2019

Deformation Correlations and Machine Learning: Microstructural inference and crystal plasticity predictions

Michail Tzimas
West Virginia University, mt0041@mix.wvu.edu

Follow this and additional works at: https://researchrepository.wvu.edu/etd

Part of the Mechanics of Materials Commons, and the Other Materials Science and Engineering Commons

Recommended Citation
Tzimas, Michail, "Deformation Correlations and Machine Learning: Microstructural inference and crystal plasticity predictions" (2019). Graduate Theses, Dissertations, and Problem Reports. 3827.
https://researchrepository.wvu.edu/etd/3827

This Thesis is brought to you for free and open access by The Research Repository @ WVU. It has been accepted for inclusion in Graduate Theses, Dissertations, and Problem Reports by an authorized administrator of The Research Repository @ WVU. For more information, please contact ian.harmon@mail.wvu.edu.
Deformation Correlations and Machine Learning: Microstructural inference and crystal plasticity predictions

Michail Tzimas

Thesis submitted
to the Benjamin M. Statler College of Engineering and Mineral Resources
at West Virginia University

in partial fulfillment of the requirements for the degree of
Master of Science in
Mechanical Engineering

Stefanos Papanikolaou, Ph.D., Chair
Ever Barbero, Ph.D.
Nasser Nasrabadi, Ph.D.

Department of Mechanical and Aerospace Engineering

Morgantown, West Virginia
2019

Keywords: Machine Learning, Thin films, Deformation History, Discrete Dislocation Dynamics, Microstructure

Copyright 2019 Michail Tzimas
Abstract

Deformation Correlations and Machine Learning: Microstructural inference and crystal plasticity predictions

Michail Tzimas

The present thesis makes a connection between spatially resolved strain correlations and material processing history. Such correlations can be used to infer and classify prior deformation history of a sample at various strain levels with the use of Machine Learning approaches. A simple and concrete example of uniaxially compressed crystalline thin films of various sizes, generated by two-dimensional discrete dislocation plasticity simulations is examined. At the nanoscale, thin films exhibit yield-strength size effects with noisy mechanical responses which create an interesting challenge for the application of Machine Learning techniques. Moreover, this thesis demonstrates the prediction of the average mechanical responses of thin films based on the classified prior deformation history and discusses the possible ramifications for modelling crystal plasticity behavior in extreme settings.
Acknowledgements

I would like to thank Dr. Papanikolaou for giving me the opportunity to work on the present project. I take pride in the skills I developed due to his guidance, and the research opportunities he introduced me to, in the past few years. I would also like to thank the friends I made while working in West Virginia University for their support and all the good times we had. You made me feel welcome in a place thousands of miles away from home. Finally, a heartfelt thank you to my family both in the United States and back home in Greece for all their support. Without them, I would not be here today.
# Table of Contents

Abstract                                                                                       ii
Acknowledgements                                                                               iii
List of Figures                                                                                v
1. Introduction & Thesis Overview                                                           1
   1.1. Introduction                                                                       1
   1.2. Thesis Overview                                                                   2
2. Literature Review                                                                       3
   2.1. Machine Learning & Applications in Material Science                                3
       2.1.1. Preprocessing tools in unsupervised Machine learning                          4
   2.2. Discrete Dislocation Dynamics                                                        6
       2.2.1. Fundamentals and Mathematical Formulation                                    7
       2.2.2. Applications and Examples of Discrete Dislocation Dynamics                  8
3. Discrete Dislocation Dynamics Simulation Data                                              10
   3.1. Theory and Data Definitions                                                          10
   3.2. High Throughput Data Generation                                                     14
   3.3. Strain history and spatial dislocation correlations                               15
4. Application of Machine Learning Algorithms on Spatial Strain Correlation Profiles:  
   The Unsupervised Approach                                                               17
   4.1. Preprocessing                                                                       17
   4.2. Clustering and classification                                                        18
   4.3. Examples                                                                           21
       4.3.1. Double Slip System                                                           21
       4.3.2. Single Slip System                                                          28
   4.4. Accuracy & Validation                                                              29
   4.5. Uncertainty sensitivity and quantification                                         33
5. Application of Machine Learning Algorithms on Spatial Strain Correlation Profiles:  
   The Supervised Approach                                                                 36
6. Prediction of mechanical responses and material properties                                37
7. Ongoing and future work                                                                  42
8. Conclusion                                                                               45
Bibliography                                                                                47
List of Figures

Figure 1: Schematic of the various loading histories & testing deformation on samples: A material has an assumed prior deformation history (Stage L, red circles). We unload the sample and obtain the Stage U (blue circles). How does the strain field, which characterizes Stage T (green circles), reflect the prior history? The testing deformation $\varepsilon_T - \varepsilon_U$ is constant in all T-U cases._________________________ 2

Figure 2: Importance of feature scaling: Example from (Lanigan and Raschka, 2007). ___________ 5

Figure 3: Spherical Indentation in 2D-DDD: From (Song et al., 2019) __________________________ 9

Figure 4: Histograms of abrupt events and cutoff dependence. From: (Papanikolaou et al., 2017). __ 9

Figure 5: The 2D discrete dislocation plasticity model of uniaxial compression of thin films: Slip planes (lines) span the sample, equally spaced at $d = 10b$, but planes close to corners are deactivated to maintain a smooth loading boundary. Surface and bulk dislocation sources are present (disks) and forest obstacles are spread homogeneously across slip planes. Initially the sample is free of stress and mobile dislocations.___________10

Figure 6: Boundary plastic steps and stress/strain profiles in dislocation configurations: The full sample is shown in undeformed (a, c) and deformed coordinates (b). (a) A dislocation snapshot at 10% strain for a sample of width $w=2$ and double slip systems. A high density of dislocation is observed, close to $p = 10^{15}$ m$^{-2}$. T's and L's have been used to describe the dislocations that create positive stress (T's) and negative stress (L's) on the different slip systems. (b) The $e_{xx}$-strain profile that corresponds to the dislocation configuration of (a) is shown in the displaced coordinates. The boundary is filled with rough plastic steps. The dashed rectangular box denotes the region for which we consider spatial strain correlations, that are then used for ML purposes. The strain map is unitless. (c) The shear stress $\tau_{xy}$ of the configuration in (a, b) is shown. The stress map is in MPa. ____________12

Figure 7: Strain profiles captured from the 2D-DDD simulation at stage L and stage T -- $w = 2 \mu m$ – Double slip system: The full sample is shown, in both deformed (a) or undeformed (b) coordinates. A sample is loaded to 10% strain, unloaded to zero stress and reloaded to testing deformation of 0.1%. (a) Strain profile for Stage L at 10% strain. Plastic steps are allowed to emerge on the film surface (Papanikolaou et al., 2017). (b) A strain profile at the Stage T, after subtracting the residual plastic deformation at Stage U. Such strain profiles are analogous to typical DIC experimental strain profiles. The strain maps are unitless._________________________13

Figure 8: A schematic of the sequence of events when loading, unloading and reloading a sample is shown, with the corresponding stress and strain graphs vs the time step of simulation. A sample is obtained from 2D-DDD simulation and it's loaded to a specific strain value – “Prior” Loaded stage (Stage L). Then, the sample is unloaded to zero stress and the remaining plastic strain can be calculated – “Prior” Unloaded stage (Stage U). Finally, the sample is reloaded to a testing deformation – “Test” Reloaded stage (Stage T). Even though a sample has been plastically deformed (Stage L), the samples obtained from experiments can be polishes, thus the surface of a sample is not able to provide information about deformation (Sample at stage U can be seen in the figure as having smooth surface). Such techniques are applicable in experiments such as digital image correlation, where randomly placed tracking nanoparticles are detected optically and contribute to correlation statistics are applied to the sample. Then, as the sample is reloaded, the permanent deformation can be observed, since there are changes in the distances between tracked nanoparticles.__________________________14

Figure 9: Variety of strain profiles in 2D-DDD simulations for smaller and larger testing strain: For the spatial scale of the figures see Fig. 7. A sample is loaded to a high deformation of strain (which could be either 1 % or 10 %) and then unloaded. This pre-deformed sample is then reloaded to a testing strain. In (a), (b) the samples are loaded to 10% strain, while in (c), (d) the samples are loaded to 1% strain (a) Small testing deformation (0.1%). $w=2 \mu m$. Double slip system simulation. (b) Large testing deformation 1%. $w=1 \mu m$. Single slip system simulation (c) Small testing deformation (0.1%). $w=1 \mu m$. Single slip system simulation. (d) Large testing deformation (1%). $w=2 \mu m$ Double slip system simulation. For description of color map see Fig. 7 (b).__15
Figure 10: \( w = 2 \mu m \), 2D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) auto-correlation. (a) Projection of data set on first two principal components. Red blobs denote samples with 0.1\% strain (stage L), blue triangles samples with 1\% strain (stage L) and green squares denote samples with 10\% strain (stage L), respectively. (b) First principal component of PCA, shown in sample coordinates (Fig. 7, Fig. 9, Sec. 3). (c) Second principal component of PCA, shown in sample coordinates (Fig. 7). The colormaps are unitless, showing the intensity the PCA-transformed correlations.

