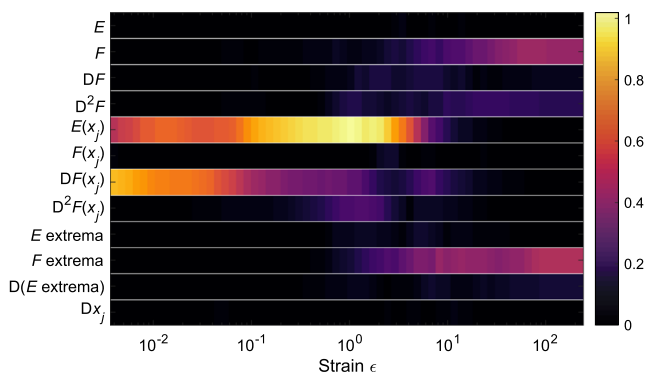


**FIG. 5.** Predictability of the stress–strain curves for the system size  $N = 64$  using hand-picked features, measured by the correlation between the predicted and actual stresses in the test set. The shown result is a sample mean over five trained models, each using their own random division into training and testing sets. Each line corresponds to a given predictive model, as indicated in the caption. The plateau at large strains corresponds to the predictability of the flow stress.

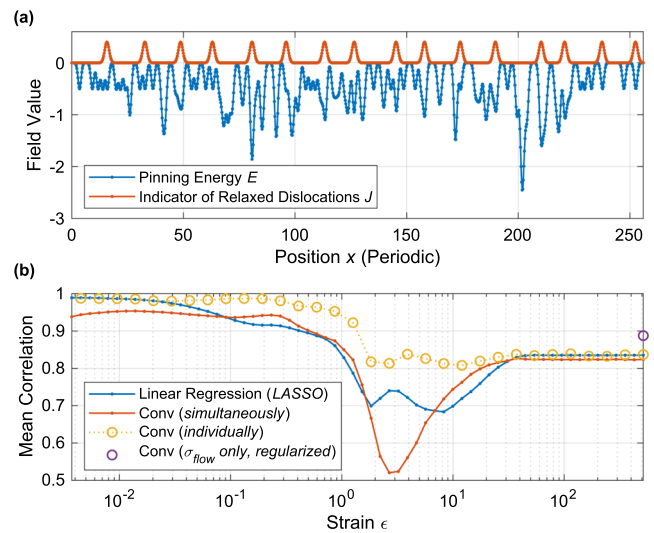
Since the models utilize heavy  $L^1$ -regularization, the analysis of the learned weights becomes straightforward. Features that contribute to predictability are allowed to have a high regularization penalty in exchange for decreasing loss, while unimportant features will not decrease loss sufficiently to justify a high regularization penalty. Therefore, if we were to omit features with a small  $L^1$  penalty, the loss would not be affected and the model would presumably do just as well, so this analysis is closely related to studying feature ablations. Figure 6 shows how much penalty is generated by the weights corresponding to each feature category for the LASSO model. At the start, the pinning force derivative at the relaxed dislocation positions,  $DF(x_j)$ , is the most active feature, but also the pinning energy at the initial dislocation positions,  $E(x_j)$ , is used. This is reasonable as such quantities could be used to estimate the initial linear response of the dislocations. Interestingly, and perhaps somewhat counter-intuitively, the pinning energy at the initial dislocation positions becomes even more important for intermediate strains in addition to a larger diversity of features used. The flow stress prediction uses mainly the quantiles of the pinning force  $F$  as well as its local extrema points. This dependence of the important features on strain indicates that predicting the response of the system to infinitesimal stresses is of a different nature compared to predicting the flow stress.

## 2. Predictability of stress-strain curves: Convolutional neural networks

We then proceed to employ CNNs as predictive models. Figure 7(b) shows the mean correlation as a function of strain as obtained using three different types of CNNs along with the LASSO result for reference. Here, the system size  $N = 16$  is considered. First, we consider a CNN that predicts the entire stress-strain curve at once [a “generalist” CNN, shown as a red line in Fig. 7(b)]. This model actually underperforms the simple linear regression model for the majority of the strain values. This could be due to the



**FIG. 6.** Weight analysis of the trained LASSO model for the  $N = 64$  system. Color indicates the total  $L^1$  penalty ( $L^1$  norm) from each feature category's weights as a function of strain, averaged over five training instances. Feature categories are explained in Sec. III B 1. A high penalty can be interpreted to indicate the high importance/usefulness of the corresponding feature category for predicting the stress at the specific strain. Notice that the properties of the initial dislocation positions are most useful at small strains, while more generic features of the force landscape become more useful at larger strains.