Figure 11: Variance of principal component: It is observed that after PCA, the first principal component retains the highest percentage of variance (~97\%) and components 2 and 3, retain about 2\% variance.

Figure 12: \( w = 1 \mu m \), 3D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10. 3 different clusters are shown like in Fig. 11. Introducing the 3rd component into the PCA map, does not affect the results.

Figure 13: \( w = 1 \mu m \) -- 2D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10. (a) Projection of data set on first two principal components. (b) First principal component of PCA, shown in sample coordinates (Fig. 7). (d) Second principal component of PCA, shown in sample coordinates (Fig. 7).

Figure 14: \( w = 0.5 \mu m \) -- 2D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10. (a) Projection of data set on first two principal components. For description of colormaps, see Fig. 10. (b) Projection of data set on first two principal components without a clustering algorithm applied to the data set, demonstrating a failure in clustering the various deformation levels. (c) First principal component of PCA, shown in sample coordinates (Fig. 7). (d) Second principal component of PCA, shown in sample coordinates (Fig. 7). For description of colormaps, see Fig. 10.

Figure 15: \( w = 1 \mu m \) -- The choice of the correlation domain and how it impacts the PCA maps -- Double slip system: \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10. (a) Projection of data set on first two principal components. We have rich phenomenology present towards the center of the domain and at the boundaries. (b) 100 \( \times \) 100 domain of the correlation matrix. A highly focused area near the center of the domain is shown, where the phenomena are focused. The smoothness present in (a) is slowly removed from this domain. (c) 200 \( \times \) 200 domain of correlation matrix. Highly smooth in the center and towards the boundaries of the domain. (d) PCA maps for 40 \( \times \) 40 domain. (e) PCA map for 100 \( \times \) 100 domain. The variance of the data has changed, and the projections have shifted. The information provided by (b) does not change the cluster formations but introduces unnecessary information that has shifted the results along the PC1 and PC2 axes. (f) PCA map for 200 \( \times \) 200 domain. (g) The variance of the data has changed even more compared to (f). The distances between the blue and green clusters have increased an order of magnitude compared to (e) and 2 orders of magnitude compared to (d). The information provided by (c) does not affect the clusters that are formed from our algorithm.

Figure 16: \( w = 0.25 \mu m, w = 0.125 \mu m \) -- 2D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10 (a). Figures on the left correspond to classified samples, with colors assigned by the algorithm. Figures on the right are samples with colors assigned by their initial load. (a) \( w = 0.25 \mu m \). The clustering algorithm fails and puts all points in a single cluster, corresponding to Stage L at 10\% strain. (b) \( w = 0.25 \mu m \). Actual representation of initial deformations. (c) \( w = 0.125 \mu m \). The clustering algorithm fails and puts all points in a single cluster, corresponding to Stage L at 10\%. (d) \( w = 0.125 \mu m \). Actual representation of initial deformations.

Figure 17: Large-reload vs. small-reload testing -- Example of PCA projection results for thin films of \( w = 0.5 \mu m \): \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10 (a). Strain profiles are created from the quantity \( \epsilon_T \) (a) Sample with stage L = 10\% strain is unloaded to zero stress and then reloaded to small testing deformation (Stage T = 0.1\%). (b) Stage T at small testing deformation (0.1\%), without a clustering algorithm applied to data set. Projection on two principal components. Actual representation of the data set, with
some mixing of the samples. The clusters have shifted closer to one another but not indistinguishable. (c) Sample with stage L = 10% strain is unloaded to zero stress and then reloaded to large testing deformation (Stage T = 1%). (d) Stage T at large testing deformation (1%), without a clustering algorithm applied to data set. Actual representation of the data set. For higher testing deformation of 1%, we can see that there is much more mixing of the samples. Reloading to higher strain values, adds plastic memory to the samples, rendering our process inapplicable for these cases. For description of color map see Fig. 7 (b).

Figure 18: Large-reload vs. small-reload testing -- Example of PCA projection results for thin films of w = 0.5 μm: h0, h0 auto-correlation. Classified Samples. The colors follow the definition of Fig. 10. (a) PCA map for classified samples of small testing deformation (0.1%). The CkNN algorithm fails to find multiple clusters due to the small distance between PCA projected samples. (b) PCA for classified samples of large testing deformation (1%). The CkNN algorithms finds two clusters, recognizing samples of 1% and 10% initial deformation, and ultimately fails.

Figure 19: First and second principal component of PCA application on thin films of w = 0.5 μm shown in sample coordinates: (a) First principal component - Stage T = 0.1% (b) First principal component - Stage T = 1% (c) Second principal component - Stage T = 0.1% (d) Second principal component - Stage T = 1%. For description of color maps, see Fig. 10.

Figure 20: w = 2μm -- 2D projection of PCA results for thin films -- Comparison between single and double-slip system: h0, h0 auto-correlation. The colors follow the definition of Fig. 10. Projection of data set on first two principal components. (a) Single slip system projection (b) Double slip system projection.

Figure 21: w = 2μm -- 2D projection of PCA results for thin films -- Comparison of principal components among double and single slip systems: Components shown correspond to the analysis of Fig. 18. (a) First principal component of double slip system. (b) First principal component of single slip system. (c) Second principal component of double slip system. (d) Second principal component of single slip system. For description of color maps, see Fig. 10.

Figure 22: w = 2μm -- 2D projection of PCA results for thin films -- Double slip system -- Validation: h0, h0 auto-correlation. Red blobs denote samples with 0.1 % strain (stage L), blue triangles samples with 1 % strain (stage L) and green squares denote samples with 10 % strain (stage L), respectively. Red stars depict testing samples of 0.1 % strain (stage L), blue stars testing samples of 1 % strain (stage L) and green stars testing samples of 10 % strain (stage L). Validated-split data set. Projection on first two principal components.

Figure 23: Measures of success for classification of samples -- 0.1 % testing deformation: h0, h0 auto-correlations. The x-axis of each graph is the base 2 logarithm of the various sample widths we examine. (a) Accuracy score for the samples. Maximum value 1 means that all the samples have been correctly classified. (b) F1-score of our 3 clusters that are formed. The line with the squares represents the cluster with samples at stage L = 10 % strain, while the line with the triangles is for the cluster with samples at stage L = 1 % strain. Finally, the line with the circles is for the cluster with samples at stage L = 0.1 % strain. For smaller sized systems we have observed that most of the samples are classified as belonging in the “square” cluster, hence the scored value for that cluster only. Since the algorithm correctly classifies the samples that were initially loaded to 10 % strain, but also classifies more samples as belonging to that cluster, then the score does not have the maximum value of 1 but lower. (c) F2-score of our 3 cluster that have formed. The definition of the colored lines follows (b). Since for F2-score we have increased weight of the recall, the 0.7 maximum value is expected for the “square” cluster. (d) F0.5-score of our 3 clusters. The colors definition follows (b). Since we have reduced weight of the precision, for lower sample widths it is expected to have lower score than F1 for the “square” cluster.

Figure 24: Measures of success for classification of samples -- 0.1 % testing strain: h0, h0 auto-correlations. The x-axis of each graph is percentage of samples tested for classification. (a) Accuracy score for samples of w = 2 μm (stars) and w = 1 μm (disks). Maximum value 1 means that all the samples have been correctly classified. (b) Averaged F1-score across the 3 clusters that have formed for samples of w = 2 μm (stars) and w = 1 μm (disks). It is obvious that we have good agreement for the classified samples even when we test less than 30 % of the total number of samples.
Figure 25: Effect of strain invariant type for pre-processing -- Examples of PCA projection maps -- 

**J2 strain tensor invariant:** The colors follow the definition of Fig. 10 (a) h0,h0 autocorrelations, \( w = 2 \mu m. \) (b) h1,h1 autocorrelations, \( w = 1 \mu m. \)  

Figure 26: Residual/Plastic strain for pre-processing: The colors follow the definition of Fig. 10. h0,h0 autocorrelations, \( w = 2 \mu m. \)  

Figure 27: Plastic strain calculations for J2 and \( \phi \) strain-tensor invariants -- Examples of PCA projection maps: The colors follow the definition of Fig. 10. (a) h0,h0 autocorrelations. \( w = 2 \mu m. \) J2 invariant. (b) h0,h0 autocorrelations. \( w = 2 \mu m. \) \( \phi \) invariant. (a) h0,h0 autocorrelations. \( w = 1 \mu m. \) J2 invariant. \( \phi \) invariant.  

Figure 28: \( w = 2\mu m, 2D \) projection of IPCA results for thin films -- Double slip system: h0, h0 autocorrelation. The colors follow the definition of Fig. 10. 

Figure 29: Schematic for obtaining the 1% strain mechanical response of unknown samples: A new test sample is mechanically loaded to 0.1% strain, spatial strain correlations are captured and used to find the classification based on the previously tested samples (whose classification is known by Supervised ML algorithms). Depending on the exported classification, the future mechanical response can be predicted, and is the average mechanical response of the previously tested samples of the same class. 

Figure 30: Prediction of the average mechanical response of 2 \( \mu m \) samples with a known prior deformation history  

Figure 31: Prediction of the average mechanical response of 1 \( \mu m \) samples with a known prior deformation history  

Figure 32: Prediction of the average mechanical response of 0.5 \( \mu m \) samples with a known prior deformation history  

Figure 33: Prediction of the average mechanical response of 0.25 \( \mu m \) samples with a known prior deformation history  

Figure 34: Prediction of the average mechanical response of 0.125 \( \mu m \) samples with a known prior deformation history  

Figure 35: Prediction of maximum stress-width relationship in thin films  

Figure 36: Supervised learning of nanoindentation samples with \( \rho = 1.0e + 11 \) [m \(-2\)]: An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples, with 100% accuracy.  

Figure 37: Supervised learning of nanoindentation samples with \( \rho = 1.0e + 12 \) [m \(-2\)]: An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples, with 100% accuracy.  

Figure 38: Supervised learning of nanoindentation samples with \( \rho = 5.0e + 12 \) [m \(-2\)]: An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples. A misclassification of a sample is seen, and the expected accuracy is lowered.
1. Introduction & Thesis Overview

1.1. Introduction

It is an axiom of materials science that microstructures controls properties, and that the microstructure is heavily influenced by material processing (e.g. cold/hot rolling, extrusion) (Cottrell, 1990). Due to this indirect link between processing history and material properties, samples of the same shape and composition but of different processing technique, may exhibit change in mechanical properties. In the context of crystal plasticity, a major component for possible property sensitivity is plastic deformation. The question then becomes: How to distinguish the processing history of crystalline samples?

Signatures of crystal plasticity, have been traditionally linked to strain localization, typically in the form of shear bands (Asaro and Lubarda, 2006; Bigoni and Hueckel, 1991). Shear bands have been identified as possible indicators of prior deformation: In particular, it has been shown that the creation of slip bands during various small-load mechanical tests strongly depend on the prior deformation history of a sample (Diehl and Hinzner, 1964; Novák et al., 1984a, 1984b). However, the presence or absence of shear bands may not suffice for characterizing prior processing. For a complete characterization of crystal plasticity history, other signatures should be considered, such as the full local structure of the stress and strain tensor fields. Moreover, the knowledge of such signatures becomes mandatory at the microscale, where size effects and stochastic deformation can be a defining factor for plastic deformation. The spatial structure of the microstructure may also be reflected in spatially resolved strain correlations (Chaikin et al., 1995; Papanikolaou et al., 2013, 2007; Raman et al., 2008) which can be thought as using augmented observables including strain gradients (Aifantis, 1987; Fleck et al., 1994; Gurtin, 2000), to successfully classify the material response.

The present thesis focuses on distinguishing the material deformation history of samples by using strain correlations signatures, and the prediction of future large-strain behavior based on the prior processing history. Simple and tractable examples of crystal plasticity are used to infer the prior deformation history of crystalline aluminum thin films generated by two-dimensional discrete dislocation dynamics (2D-DDD) simulations with well-controlled mechanical processing histories. Given that dislocation movements are the principal instigators of plastic deformation on materials, this study captures a large number of experimentally relevant cases and can provide a transparent application framework (der Giessen and Needleman, 1995; Nicola et al., 2006; Papanikolaou et al., 2017; Song et al., 2016). The mechanical properties of test samples are examined by small strain loading, but only after subtracting prior mechanical memory. Using unsupervised and supervised machine learning (ML), this thesis shows that spatial correlations encoded in the strain images of the reloaded samples, contain adequate
information to produce a full classification of dislocation-driven deformation history at multiple scales, and that a subsequent testing will produce a mechanical response that may be described as the average response of specimens with the same processing history.

1.2. Thesis Overview

In this study, the simulated systems begin in a state where they have been compressed and released for one cycle, with different amplitudes of compression (prior deformation history). After release, samples are then reloaded to a second, small amplitude compression (testing deformation). The strain signatures that are observed during the second compression are imaged (in the form of shear bands) and are then used as input to the ML process. The goal is that the ML process, accurately captures the prior deformation history of the samples. This process is meant to mimic a readily-accessible, non-destructive (small strain testing) experimental technique, namely, digital image correlation (DIC) which has approached nanoscale resolutions due to improved quality in microscopy techniques (Kammers and Daly, 2013a, 2013b; Peters et al., 1983; Peters and Ranson, 1982; Schreier et al., 2009; Sutton et al., 2007, 1986, 1983), but images originating from other material characterization techniques (e.g. SEM (Goldstein et al., 2017), TEM (Williams and Carter, 1996)) may be used. Figure 1 shows a schematic of the process for three different amplitudes of strain:

![Schematic](image_url)

*Figure 1: Schematic of the various loading histories & testing deformation on samples: A material has an assumed prior deformation history (Stage L, red circles). We unload the sample and obtain the Stage U (blue circles). How does the strain field, which characterizes Stage T (green circles), reflect the prior history? The testing deformation \( \epsilon_T - \epsilon_U \) is constant in all T-U cases.*
2. Literature Review

2.1. Machine Learning & Applications in Material Science

Machine learning, artificial intelligence, computer vision and data mining are words that are often used interchangeably in the computer science community to describe the same process: A computer is taught to perform tasks without instructions but relying on past actions and information to infer new actions. It is worth noting, that while the previous words may be used interchangeably, they are all different technical terms that are normally subsets of one another and should not be confused. In that regard, the present thesis will focus on the word Machine Learning.

Machine learning (ML) involves the study of computer algorithms that are used to teach a computer how to perform actions without prior given instructions (Bishop, 2006). In other words, the computer “learns” by the algorithm how to operate/perform tasks. At its core, a ML algorithm is a way to mathematically model known data, so that the model will provide a prediction or a decision on a set of unknown data. Machine learning can be split into two main categories: Supervised learning, which is the study of known inputs and known outputs (i.e. classification and regression) (Alpaydin, 2010), and unsupervised learning, where the inputs are known but the outputs are unknown (clustering techniques).

Supervised Learning is used in tasks of regression like predicting the sale price of a house, based on sale prices of houses in the same neighborhood, with the same number of bedrooms/bathrooms and features in general. Supervised learning is highly used in everyday life: SPAM messages, Google Searches, Google recommendations on email responses, are all based on supervised learning. Huge collections of data are used to train an algorithm which should recognize which mails are spam, based on similar features (for example, spam messages contain certain words and/or numbers in the title). Based on previous actions of users who have marked emails as spam, the algorithm learns to identify new spam mail. One of the adages of ML is that “more data gives better results”, which is heavily featured in supervised ML algorithms. In supervised algorithms, the dataset is split into a training set and a test set: The algorithm is trained on the training set, and the accuracy and validity are further examined on the test set.

Unsupervised learning is a class of Machine Learning techniques to find commonalities/patterns in data. The data that is given to the algorithm is unlabeled, i.e. only the features are fed into the algorithm, with no corresponding knowledge of the output/expected variables. In an unsupervised algorithm, the machine is trying to discover similar structures in the data and aims to cluster data with similarities. For that reason, the most common unsupervised method is cluster exploratory data analysis.

In material science, machine learning is critical in areas such as new materials discovery and material property prediction (Liu et al., 2017; Ramprasad et al., 2017). Studies have been focused on predicting electronic properties, designing new materials
with desirable properties (Liu et al., 2015; Pilania et al., 2013), identifying material properties in nanoindentation (Meng et al., 2015), predicting phase behaviors of grain boundaries (Zhu et al., 2018), and creating links between microstructure and properties (Choudhury et al., 2016; Kalidindi, 2012; Khosravani et al., 2017), typically in a supervised learning problem.

The present thesis is separated into two parts: First, an unsupervised learning problem: Clustering samples based on their prior deformation history. The machine is not aware of the amplitudes of prior strain, and spatial correlations on strain signatures are used to find common patterns in the samples. The result would be to find as many clusters as prior deformation amplitudes. The strain signatures that are observed during the second compression are imaged (in the form of shear bands) and are then used as input to the machine learning process. The goal is that the ML process, accurately describes the prior deformation history of the samples. The second part, describes the use of supervised machine learning algorithms for predicting the mechanical response of samples. Decision Trees (Quinlan, 1986) are used for training samples with known prior deformation history, and we show that the mechanical response of materials in further deformation, are influenced by the prior deformation and their size.

2.1.1. Preprocessing tools in unsupervised Machine learning

The goal in an unsupervised ML method is too find common patterns in the given dataset and learn to classify samples based on these common patterns. However, for any given dataset, the descriptors/features of it, may be counted in tens or hundreds. This high-dimensional nature of datasets prohibits the immediate visualization of the dataset, and statistical tools may need to be applied on the features. Moreover, the magnitude of features may be different per feature, and even in the case where visualization is possible (low-dimensional dataset), the features that have the greatest magnitude will be thought of as the primary commonality that may be found in the dataset – a wrong assumption of the ML algorithm that can be easily avoided.

Given the above restrictions, in ML approaches, and after the dataset has been collected, there are two steps that need to be completed before finding common patterns: First, the standardization of each feature of the dataset. Standardization involves rescaling the features such that they have the properties of a standard normal distribution with a mean of zero and a standard deviation of one. With that step, only the features that truly contribute to patterning of the dataset will be used by the ML algorithms. Secondly, to allow for easier visualization of high-dimensional datasets, dimensionality reduction methods are used on the dataset, the most common one being Principal Component Analysis (PCA, (Shlens, 2003)).