**FIG. 7.** (a) Example input fields for a CNN with  $N = 16$ . Dot markers show the spatial resolution. (b) Stress-strain curve predictability at  $N = 16$  for different models: LASSO (blue line), a “generalist” CNN predicting the entire stress-strain curve at once (red line), individual “specialist” CNNs trained specifically to predict the stress at a single strain (yellow circles), and an  $L^1$ -regularized CNN trained to only predict the flow stress (purple circle).

fundamentally different nature of the problem for different strain values, as illustrated by Fig. 6, showing that different features are important for different strains. Hence, the “generalist” CNN might get exhausted and not be able to optimally learn the entire stress-strain curve at once.

Thus, we also consider a “specialist” CNN that is trained separately for each strain value to learn the corresponding value of stress. The result is shown as yellow circles in Fig. 7(b). Indeed, as compared to the generalist model, the specialist model is able to learn the deformation dynamics of the pileup model much better, especially for intermediate strains, where the generalist CNN exhibits a deep dip in the correlation coefficient, presumably due to the importance of hard to predict dislocation avalanches for those strains. It is interesting that the specialist model is able to reach a correlation coefficient exceeding 0.8 also in this regime.

Notably, neither of the models discussed above is able to predict the flow stress much better than by just considering the correlation with the 5% quantile of the pinning force  $F$  (see Fig. 2). We, therefore, try to further improve the specialist CNN by switching on  $L^1$ -regularization and using only the pinning force  $F$  as the input field (this was shown above to be the most informative quantity for flow stress prediction). The individual purple point at a large strain value in Fig. 7(b) is the resulting flow stress prediction with a correlation coefficient of 0.89, i.e., significantly better correlation than using any of the other models considered.

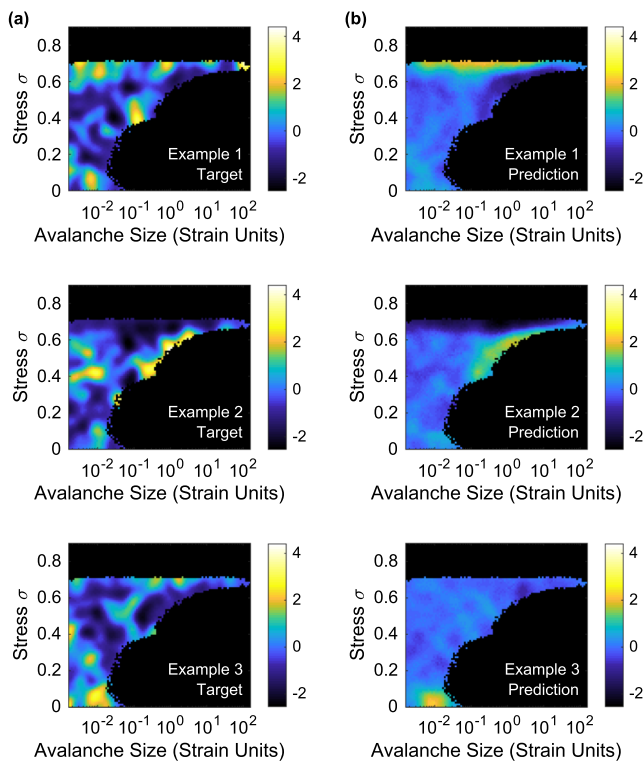
## C. Predictability of strain bursts

We also explore the possibility of predicting individual strain bursts along the stress-strain curves. First, we need to formulate

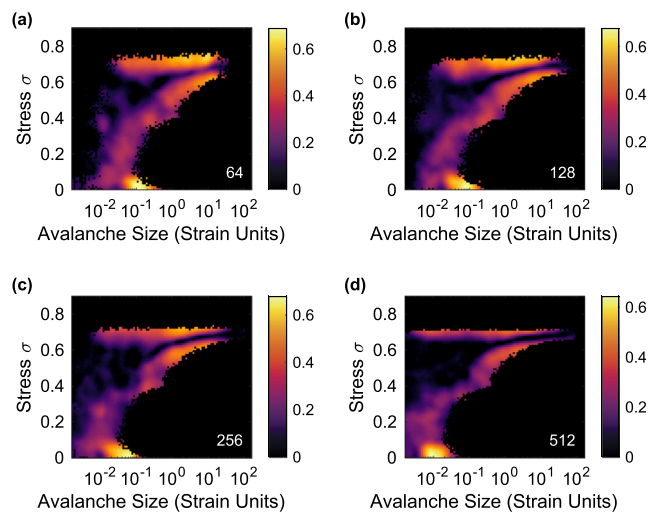


a problem that is suitable for analysis using the predictive models at hand. To this end, we generalize the bivariate avalanche histograms (left panels of Fig. 4) to the case of individual samples. As such, these would be rather sparse sets of points in the space spanned by the avalanche size and stress. We, therefore, smooth out these maps by convolving with a Gaussian with a standard deviation of 0.03 units of stress in the vertical direction and 0.15 units of  $\log_{10}(\text{avalanche size})$  in the horizontal direction; both correspond to three times the bin size used in Figs. 4 and 8. The studied region is limited to bins where at least one avalanche appears within the dataset so that the Gaussian smoothing does not cause distribution boundaries to extend to impossible areas. Finally, the maps are standardized by removing the mean and dividing by the standard deviation at each bin. We then employ a simple neural network with 64 hidden units to find a mapping from the hand-picked input features to the smoothed sample-specific bivariate avalanche histograms.

Three different examples of smoothed avalanche maps are shown in the left column of Fig. 8 with the corresponding predicted avalanche maps in the right column. While the predictions are clearly not perfect, the algorithm is able to capture some key features of these avalanche maps. Figure 9 shows the correlation maps between the predicted and actual avalanche densities at four different system sizes, averaged over five training instances. These show a number of interesting features: (i) Large parts of the sample-specific



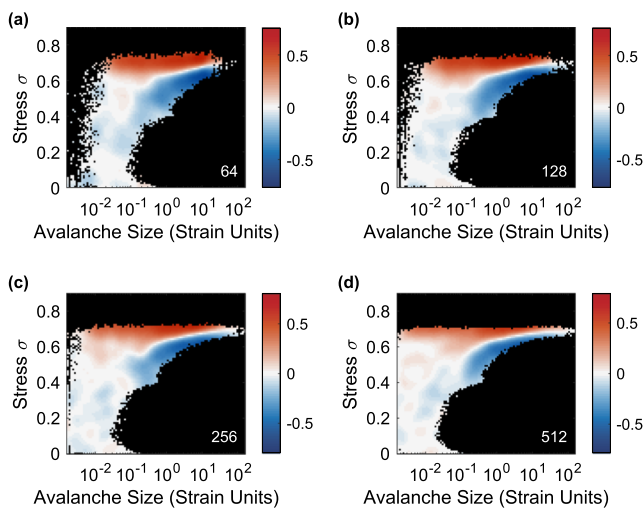
**FIG. 8.** Three examples (rows) of strain burst maps of individual systems for  $N = 512$ . The left column (a) shows three target maps, and on the right (b) are the corresponding predictions of a simple neural network.



**FIG. 9.** Correlation maps between the predicted and actual sample-specific avalanche maps for different system sizes, (a)  $N = 64$ , (b)  $N = 128$ , (c)  $N = 256$ , and (d)  $N = 512$ . Notice the positive correlations for the largest avalanches for each stress value (avalanche cutoff) and for most of the avalanches very close to the flow stress, as well as the approximately zero correlation for avalanches belonging to the power-law scaling regime (see Fig. 4).