PCA is a statistical approach that finds an ordered set of orthogonal basis vectors that efficiently describes the variance in a data set (Shlens, 2003) similar to Singular Value Decomposition (SVD), where the singular vectors are called principal components and
the basis is defined in such a way that the first principal component has the largest possible variance (that is, the first singular value accounts for as much of the variability in the data as possible), and the following components in turn have the highest variance possible, while still being orthogonal to all the preceding components. For this reason, it is critical that all samples in the data set were acquired under identical controlled conditions with a clear understanding of the origin of the variability (hence the use of standardization methods (Pedregosa et al., 2011)). PCA is commonly used as one step in a series of analyses; one can use it to reduce the number of variables, especially when there are too many predictors relative to the number of observations. To illustrate the importance of these two steps, and how they work, Fig. 2 shows how scaling of the features helps PCA, and how PCA allows for classification of a dataset with 13 features:

![Figure 2: Importance of feature scaling: Example from (Lanigan and Raschka, 2007).](image)

Mathematically, given a data matrix $D$, whose columns each contain the components of a single data point, $D$ may be decomposed to a diagonal matrix of singular values, $S$, and left/right singular vector matrices $V$ and $U$, with $D = V^T S U$. The columns of $V$ and $U$ are the left and right singular vectors of $D$. The $V$ vectors that correspond to the largest singular values capture the most characteristic features.

The difference between PCA and SVD lies in the fact that PCA can also be derived through the covariance matrix $C = D^T D / (n - 1)$, where $n$ is the number of samples in the data matrix. The covariance symmetric matrix $C$ can be further diagonalized into $C = V^T L V$, where projections of the data points on eigenvectors $V$ are called principal
components. To find the principal components through SVD, after decomposing the data matrix, one needs to further consider that the covariance matrix from SVD can be found through:

\[
C = DD^T = \frac{V^T S U U^T S V}{n - 1} = V^T \frac{S^2}{n - 1} V
\]

Where the connection between the two expressions of the covariance matrix is the same, as long as \( L = S^2 / (n - 1) \). Given a set of vectors \( V \), data samples can be projected onto subspaces where each subspace can be any pair of vectors \( V_i \). In the present thesis, we examine subspace defined by the first 3 vectors of \( V \), (principal components) denoted as PC1, PC2, and PC3.

### 2.2. Discrete Dislocation Dynamics

In materials science, crystals can be typically visualized by their arrangement of atoms in what is known as a unit cell, and a crystal is called perfect when the alignment of atoms in the cell is perfect. Dislocations are line defects that introduce an extra half plane of atoms in the unit cell, creating stress fields that influence the atoms in the rest of the unit cell. Physically, a crystal can be deformed irreversibly by slip, the shifting of the crystal along one of its planes. If the shift is equal to the lattice vector of the crystal, then the displacement is permanent. This first understanding of deformation by slip was observed in the late 19\textsuperscript{th} century by (Ewing and Rosenhain, 1900). They observed formation of slip bands on the surface after deformation of a metal specimen.

Over the past decades, it has been well established that dislocation are the principle instigators of crystal plasticity, and that they strongly influence the mechanical properties of materials (Hirth and Lothe, 1982). Due to the nature of dislocations (defects in the crystal) and their typical size (nano/atomic scale), continuum crystal plasticity models do not typically include dislocation mechanisms, due to the complexity of describing the overall motions of dislocations and incorporating these motions of atomistic defects into a continuum representation.

To overcome the limitations of continuum models, various discrete dislocation dynamics (DDD) models have been developed since the 1980’s. The first discrete dislocation models, were two-dimensional (2D), using periodic boundary conditions (i.e. escaping dislocations would re-enter the sample), and contained dislocations where the movements and behavior was governed by simple rules (Amodeo and Ghoniem, 1988; der Giessen and Needleman, 1995; Groma and Pawley, 1993; Le and Stumpf, 1996). Although the main mechanisms of dislocation were captured in these models, the information was limited due to the 2D character of the models. The last three decades, with the emergence and rapid evolution of computer hardware, new dislocation dynamics (DD) approaches have been developed in three-dimensional space (3D) and a hope for breakthrough in dislocation behavior and understanding of crystal plasticity was generated (Canova et al., 1993, 1992; Hirth et al., 1996; Zbib et al., 1998).
In these 3D models, the motion, formation and multiplication of dislocations as well as the effect and interactions of dislocations with other defects is usually considered. However, new complications emerged, due to the overall complexity of the problem, since these models must keep track of all these interactions and mechanisms accurately and efficiently (fast). The progress and complexity of DDD models has been further magnified by the idea to couple DD with continuum mechanics analysis in association with computational algorithms such as finite elements. Such coupling will allow material scientist to better understand crystal plasticity mechanisms in multi-scale approaches, increasing the potential for future applications of this method in material, mechanical, structural and process engineering analyses. In the following, the fundamentals of a mathematical formulation of DD analysis are presented, to better understand the model presented in Section 3.

2.2.1. Fundamentals and Mathematical Formulation

A dislocation is a line defect in a crystal (perfect otherwise) described by its line sense vector $\xi$ and Burgers vector $b$. The Burgers vector has two components: edge, perpendicular to its line sense vector, and screw, parallel to its line sense vector. In that sense, when the Burgers vector is perpendicular to $\xi$, the crystal has an edge dislocation; Screw dislocations exist on the second case. Under loading, dislocations glide and propagate on slip planes causing deformation and change of shape. In the case where there is a screw character of the local line (parallel to burgers vector), there is a change of cross-slip, or the change of slip planes by the dislocation. Cross-slip is a very important phenomenon which must be dealt with in DD and is the main instigator of the three-dimensional character of the motion of dislocations. Other phenomena that need to be accounted for in 3D-DDD and describe most if not all the mechanisms of dislocations include:

1. The dislocation topology; 3D geometry, Burgers vector and vector $\xi$.
2. Identification of all possible slip planes for each dislocation.
4. Multiplication and annihilation of dislocation segments.
5. Formation of complex connections and intersections such as junctions, jogs and branching of the dislocation in multiple directions.

Dislocation segments may be annihilated if their Burgers vector is equal in magnitude and opposite in direction. Moreover, junctions, jogs and branching in multiple directions may happen between cross-slipped dislocations which are perpendicular to each other, or otherwise connected. It can be easily inferred that the motion of dislocations is a complex problem. Typically, the dislocation dynamics are governed by a “Newtonian” equation of motion, including inertia, damping, driving force terms from all possible interactions (short or long range). In general, due to these terms, a dislocation needs to overcome internal drag forces, lattice friction parameters (Peierls stresses (Joós
and obstacles (stacking faults, vacancies). Most of these parameters exert a short-range interaction on the dislocations, by way of an internal strain field, which leads to stochastic (noisy) perturbation on the encountered dislocations. Therefore, the internal strain fields of local obstacles will irregularly perturb the equation of motion and this needs to be accounted for. In summary, the dislocation may encounter the following set of forces:

- Drag force, \( Bv \), where \( B \) is the drag coefficient and \( v \) is the dislocation velocity.
- Peierls stress \( F_{\text{Peierls}} \).
- Force due to externally applied loads, \( F_{\text{External}} \).
- Dislocation-dislocation interaction force \( F_D \).
- Dislocation self-force \( F_{\text{Self}} \).
- Dislocation-obstacle interaction force \( F_{\text{Obstacle}} \).
- Image force \( F_{\text{Image}} \).
- Osmotic force \( F_{\text{Osmotic}} \) resulting from climb of dislocations and results in the absorption or emission of point defects.
- Thermal force \( F_{\text{Thermal}} \) to account for thermal fluctuations.

### 2.2.2. Applications and Examples of Discrete Dislocation Dynamics

Over the past decade, discrete dislocation dynamics have been utilized by several researchers to investigate many complicated small-scale crystal plasticity phenomena that occur under a wide range of loading and boundary conditions and covering a wide spectrum of strain rates. Some of the major phenomena that have been addressed include:

- The role of dislocation mechanisms in strain hardening (Canova et al., 1993; Hiratani and Zbib, 2002; Zbib et al., 2011).
- Dislocation pattern formation during monotonic and cyclic loading.
- Dislocation-defect interaction problems, including dislocation-void interaction (Ghoniem and Sun, 1999), dislocation-SFT/void-clusters interaction in irradiated materials and the role of dislocation mechanisms in the formation of localized shear bands (Hiratani et al., 2003; Khraishi and Zbib, 2002a; Zbib et al., 2011).
- The effect of free surfaces on the plasticity of a small material volume (Khraishi and Zbib, 2002b).
- Crack tip plasticity and dislocation-crack interaction (Van der Giessen and Needleman, 2002).
- Size effects in polycrystalline metals (Ohashi et al., 2007).
- The role of various dislocation patterns such geometrically necessary boundaries (GNB’s) in hardening phenomena (Zbib et al., 2004).
• Plastic zone and hardening in Nano-indentation tests (Fivel et al., 1998).
• Examination of abrupt plastic events in nanopillars (Papanikolaou et al., 2017).
• Analysis of rate-dependent friction (Song et al., 2016).
• Indentation and the effect of pre-stress (Song et al., 2019) or indentation size effects (Balint et al., 2006).