avalanche histograms appear to be completely unpredictable, with the correlation coefficient assuming a value very close to zero (shown in black in Fig. 9). This is especially true for most of the regions corresponding to the power-law scaling regime of the avalanche size distribution in agreement with the idea that critical avalanches should be intrinsically hard to predict.<sup>41</sup> (ii) Avalanches with the largest size for each stress value (i.e., those belonging to the cutoff of the stress-resolved avalanche size distribution) appear to exhibit some degree of predictability, although the related correlation coefficient values are somewhat lower than those found for the predictability of the stress-strain curves. This is in agreement with the idea that avalanches in the cutoff are not critical and hence somewhat predictable. (iii) Surprisingly, we also find the non-vanishing predictability of avalanches of any size occurring in the immediate proximity of the flow stress (bright horizontal segments in the upper parts of the panels of Fig. 9). This is surprising because those events are expected to be critical and hence should be unpredictable.

In order to shed some light on the origin of the key features of avalanche predictability discussed above, we finish by considering the correlation between the avalanche count and flow stress. The logic here is that since we have demonstrated above that the sample-specific flow stress can be predicted quite well, if the avalanche maps are correlated with the flow stress, these should be predictable as well. Figure 10 reveals the two stand-out features: (i) The avalanches taking place in the immediate proximity of the flow stress exhibit clear positive correlation with the flow stress and (ii) the largest avalanches for stresses smaller than the flow stress (i.e., avalanches belonging to the stress-dependent cutoff of the avalanche size distribution) are negatively correlated with the flow stress. This means that both a *higher* than average number of avalanches very close to



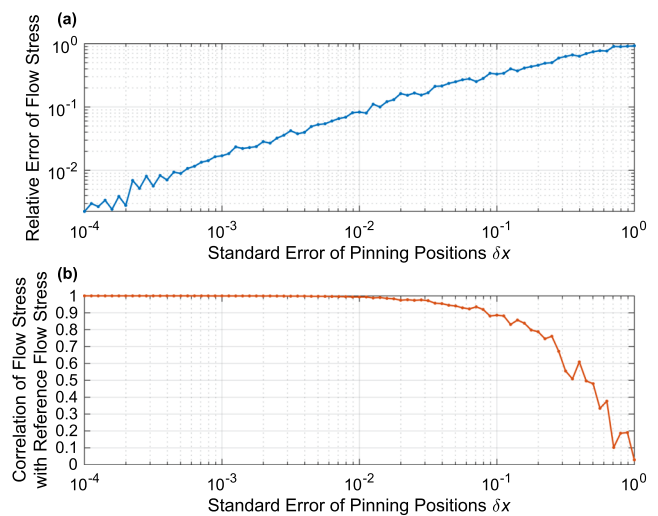
**FIG. 10.** Correlation maps between the avalanche count and flow stress for the avalanche maps considered in Fig. 8. (a)  $N = 64$ , (b)  $N = 128$ , (c)  $N = 256$ , and (d)  $N = 512$ . Notice the positive correlation between the flow stress and avalanches close to the flow stress, as well as the negative correlation between the flow stress and the largest avalanches occurring for smaller stresses.

the flow stress and a lower than average number of large avalanches at smaller stresses imply a higher than average flow stress value. The good level of predictability of the flow stress demonstrated above thus translates into the reasonable predictability of these features of the sample-specific bivariate avalanche histograms. Notice that the widths of the predictable bands in the avalanche maps (Fig. 9) appear to become thinner with increasing system size, so we cannot exclude the possibility that this avalanche predictability would be a finite size effect.

#### D. Robustness of the pileup model against small perturbations

We finish by considering the effect of small perturbations of the pinning point positions on the flow stress value. Notice that these perturbations also affect important features such as the extreme quantiles of the pinning force landscape as large pinning force values originate from overlapping force fields of more than one pinning point. This, therefore, serves as a measure of the robustness of our results on deformation predictability in situations where the initial microstructure can be characterized with only a finite precision. To that end, we perturb the initial pinning point positions by random amounts drawn from a normal distribution with mean 0 and standard deviation  $\delta x$ . Figure 11 compares the resulting flow stress with the unperturbed one as a function of  $\delta x$ , by measuring the relative difference [Fig. 11(a)] and the correlation between the perturbed and unperturbed flow stress [Fig. 11(b)]. The results are calculated for the system of 64 dislocations using a set of 100 reference cases, which are compared with the corresponding perturbed systems with varying  $\delta x$ . Although slightly noisy, the overall shape of the curves is clearly visible.

Figure 11 shows that perturbing the pinning point positions has essentially no effect below a standard error of  $\delta x \sim 10^{-2}$ , and a



**FIG. 11.** (a) Relative error of the flow stress (standard error relative to the theoretical upper limit that is the expected error between two uncorrelated samples from the flow stress distribution) and (b) correlation of the flow stress with a reference (unperturbed) flow stress value as a function of  $\delta x$ , the standard error of pinning point positions (the standard deviation of the normal distribution from which the perturbations are drawn from).

significant effect of perturbing the pinning point positions becomes visible only when the perturbation magnitude clearly exceeds that threshold value. This suggests that the dislocation pileup model and hence our results concerning deformation predictability in that model are essentially unchanged even if the pinning point positions (or, more generally, the features describing the pinning potential) are perturbed a little.

#### V. DISCUSSION AND CONCLUSIONS

We have established that predictive models ranging from linear regression to CNNs can be used to predict the stress–strain curves of the pileup model with a good accuracy. While the different models employed give rise to somewhat different results, a practical conclusion is that the model giving the highest correlation coefficient for a given strain provides a lower limit for the deformation predictability of the pileup system. Notably the “specialist” CNN discussed above results in correlation coefficients exceeding 0.8 over the entire range of strains considered, and adding regularization further improves the correlation coefficient of the flow stress prediction to 0.89. These should be interpreted as lower limits of deformation predictability as we obviously cannot exclude the possibility that a hypothetical predictive model not considered here might outperform our models. For flow stress prediction, the 5% quantile of the pinning force landscape turned out to be a surprisingly information-rich feature, highlighting the fact that our method of using quantiles as features might be useful also in other contexts.

While the predictability of the stress–strain curves turns out to be quite good, it is not perfect. This is to be expected because the stress–strain curves consist of critical-like strain bursts that are expected to be inherently hard to predict. Indeed, our attempts

to predict the sample-level bivariate avalanche count resulted in essentially zero predictability for most of the critical dislocation avalanches belonging to the power-law distributed part of the avalanche size distribution. On the other hand, these sample-level avalanche distributions also showed an interesting regularity: A larger than average flow stress value indicates a lower than average number of the largest avalanches for stresses below the flow stress and a larger than average number of avalanches in the immediate vicinity of the flow stress. It would be interesting to extend such an analysis from our stress-controlled simulations to strain-controlled ones. One should point out that because of the unpredictable nature of critical dislocation avalanches, the ML models cannot fully replace the integration of the equations of motion as they are unable to fully resolve such strain bursts. However, the fact that the overall shape of the stress–strain curve is still predictable to some degree suggests that the trained ML models could be used for simple classification tasks like determining if a given sample is “weak” or “strong” compared to the average, using significantly less computational resources than what is required for integrating the equations of motion.

An interesting issue that is also of practical relevance has to do with the accuracy of the input data used for prediction. Here, we have access to the precise microscopic information of the pinning force landscape, and indeed, the features we use have been constructed from such precise information that might not be available in real experimental systems. However, our robustness analysis of the pileup model suggests that extreme precision is not needed for good predictions, and the same conclusion can be made by remembering that our hand-picked features are constructed in a way that does not preserve the information of the precise locations of the pinning points. Thus, we expect our general methodology to be applicable also in situations where the initial microstructure can be characterized with a finite precision only.

To put our work in the context of the broader literature on applying ML to physics problems, let us point out some details: First, incorporating the inductive biases from different physical systems into machine learning architectures has recently received attention,<sup>42–44</sup> the key idea being to use ML to learn missing physical relations from data by constructing models that mimic the equations of motion and then applying them to interpolate or extrapolate the original data. One should note that our aim here is simply to find a mapping from features characterizing the initial microstructure to the plastic response of the sample, without explicitly incorporating information about the laws of physics in the predictive models. When constructing the CNN, we make use of the fact that the system we study is periodic, but otherwise it is up to the ML model(s) to learn the above-mentioned mapping without explicit input related to the physics of the system. A second relevant development in ML has to do with incorporating the domain knowledge in ML models.<sup>45</sup> In the context of our study, the feature engineering process that resulted in the features we use for LASSO and simple neural networks can be argued to incorporate some “domain knowledge,” as the features used reflect, at least to some extent, our physical intuition as to what the important features might be. Third, our study dealing with a rather simple dislocation pileup model allowing us to calculate an extensive training database with relative ease could perhaps serve as a “sister system” for related systems such as more realistic/complicated models as well as experiments where