A setup of DDD simulation can be seen in Fig. 3 for indentation (Song et al., 2019) where the effect of pre-stress on hardness was examined and in Fig. 4 for abrupt plastic events in thin films (Papanikolaou et al., 2017) where the yield strength size dependence at the nanoscale was examined:

![Figure 3: Spherical Indentation in 2D-DDD: From (Song et al., 2019)](image)

![Figure 4: Histograms of abrupt events and cutoff dependence. From: (Papanikolaou et al., 2017).](image)
3. Discrete Dislocation Dynamics Simulation Data

3.1. Theory and Data Definitions

The 2D DDD model presented in this thesis is a variation of the Van Der Giessen-Needleman model (der Giessen and Needleman, 1995). The geometry of the model problem is shown in Fig. 5 double slip systems. In this work, our primary focus will be on the double slip system samples, but there will also be a comparison with single slip systems. Samples are modeled (Papanikolaou et al., 2017) by a rectangular profile of width \( w \) height \( h \) and aspect ratio \( \alpha = h/w \).

In our DDD simulation, plastic flow occurs by the nucleation and glide of edge dislocations. The typical Burgers vector offcc crystals being \( b = 0.25 \text{ nm} \), and we study thin films for width which varies in powers of 2 from \( w = 0.125 \mu \text{m} \) to \( w = 2 \mu \text{m} \) with \( \alpha = h/w = 4 \). Dislocations exit the sample at the top and bottom edges (\( x = 0, w \)) which are traction free, while loading is displacement controlled (strain control), by prescribing the \( y \)-displacement at the lateral edges (\( y = 0, h \)). The applied strain rate (for both loading and unloading regimes), \( \dot{h}/h = 10^4 \text{ s}^{-1} \), is held constant across all our simulations, similar to experimental practice. In the model, slip planes are separated by \( 10b \) and oriented at \( \pm 30^\circ \) from the loading direction (Fig. 5). In the single slip model, planes are also separated by \( 10b \) but are oriented in just one direction (\( 30^\circ \) from the loading direction).

![Figure 5: The 2D discrete dislocation plasticity model of uniaxial compression of thin films: Slip planes (lines) span the sample, equally spaced at \( d = 10b \), but planes close to corners are deactivated to maintain a smooth loading boundary. Surface and bulk dislocation sources are present (disks) and forest obstacles are spread homogeneously across slip planes. Initially the sample is free of stress and mobile dislocations.](image)

Plastic deformation of the crystalline samples is described using the discrete dislocation framework for small strains (der Giessen and Needleman, 1995), where the determination of the state in the material employs superposition: Each dislocation is treated as a singularity in a linear elastic background solid with Young's modulus \( E \) and Poisson ratio \( \nu \), whose analytic solution is known at any position, and because the sample does not extend to infinity the displacement, strain and stress fields need to be corrected by smooth image fields (denoted by \( ^\wedge \) below) to satisfy boundary conditions at the edges. Hence, the displacements \( u_i \), strains \( \varepsilon_{ij} \), and stresses \( \sigma_{ij} \) are written as:
\[ u_i = \bar{u}_i + \tilde{u}_i, \varepsilon_{ij} = \bar{\varepsilon}_{ij} + \tilde{\varepsilon}_{ij}, \sigma_{ij} = \bar{\sigma}_{ij} + \tilde{\sigma}_{ij}, \]  
(1)

where the \((\bar{\cdot})\) field is the sum of the fields of all \(N\) dislocations in their current positions, i.e.

\[ \bar{u}_i = \sum_{j=1}^{N} \bar{u}_i^{(j)}, \bar{\varepsilon}_{ij} = \sum_{j=1}^{N} \bar{\varepsilon}_{ij}^{(j)}, \bar{\sigma}_{ij} = \sum_{j=1}^{N} \bar{\sigma}_{ij}^{(j)}. \]  
(2)

The image fields are obtained by solving a linear elastic boundary value problem using finite elements (320x80 elements meshing) with the boundary conditions changing as the dislocation structure evolves under load. At the beginning of the calculation, the crystal is stress free and there are no mobile dislocations. This corresponds to a well-annealed sample, yet with pinned dislocation segments left that can act either as dislocation sources or as obstacles.

Dislocations multiplication happens from sources when the resolved shear stress \(\tau\) at the source location is sufficiently high (\(\tau_{nuc}\)) for a sufficiently long time (\(t_{nuc}\)). We consider bulk dislocation sources and obstacles (der Giessen and Needleman, 1995). A dislocation configuration of one of the simulations, at 10% total strain, is shown in Fig. 6 (a) together with the \(xx\) component of the total strain (b) and the shear stress (c).

Bulk sources are randomly distributed over slip planes and locations at a density \(\rho_{nuc}^{bulk} = 60 \mu m^{-2}\), while their strength is selected randomly from a Gaussian distribution with mean value \(\tau_{nuc} = 50 MPa\) and 10\% standard deviation. Bulk sources are designed to mimic the Frank-Read mechanism in two dimensions (Hirth and Lothe, 1982), such that they generate a dipole of dislocations at distance \(L_{nuc}\), when activated. The initial distance between the two dislocations in the dipole is

\[ L_{nuc} = \frac{E}{4\pi(1-\nu^2)\tau_{nuc}} \frac{b}{\tau_{nuc}} \]  
(3)

at which the shear stress of one dislocation acting on the other is balanced by the local shear stress. Surface dislocation sources are successively placed at opposite ends of slip planes, which corresponds to a surface density of around \(\rho_{nuc}^{sur} = 175 / \mu m\). Once a single dislocation is generated from a surface source, it is put at 10\(b\) from the free surface. Under this circumstance, our surface nucleated dislocation has an effective nucleation strength of 312 MPa (Papanikolaou et al., 2017). The model considers only glide of dislocations, neglecting the possibility of climb. In order to find the evolution of dislocations, we solve for the component of the Peach-Koehler force in the slip direction, which for the \(I\) -th dislocation, is given by:
Figure 6: Boundary plastic steps and stress/strain profiles in dislocation configurations: The full sample is shown in undeformed (a, c) and deformed coordinates (b). (a) A dislocation snapshot at 10% strain for a sample of width w=2 and double slip systems. A high density of dislocation is observed, close to $\rho = 10^{15}/m^2$. T’s and $\perp$’s have been used to describe the dislocations that create positive stress (T’s) and negative stress ($\perp$’s) on the different slip systems. (b) The $\epsilon_{xx}$-strain profile that corresponds to the dislocation configuration of (a) is shown in the displaced coordinates. The boundary is filled with rough plastic steps. The dashed rectangular box denotes the region for which we consider spatial strain correlations, that are then used for ML purposes. The strain map is unitless. (c) The shear stress $\tau_{xy}$ of the configuration in (a, b) is shown. The stress map is in MPa.

$$f^{(I)} = \mathbf{n}^{(I)} \cdot \left( \hat{\sigma} + \sum_{j \neq I} \tilde{\sigma}^{(j)} \right) \cdot \mathbf{b}^{(I)}$$  \hspace{1cm} (4)$$

where $\mathbf{n}^{(I)}$ is the slip plane normal and $\mathbf{b}^{(I)}$ is the Burgers vector of dislocation $I$. This force will cause the dislocation $I$ to glide, following over-damped dynamics, with velocity:

$$v^{(I)} = \frac{f^{(I)}}{B}$$  \hspace{1cm} (5)$$
where $B$ is the drag coefficient. In this paper, its value is taken as $B = 10^{-4} Pa \cdot s$, which is representative for aluminum.

Each sample contains a random distribution of forest dislocation obstacles and surface dislocation sources, as well as a random distribution of bulk dislocation sources. Once nucleated, dislocations can either exit the sample through the traction-free sides, annihilate with a dislocation of opposite sign when their mutual distance is less than $6b$, or become pinned at an obstacle. Point obstacles are randomly distributed over slip planes and locations with a constant density that corresponds, on average, to one source, either surface or bulk, for every 8 randomly-distributed obstacles. In this way the densities of sources and obstacles remains the same as the sample dimensions change. We model obstacles in a simple way where a dislocation stays pinned until its Peach-Koehler force exceeds the obstacle-dependent value $\tau_{obs}$. If dislocations approach the physical boundary of the sample then a geometric step is created on the surface along the slip direction (see Figs. 6, 7). The strength of the obstacles $\tau_{obs}$ is taken to be $300 \text{ MPa}$ with 20% standard deviation. Our simulations are carried out for material parameters that are reminiscent of aluminum: $E = 70 \text{ GPa}$, $\nu = 0.33$. The effective Young’s modulus for plane stress problems is $E_{eff} = E / (1 - \nu^2) = 78.55 \text{ GPa}$.