data are scarce, in analogy to the approach of Hoffmann *et al.*<sup>31</sup> For instance, in systems with a higher dimensionality than the one considered here, such an approach might work well especially for small strains, where the dislocation dynamics is governed by the linear response independent of the system’s spatial dimensionality. For larger strains, the formation of tangled dislocation structures resulting in strain hardening might, however, complicate the situation, as such features are not captured by our one-dimensional pileup model.

We finish by noting that the pileup model studied here has the advantage that the dislocation system will sample the entire one-dimensional random energy landscape, and hence, extracting relevant descriptors of the sample-specific quenched disorder is straightforward. This is in contrast to other systems exhibiting a depinning phase transition such as magnetic domain walls in disordered ferromagnetic thin films,<sup>46</sup> planar crack fronts in disordered solids,<sup>39,40,47</sup> or individual dislocation lines interacting with point defects.<sup>48</sup> In such systems, driven by an external force in a direction perpendicular to the average elastic line direction, the disorder sampled by a given realization is not known *a priori*. Nevertheless, it would be interesting to extend the present study to consider the predictability of depinning dynamics in such systems. Finally, we point out two crucial outstanding questions related to deformation predictability: The problem remains to be addressed in realistic three-dimensional DDD simulations, with<sup>29,49</sup> or without<sup>26</sup> additional defects interfering with dislocation motion, as well as in experiments where descriptors of the initial sample microstructure could be extracted, e.g., by x-ray measurement techniques.<sup>50–52</sup>

## ACKNOWLEDGMENTS

The authors wish to thank Henri Salmenjoki for technical help on machine learning. We acknowledge the financial support of the Academy of Finland via the Academy Project COPLAST (Project No. 322405).

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- 1 Y. Q. Cheng and E. Ma, “Atomic-level structure and structure–property relationship in metallic glasses,” *Prog. Mater. Sci.* **56**, 379–473 (2011).
- 2 M. Delincé, Y. Bréchet, J. D. Embury, M. G. D. Geers, P. J. Jacques, and T. Pardoen, “Structure–property optimization of ultrafine-grained dual-phase steels using a microstructure-based strain hardening model,” *Acta Mater.* **55**, 2337–2350 (2007).
- 3 A. J. Ardell, “Precipitation hardening,” *Metall. Trans. A* **16**, 2131–2165 (1985).
- 4 M. D. Uchic, D. M. Dimiduk, J. N. Florando, and W. D. Nix, “Sample dimensions influence strength and crystal plasticity,” *Science* **305**, 986–989 (2004).
- 5 D. M. Dimiduk, C. Woodward, R. LeSar, and M. D. Uchic, “Scale-free intermittent flow in crystal plasticity,” *Science* **312**, 1188–1190 (2006).
- 6 D. M. Dimiduk, M. D. Uchic, and T. A. Parthasarathy, “Size-affected single-slip behavior of pure nickel microcrystals,” *Acta Mater.* **53**, 4065–4077 (2005).
- 7 L. Zdeborová, “Machine learning: New tool in the box,” *Nat. Phys.* **13**, 420–421 (2017).