![Strain profiles captured from the 2D-DDD simulation at stage L and stage T -- w = 2 μm -- Double slip system](image)

Figure 7: Strain profiles captured from the 2D-DDD simulation at stage L and stage T -- $w = 2 \mu m$ -- Double slip system: The full sample is shown, in both deformed (a) or undeformed (b) coordinates. A sample is loaded to 10% strain, unloaded to zero stress and reloaded to testing deformation of 0.1%. (a) Strain profile for Stage L at 10% strain. Plastic steps are allowed to emerge on the film surface (Papanikolaou et al., 2017). (b) A strain profile at the Stage T, after subtracting the residual plastic deformation at Stage U. Such strain profiles are analogous to typical DIC experimental strain profiles. The strain maps are unitless.

The simulation is carried out in an incremental manner, using a time step that is a factor 20 smaller than the nucleation time $t_{nuc} = 10 \text{ ns}$. At the beginning of every time increment, nucleation, annihilation, pinning at and release from obstacle sites are evaluated. After updating the dislocation structure, the new stress field in the sample is determined, using the finite element method to solve for the image fields (der Giessen and Needleman, 1995). According to the literature review, the model uses the most important characteristics of dislocation motion in a crystal, allowing us to examine an average scope of crystal plasticity in thin films.
### 3.2. High Throughput Data Generation

The test that we wish to imitate would measure the strain field in the sample after it has been strained and relaxed, as described above, and then subjected to a subsequent “testing” deformation. We consider a testing reload regime that is governed mainly by the degree of invasiveness we introduce to the data set. All tests have been carried out for prior deformation (see Sec 1.2) in three different amplitudes (0.1, 1, 10 %) of total strain. In the case of double-slip systems, we consider 50 random realizations of sources and obstacles in each parameter case (loading of 0.1, 1, 10%) for a total of \( n = 150 \) samples, to obtain a statistically significant dataset (i.e. the strain and stress information across samples show sufficient variance). Moreover, the testing deformation is set to 0.1% for the “small reload” dataset and to 1% for the “large reload” data set for the same realizations, and the total number of samples becomes \( n = 300 \); 150 for each testing deformation. For single-slip systems, we consider 9 random realizations for each parameter case for a total of \( n = 27 \) samples. Similar to the double slip systems, we have a total of \( n = 54 \) samples; 27 for each testing deformation.

Figs. 1, 8 show a schematic of the way we create our data set: “As annealed” samples (see Fig. 8) are loaded to 3 different amplitudes (L stages). For each stage L, we unload (at 0 applied stress) to obtain U stages. In stage U, the samples are stress free, but there is some remaining strain due to plasticity. We then reload the samples to a specific testing deformation (stage T).

![Diagram](image-url)

*Figure 8: A schematic of the sequence of events when loading, unloading and reloading a sample is shown, with the corresponding stress and strain graphs vs the time step of simulation. A sample is obtained from 2D-DDD simulation and it’s loaded to a specific strain value – “Prior” Loaded stage (Stage L). Then, the sample...*
is unloaded to zero stress and the remaining plastic strain can be calculated – “Prior” Unloaded stage (Stage U). Finally, the sample is reloaded to a testing deformation – “Test” Reloaded stage (Stage T). Even though a sample has been plastically deformed (Stage L), the samples obtained from experiments can be polished, thus the surface of a sample is not able to provide information about deformation (Sample at stage U can be seen in the figure as having smooth surface). Such techniques are applicable in experiments such as digital image correlation, where randomly placed tracking nanoparticles are detected optically and contribute to correlation statistics are applied to the sample. Then, as the sample is reloaded, the permanent deformation can be observed, since there are changes in the distances between tracked nanoparticles.

3.3. Strain history and spatial dislocation correlations

By running 2D-DDD simulations, we acquire strain information at the L, U and T stages. In order to remove the prior memory, we form the quantity $T - U = \epsilon_T(\vec{r}) - \epsilon_U(\vec{r})$, or $\epsilon_{TU}$, which is the testing deformation in Fig. 1. This process is similar to polishing a sample, applying speckles and tracking their movement as the sample is further deformed, which is naturally similar to DIC (Kammers and Daly, 2013a, 2013b; Peters et al., 1983; Peters and Ranson, 1982; Schreier et al., 2009; Sutton et al., 2007, 1986, 1983).

After removing the prior strain, it is not straightforward to characterize the plastic behavior of the samples without prior knowledge (i.e. the degree of plasticity incurred from L and T stages). For example, in Fig. 9, without prior knowledge we would not know that the samples in (a), (b) are loaded to 10% strain while in (c), (d) the samples are loaded to 1% strain. The figures appear to be quite different, and the similarity of their histories is not recognizable by eye. However, ML’s trained eye can detect the initial deformation history. Indeed, in later sections we will show how ML algorithms can show that the figures are quantifiably and fundamentally different. With the help of ML, we can find the initial deformation history of various samples, as long as the testing deformation does not overwrite it.

Figure 9: Variety of strain profiles in 2D-DDD simulations for smaller and larger testing strain: For the spatial scale of the figures see Fig. 7. A sample is loaded to a high deformation of strain (which could be either 1 % or 10 %) and then unloaded. This pre-deformed sample is then reloaded to a testing strain. In (a), (b) the samples are loaded to 10% strain, while in (c), (d) the samples are loaded to 1% strain (a) Small testing deformation (0.1%). $w=2 \mu m$. Double slip system simulation. (b) Large testing deformation 1%. $w=1 \mu m$. Single slip system simulation (c) Small
We compute the strain \( \epsilon_{TU} \) on a grid of 2000 \( \times \) 500 points, which is overlaid on the finite element mesh of the DD model (320 \( \times \) 80 nodes, see Sec.3.1), using finite element interpolation to compute the strain at each nodal point. For each node of the interpolated grid, we construct a scalar variable, the determinant of the deviatoric total strain, \( \phi \equiv \left( \frac{1}{2} \epsilon_{xx}^2 + \epsilon_{yy}^2 \right) - \epsilon_{xx} \epsilon_{yy} + 2\epsilon_{xy}^2 \), where the tensor \( \epsilon \) is the total strain. The scalar \( \phi \) can be interpreted as a value on the overlaid grid, to form a strain profile image, such as those shown in Figs. 7b, 9, where the different colors represent different values of \( \phi \). In Sec. 4.5 we discuss other scalar variables that create strain profile images. The strain profiles are then put through a correlation algorithm, following the scheme of the Materials Knowledge System (MKS) (Fast and Kalidindi, 2011).

In the general MKS scheme (Steinmetz et al., 2016), one selects spatially-varying quantity(ies) which characterize the microstructure. The space of all possible values of these quantities is called the local state space, \( \mathcal{H} \), and a point in this space is denoted \( h \). In this study, our quantity is the determinant of the deviatoric total strain invariant (\( \phi \)). In the MKS method, one further considers a “microstructure function”, defined on the product space of the microstructure state variables \( \mathcal{H} \), and physical space \( x, m(h, x) \). In general use, this function may be thought of a probability density on these spaces.

In order to obtain data suitable for constructing two-point correlations (Niezgoda et al., 2008), it is necessary to bin the state variables. We make use of the PyMKS software (Wheeler et al., 2014) which offers tools to accomplish this. The most basic \( h \)-axis discretization scheme is the so-called “primitive basis” scheme, in which one selects some number \( P \) of evenly-spaced levels, \( h_0, h_1, \ldots, h_P \), and, at a point in space where the state variable has value \( h \), selects amplitudes \( \omega_i \) for these levels such that \( \sum_i \omega_i h_i = h \), with the additional restriction that only the \( h_i \)'s such that \( h_{i-1} < h_i < h_{i+1} \) are nonzero, and \( \sum_i \omega_i = 1 \). The entire system is thus described by a set of values \( \omega_i \) in each spatial point \( x \). In our simulations we discretize the state space into 3 different bins, corresponding to 3 local states \( h_0, h_1, \) and \( h_2 \) at low, intermediate, and high local strains. 6 possible correlations are observed for \( (i, j) \in Z \), where space \( Z \) is defined by the values of \( (i, j) = \{(0,0), (0,1), (0,2), (1,1)(1,2), (2,2)\} \) for the 3 local states. A correlation basically describes the probability of having the head and tail of a randomly spaced vector in the microstructure fall in the desired local states.

The two-point correlations (Niezgoda et al., 2008) are:

\[
C^{[k]}[r|l, l'] = \frac{1}{s} \sum_s m[s, l]m[s + r, l']
\]  

(6)

Where \( C^{[k]}[r|l, l'] \) is the conditional probability of finding the local states \( l \) and \( l' \) at a distance and orientation away from each other defined by the vector \( r \), for the
The application of spatial correlations on strain profile images, is done to ensure that any noise associated to the finite element solution of the DDD approach was smoothened out, since correlations are invariant to noise and sizes. This becomes especially important for strain images originating in experiments; Due to resolutions of optical microscopy, as well as the possible changes that may arise in strain measurements due to positioning, the application of correlation functions ensures that these issues will not affect the creation of the dataset.