- <sup>8</sup>G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, "Machine learning and the physical sciences," *Rev. Mod. Phys.* **91**, 045002 (2019).
- <sup>9</sup>S. Papanikolaou, M. Tzimas, A. C. Reid, and S. A. Langer, "Spatial strain correlations, machine learning, and deformation history in crystal plasticity," *Phys. Rev. E* **99**, 053003 (2019).
- <sup>10</sup>M. Raissi and G. E. Karniadakis, "Hidden physics models: Machine learning of nonlinear partial differential equations," *J. Comput. Phys.* **357**, 125–141 (2018).
- <sup>11</sup>P. Baldi, P. Sadowski, and D. Whiteson, "Searching for exotic particles in high-energy physics with deep learning," *Nat. Commun.* **5**, 4308 (2014).
- <sup>12</sup>K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, "Machine learning for molecular and materials science," *Nature* **559**, 547–555 (2018).
- <sup>13</sup>P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, and A. J. Norquist, "Machine-learning-assisted materials discovery using failed experiments," *Nature* **533**, 73–76 (2016).
- <sup>14</sup>G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, "Accelerating materials property predictions using machine learning," *Sci. Rep.* **3**, 2810 (2013).
- <sup>15</sup>Z.-L. Liu, P. Kang, Y. Zhu, L. Liu, and H. Guo, "Material informatics for layered high- $T_c$  superconductors," *APL Mater.* **8**, 061104 (2020).
- <sup>16</sup>R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakithodi, and C. Kim, "Machine learning in materials informatics: Recent applications and prospects," *npj Comput. Mater.* **3**, 54 (2017).
- <sup>17</sup>H. Salmenjoki, M. J. Alava, and L. Laurson, "Machine learning plastic deformation of crystals," *Nat. Commun.* **9**, 5307 (2018).
- <sup>18</sup>Z. Yang, S. Papanikolaou, A. C. Reid, W.-k. Liao, A. N. Choudhary, C. Campbell, and A. Agrawal, "Learning to predict crystal plasticity at the nanoscale: Deep residual networks and size effects in uniaxial compression discrete dislocation simulations," *Sci. Rep.* **10**, 8262 (2020).
- <sup>19</sup>H. Salmenjoki, M. J. Alava, and L. Laurson, "Mimicking complex dislocation dynamics by interaction networks," *Eur. Phys. J. B* **91**, 275 (2018).
- <sup>20</sup>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. Research* **2**, 023338 (2020).
- <sup>21</sup>S. Biswas, D. F. Castellanos, and M. Zaiser, "Prediction of creep failure time using machine learning," *Sci. Rep.* **10**, 16910 (2020).
- <sup>22</sup>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**, 9276–9282, <https://doi.org/10.1002/2017gl074677> (2017).
- <sup>23</sup>M.-C. Miguel, A. Vespignani, M. Zaiser, and S. Zapperi, "Dislocation jamming and Andrade creep," *Phys. Rev. Lett.* **89**, 165501 (2002).
- <sup>24</sup>L. Laurson, M.-C. Miguel, and M. J. Alava, "Dynamical correlations near dislocation jamming," *Phys. Rev. Lett.* **105**, 015501 (2010).
- <sup>25</sup>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).
- <sup>26</sup>A. Lehtinen, G. Costantini, M. J. Alava, S. Zapperi, and L. Laurson, "Glassy features of crystal plasticity," *Phys. Rev. B* **94**, 064101 (2016).
- <sup>27</sup>M. Ovaska, A. Lehtinen, M. J. Alava, L. Laurson, and S. Zapperi, "Excitation spectra in crystal plasticity," *Phys. Rev. Lett.* **119**, 265501 (2017).
- <sup>28</sup>M. Ovaska, L. Laurson, and M. J. Alava, "Quenched pinning and collective dislocation dynamics," *Sci. Rep.* **5**, 10580 (2015).
- <sup>29</sup>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. Materials* **4**, 083602 (2020).
- <sup>30</sup>P. Moretti, M.-C. Miguel, M. Zaiser, and S. Zapperi, "Depinning transition of dislocation assemblies: Pileups and low-angle grain boundaries," *Phys. Rev. B* **69**, 214103 (2004).
- <sup>31</sup>J. Hoffmann, Y. Bar-Sinai, L. M. Lee, J. Andrejevic, S. Mishra, S. M. Rubinstein, and C. H. Rycroft, "Machine learning in a data-limited regime: Augmenting experiments with synthetic data uncovers order in crumpled sheets," *Sci. Adv.* **5**, eaau6792 (2019).
- <sup>32</sup>F. Leoni and S. Zapperi, "Slip line growth as a critical phenomenon," *Phys. Rev. Lett.* **102**, 115502 (2009).
- <sup>33</sup>F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).
- <sup>34</sup>F. Chollet *et al.*, Keras, <https://keras.io>, 2015.
- <sup>35</sup>R. Tibshirani, "Regression shrinkage and selection via the LASSO," *J. Roy. Stat. Soc. B* **58**, 267–288 (1996).
- <sup>36</sup>Y. Miyazawa, F. Briffod, T. Shiraiwa, and M. Enoki, "Prediction of cyclic stress-strain property of steels by crystal plasticity simulations and machine learning," *Materials* **12**, 3668 (2019).
- <sup>37</sup>D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014).
- <sup>38</sup>J. F. Joanny and P. G. De Gennes, "A model for contact angle hysteresis," *J. Chem. Phys.* **81**, 552–562 (1984).
- <sup>39</sup>D. Bonamy, S. Santucci, and L. Ponson, "Crackling dynamics in material failure as the signature of a self-organized dynamic phase transition," *Phys. Rev. Lett.* **101**, 045501 (2008).
- <sup>40</sup>L. Laurson, S. Santucci, and S. Zapperi, "Avalanches and clusters in planar crack front propagation," *Phys. Rev. E* **81**, 046116 (2010).
- <sup>41</sup>C.-K. Pun, S. Matin, W. Klein, and H. Gould, "Prediction in a driven-dissipative system displaying a continuous phase transition using machine learning," *Phys. Rev. E* **101**, 022102 (2020).
- <sup>42</sup>Y. Bar-Sinai, S. Hoyer, J. Hickey, and M. P. Brenner, "Learning data-driven discretizations for partial differential equations," *Proc. Natl. Acad. Sci. U. S. A.* **116**, 15344–15349 (2019).
- <sup>43</sup>M. Cranmer, A. Sanchez-Gonzalez, P. Battaglia, R. Xu, K. Cranmer, D. Spergel, and S. Ho, "Discovering symbolic models from deep learning with inductive biases," [arXiv:2006.11287](https://arxiv.org/abs/2006.11287) (2020).
- <sup>44</sup>A. Beatson, J. T. Ash, G. Roeder, T. Xie, and R. P. Adams, "Learning composable energy surrogates for PDE order reduction," [arXiv:2005.06549](https://arxiv.org/abs/2005.06549) (2020).
- <sup>45</sup>Z. Ren, F. Ovidio, M. Thway, S. I. Tian, Y. Wang, H. Xue, J. D. Perea, M. Layurova, T. Heumueller, E. Birgersson *et al.*, "Embedding physics domain knowledge into a Bayesian network enables layer-by-layer process innovation for photovoltaics," *npj Comput. Mater.* **6**, 9 (2020).
- <sup>46</sup>S. Zapperi, P. Cizeau, G. Durin, and H. E. Stanley, "Dynamics of a ferromagnetic domain wall: Avalanches, depinning transition, and the Barkhausen effect," *Phys. Rev. B* **58**, 6353 (1998).
- <sup>47</sup>L. Laurson, X. Illa, S. Santucci, K. T. Tallakstad, K. J. Måløy, and M. J. Alava, "Evolution of the average avalanche shape with the universality class," *Nat. Commun.* **4**, 2927 (2013).
- <sup>48</sup>S. Zapperi and M. Zaiser, "Depinning of a dislocation: The influence of long-range interactions," *Mater. Sci. Eng., A* **309–310**, 348–351 (2001).
- <sup>49</sup>A. Lehtinen, L. Laurson, F. Granberg, K. Nordlund, and M. J. Alava, "Effects of precipitates and dislocation loops on the yield stress of irradiated iron," *Sci. Rep.* **8**, 6914 (2018).
- <sup>50</sup>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**, 602–607 (2001).
- <sup>51</sup>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**, 619–622 (2006).
- <sup>52</sup>E. Schafler, K. Simon, S. Bernstorff, P. Hanák, G. Tichy, T. Ungár, and M. J. Zehetbauer, "A second-order phase-transformation of the dislocation structure during plastic deformation determined by *in situ* synchrotron x-ray diffraction," *Acta Mater.* **53**, 315–322 (2005).