4. Application of Machine Learning Algorithms on Spatial Strain Correlation Profiles: The Unsupervised Approach

4.1. Preprocessing

After obtaining a set of samples with different plastic deformation histories and calculating statistical correlation functions of each sample, we must show that differences between the correlations can be used to distinguish samples with different prior plastic strain levels by using an ML algorithm. The inputs to our ML algorithm are the correlation functions of the MKS local states discussed in Sec. 3. Having computed the correlation functions, we now wish to see if ML can extract prior histories from them. However, the statistical correlations obtained lie in a high dimensional space that prohibits visualization. Moreover, since the statistical correlation function values differ between samples, the data matrix \( D \) needs to have a mean zero value per feature vector. We employ preprocessing tools that exist in the sklearn model (Pedregosa et al., 2011) to transform the dataset. Our ML workflow will: (1) convert correlation functions to vectors (raster-wise flattening of the correlation functions), (2) find the significant features of the vectors by PCA, and (3) apply a clustering algorithm to identify samples with similar histories.

The correlation functions are evaluated at positions \( r = (r_x, r_y) \) in a box around \((0,0)\). The range of the correlation function can be limited by choosing the size of the box. The size of the box in the present this is a \( 40 \times 40 \) square. By assigning integers \( v \) to each \( r \), we can convert the 2-dimensional set of points to a list, and thus convert the correlation function to a high dimensional vector:
\[ d_{ij} = (C^{[k]}[r_1|h_i h_j], C^{[k]}[r_2|h_i h_j], ..., C^{[k]}[r_q|h_i h_j]) \]  

(7)

Here, \( q \) is the total number of points in the box, \((i, j)\) label the MKS local states from which the correlation function was computed, and \( k \) labels the samples \textit{i.e.} the simulation run. Thus, the Stage T strain from each simulation run has been mapped to a point in a high dimensional space. Our goal is to see if different Stage L strains show up as clusters in this space.

At the end we have a matrix \( D \) in the form \( n \times m \), with \( n \) rows, where \( n \) is the number of statistical samples. Each row contains the vector \( d_{ij} \) which may or may not be truncated (\textit{i.e.} the correlation function range may be limited by the 40 \( \times \) 40 box or not). The matrix \( D \) has \( m \) columns, where \( m \) is the number of spatial correlation instances:

\[ D = \begin{bmatrix}
C^{[1]}[r_1|h_i h_j] & \cdots & C^{[1]}[r_m|h_i h_j] \\
\vdots & \ddots & \vdots \\
C^{[n]}[r_1|h_i h_j] & \cdots & C^{[n]}[r_m|h_i h_j]
\end{bmatrix} \]  

(8)

The rows of \( D \), are the data vectors on which PCA operates, and the resulting principal components are linear combinations of the basis vectors of this set ((Shlens, 2003) also see Sec 2.1.1).

Having projected the correlation data to the PCA-derived subspace, we can address our main objective, to find the prior deformation history of the samples from the strain correlation measurements. By projecting on the PCA-subspace, we can now visually examine any cluster formations and we are able to identify the samples that belong to each cluster, which we use as an input for examining the accuracy of classification algorithm. The set of points belonging to the 3D subspace of PCA (PC1, PC2, PC3) are used as an input to the classification algorithm to examine whether the visually identified clusters are replicated by the classification algorithm.

4.2. Clustering and classification

We use the Continuous k-Nearest Neighbors (CkNN) algorithm (Berry and Sauer, 2016) to classify samples after running PCA on the data set. The CkNN algorithm is a clustering algorithm, with the advantage that the number of clusters is not arbitrarily defined by the user, as in K-Means clustering (Hartigan and Wong, 1979), but is calculated through a distance based approach. In particular, CkNN recognizes samples that are close to each other and calculates the most probable number of clusters for the data set. After the number of clusters has been found, the algorithm classifies the samples like the K-Means approach. The algorithm is an unsupervised method that detects natural clusters within a data set, and our interest in it is the degree to which the natural clusters correspond to the prior deformation (stage L). The input to this algorithm must consist of a set of points, which in our case is the projections of the correlation matrix on the three
principal components. As an output, the algorithm produces the classified samples, based on the cluster to which they belong (see Fig. 10).

The size of the data set (see Sec. 3), as in most classification algorithms, imposes a limitation on the algorithm. The algorithm groups data samples with similar PCA vectors into one cluster. We find that the algorithm works better for larger data sets. This introduces a limitation on classification, especially for single slip systems. The method is considered successful if the samples with different prior loading are grouped into different clusters. Note that the clustering is done in three dimensions using all three principal components. The choice of limiting the clustering in two dimensions is twofold: Firstly, the highest variance of the dataset can be observed in the first three principal components (see Fig. 11), and secondly, for visualization purposes it is enough to show two dimensional plots, which holds true for most PCA maps in the present thesis.

Figure 10: $w = 2\mu m$, 2D projection of PCA results for thin films -- Double slip system: $h_0, h_0$ auto-correlation. (a) Projection of data set on first two principal components. Red blobs denote samples with 0.1\% strain (stage L), blue triangles samples with 1\% strain (stage L) and green squares denote samples with 10\% strain (stage L), respectively. (b) First principal component of PCA, shown in sample coordinates (Fig. 7, Fig. 9, Sec. 3). (c) Second principal component of PCA, shown in sample coordinates (Fig. 7). The colormaps are unitless, showing the intensity the PCA-transformed correlations.
Figure 11: Variance of principal component: It is observed that after PCA, the first principal component retains the highest percentage of variance (~97%) and components 2 and 3, retain about 2% variance.

In Fig. 12 we show a 3D PCA map for material samples of \( w = 1 \mu m \). We observe that the clustering isn’t affected by the 3rd dimension (evident in Fig. 10), and in this case the information provided by PC3 is irrelevant to our results.

Figure 12: \( w = 1\mu m \), 3D projection of PCA results for thin films -- Double slip system: \( h_0, h_0 \) autocorrelation. The colors follow the definition of Fig. 10. 3 different clusters are shown like in Fig. 11. Introducing the 3rd component into the PCA map, does not affect the results.
The results shown on the present thesis, except for results shown in section 4.4, are extracted by applying PCA and the CkNN algorithm to the whole data set. The same PCA and CkNN steps are applied to all simulations. In the remaining sections, we discuss the clustering algorithm results in various situations.

4.3. Examples

4.3.1. Double Slip System

Crystal plasticity can take place in various configurations, depending on the external loading condition and sample geometries. In uniaxial compression, the features of crystal plasticity depend drastically on the particular Schmid factor involved (Hirth and Lothe, 1982). A quite generic case takes place when two active slip systems exist, and this is the particular focus of this thesis. Multiple tests were examined for thin films of various widths $w$. For large $w (> 0.5 \mu m)$, our algorithm correctly clusters and classifies data into 3 different groups, one for each of the prior strain values, which was the main objective of our work.

Fig. 10 (a) shows that clustering is easily observed for $w = 2 \mu m$, where 3 distinct clusters appear in the PCA of the $h_0, h_0$ autocorrelation. It is clear that there is enough cluster separation to reliably classify plastically deformed metals into heavily deformed and less deformed categories. In Fig. 13, we observe the clustering results for samples of $w = 1 \mu m$. We can see that the 3 prior deformation amplitudes have been correctly identified, however the circle and triangle cluster distance has decreased compared to samples of $w = 2 \mu m$. For these larger sized systems, the CkNN algorithm has 100% accuracy, but for smaller sized systems with $w \leq 0.5 \mu m$ the clustering algorithm fails to cluster data points according to their deformation state.

That is evident in Fig. 14 (b), where one can see what a correct clustering and classification would look like for specimens of $w = 0.5 \mu m$. In Fig. 14 (a) one can observe the results after the CkNN algorithm is applied to the data set. The plastic noise fluctuations in the system, as well as the finite size of the system, interferes with the classification of smaller sized data samples, while for larger $w$ the samples are classified correctly.
Figures 10 (b) and (c), as well as 13 (b) and (c) show the representation of the first two principal components of the data matrix $D$ for samples of $w \geq 1 \mu m$, shown in their natural sample coordinates (i.e., the PCA vectors have been converted back to the 2D grid representation of a correlation function, see Sec. 3.3). Note that the first principal component is roughly isotropic, while the second is strongly anisotropic. Figures 10, 13 and 14 show the progression of our ML work-flow as sample width decreases. We can observe that the first PCA component at larger $w$ is relatively isotropic.

While in Fig. 10 (b) we notice a concrete isotropy of the first principal component of the analysis, it gradually becomes anisotropic as the sample width decreases (Fig. 14 (b)). This change is correlated with the onset of stochastic fluctuations at small scales and mechanical annealing (Shan et al., 2008) that promotes concrete slip bands even at small testing strains. While both principal components for $w = 2 \mu m$ (Fig. 10 (b, c)) are smooth, they gradually become less structured as $w$ decreases (Fig. 14 (c, d)), naturally an effect of stochastic fluctuations at small length scales. For $w = 2 \mu m$ there is a distinct difference between the first and second principal components, related to a spatial
symmetry breaking. This distinction disappears as $w$ decreases. For smaller $w$, due to the emerging crystal plasticity size effects (Papanikolaou et al., 2017), the data set is not as distinguishable as we would have wanted with our clustering technique, because of the noise associated with strengthening.

![Figure 14: $w = 0.5 \mu m$ -- 2D projection of PCA results for thin films -- Double slip system: $h_6, h_3$ auto-correlation. The colors follow the definition of Fig. 10. (a) Projection of data set on first two principal components with a clustering algorithm applied to the data set, demonstrating a failure in clustering the various deformation levels. (b) Projection of data set on first two principal components without a clustering algorithm applied to the data set, justifying (a). (c) First principal component of PCA, shown in sample coordinates (Fig. 7). (d) Second principal component of PCA, shown in sample coordinates (Fig. 7). For description of colormaps, see Fig. 10.](image)

The area of the correlations, with respect to the sample area can be calculated by:

$$A_{\text{corr}} = N_x \times N_y \times w^2 / 500^2 \ \mu m^2,$$

where $N_x, N_y$ is the number of nodes in $x, y$ directions respectively. For example, for short range (40 $\times$ 40) correlations: $A_{\text{corr}} = N_x \times N_y \times w^2 / 500^2 \ \mu m^2 = (4/625) \times w^2 \ \mu m^2$. Figure 15 shows how $A_{\text{corr}}$ can influence the results. Differences can be observed in Figs. 15 (d), (e), (f), only with respect to the variance of the projected points. The principal components in (a), (b), (c) have small differences, mostly on their intensity. We deduce that our results do not depend on the examined area of correlations, and in order to reduce computational resources and time, we examine short (40 $\times$ 40) range correlations, from the center of the sample.
One deficiency of our ML work-flow emerged as we examined the results: as $w$ decreases, the distance between the PCA-transformed samples also decreases (see Figs 14, 16). It is known that classification algorithms have an inherent limitation: when the distance between points in one cluster is similar to the distance separating two clusters, then the algorithm has difficulty distinguishing the clusters. Fig. 10 shows that the cluster distances in the PC1 direction are of order of magnitude $10^{-2}$ to $10^{-1}$. For $w \leq 0.5 \mu m$ (see Figs. 14 (a), 16 (a), 16 (c)) the cluster between PCA-transformed samples is on the order of $10^{-3}$ to $10^{-2}$, similar to the distance between the samples, and the data samples cannot be classified correctly. For smaller systems, it is evident that samples with stage $L = 0.1\%$ or $1\%$ strain (red circles and blue triangles, respectively) are so close to each other that the classifier regards them as belonging to the same cluster.

Figure 15: $w = 1 \mu m$ -- The choice of the correlation domain and how it impacts the PCA maps -- Double slip system: $h_0, h_0$ auto-correlation. The colors follow the definition of Fig. 10. (a). Projection of data set on first two principal components. For description of colormaps, see Fig. 10 (a) $40 \times 40$ domain of correlation matrix. Highly smooth in the center and towards the boundaries of the domain. (b) $100 \times 100$ domain of the correlation matrix. A highly focused area near the center of the domain is shown, where the phenomena are focused. The smoothness present in (a) is slowly removed from this domain. (c) $200 \times 200$ domain of correlation matrix. We have rich phenomenology present towards the center of the correlation matrix and at the boundaries. (d) PCA maps for $40 \times 40$ domain. (e) PCA map for $100 \times 100$ domain. The variance of the data has changed, and the projections have shifted. The information provided by (b) does not change the cluster formations but introduces unnecessary information that has shifted the results along the PC1 and PC2 axes. (f). PCA map for $200 \times 200$ domain. The variance of the data has changed even more compared to (e). The distances between the blue and green clusters have increased an order of magnitude compared to (e) and 2 orders of magnitude compared to (d). The information provided by (c) does not affect the clusters that are formed from our algorithm.
Figure 16: $w = 0.25 \mu m, w = 0.125 \mu m$ -- 2D projection of PCA results for thin films -- Double slip system: $h_0, h_0$ auto-correlation. The colors follow the definition of Fig. 10 (a). Figures on the left correspond to classified samples, with colors assigned by the algorithm. Figures on the right are samples with colors assigned by their initial load. (a) $w = 0.25 \mu m$. The clustering algorithm fails and puts all points in a single cluster, corresponding to Stage L at 10% strain. (b) $w = 0.25 \mu m$. Actual representation of initial deformations. (c) $w = 0.125 \mu m$. The clustering algorithm fails and puts all points in a single cluster, corresponding to Stage L at 10%. (d) $w = 0.125 \mu m$. Actual representation of initial deformations.

As mentioned in section 1.2, the methods shown in this thesis are applicable to experimental micrographs from DIC. For this reason, we have considered strong statistical variations in the initial microstructures. Frank-Read sources (see Sec. 3) are distributed randomly with a random nucleation stress. Obstacles (mimicking precipitates) are also distributed randomly, with a random resistance stress. This variability causes strong noise and limited spatio-temporal resolution (as can be seen for samples of $w \leq 0.5 \mu m$). Furthermore, this noise propagates into PCA maps where the variance for samples loaded to 1 or 10% strain is very high. However, these variations do not affect the successful application of the ML work-flow, and this is one of the main findings of this work.

The results from our "large reload" data set, with 1% testing strain, show that delicate handling is required to obtain the desired cluster separation. 1% testing strain does not produce the clear separation obtained with 0.1% testing deformation. As the testing deformation increases so do strain localization features and shear band sizes.
With a shear band spanning the whole specimen, we expect that the statistical correlations differ significantly from the statistical correlations of the “small-reload” data set. That is due to the overall effect of localization, from a structural correlation viewpoint.

High loads lead to strain localization in the form of shear bands, which are inhomogeneous and anisotropic, unlike the low strain plastic response. Our methods pick up the transition between the two responses. Indeed, even in the case of low reload strain, the distance between clusters is small and in smaller systems (Figs. 14,16) the samples are unclassifiable.

**Figure 17:** Large-reload vs. small-reload testing -- Example of PCA projection results for thin films of \( w = 0.5 \mu m \); \( h_0, h_0 \) auto-correlation. The colors follow the definition of Fig. 10 (a). Strain profiles are created from the quantity \( \epsilon_{TU} \) (a) Sample with stage L = 10% strain is unloaded to zero stress and then reloaded to small testing deformation (Stage T = 0.1%). (b) Stage T at small testing deformation (0.1%), without a clustering algorithm applied to data set. Projection on two principal components. Actual representation of the data set, with some mixing of the samples. The clusters have shifted closer to one another but not indistinguishable. (c) Sample with stage L = 10% strain is unloaded to zero stress and then reloaded to large testing deformation (Stage T = 1%). (d) Stage T at large testing deformation (1%), without a clustering algorithm applied to data set. Actual representation of the data set. For the higher testing deformation of 1%, we can see that there is much more mixing of the samples. Reloading to higher strain values, adds plastic memory to the samples, rendering our process inapplicable for these cases. For description of color map see Fig. 7 (b).

Fig. 17 compares the small and large reload testing regimes. Figs.17 (b) and (d) show the results of PCA without a clustering algorithm applied to the data set. From Fig.17 (d) we can see that higher testing deformation renders samples indistinguishable in PCA coordinates. The separation that was present for the low testing deformation (0.1%) is missing for higher values. Figs. 17 (a, c) show the strain profiles captured when the sample is reloaded to low (a) and high (c) testing strains. It is obvious that for higher testing deformation there is much more mixing of the samples and the CkNN algorithm
fails (see Fig. 18 (b)). In Fig. 18, we can see the CkNN classification of samples with \( w = 0.5 \, \mu m \). In 18 (a), we observe the failure of classifying samples with small testing deformation (0.1 %), when the correct PCA projected samples can be seen in Fig. 17 (b). In 18 (b) the results of CkNN classification on samples with large testing deformation are seen. Similarly, the CkNN algorithm fails, which is expected, given the observed mixing in Fig. 17 (d). The final difference of the two regimes can be seen in Fig. 19. For small reload strain the first principal component (a) is nearly isotropic, while it becomes highly anisotropic for larger reload strain (c). This observation extends to other components (e.g., 2nd (b, d)) and is correlated to the emergent anisotropy of strain localization.

Figure 18: Large-reload vs. small-reload testing -- Example of PCA projection results for thin films of \( w = 0.5 \, \mu m \): \( h_0, h_0 \) auto-correlation. Classified Samples. The colors follow the definition of Fig. 10. (a) PCA map for classified samples of small testing deformation (0.1%). The CkNN algorithm fails to find multiple clusters due to the small distance between PCA projected samples. (b) PCA for classified samples of large testing deformation (1%). The CkNN algorithms finds two clusters, recognizing samples of 1% and 10% initial deformation, and ultimately fails.
4.3.2. Single Slip System

In section 3.1, we mentioned that this thesis focuses on double slip systems, but that single slip systems will also be examined. So far, we have shown how emergent shear bands can be observed in our simulations for both systems (Fig. 9) as well as PCA results for double slip (Figs. 10,12-19). PCA results for single slip are consistent with double slip, as shown in Fig. 20. Specifically, Fig. 20 (a) shows results of single slip system simulations for \( w = 2\mu m \). The clustering properties for these larger sized systems are like the properties observed for similar systems for double-slip simulations (Fig. 20 (b)). The PCA results contain distinctly separated clusters. In Fig. 21 we show how principal components in sample coordinates compare between single (b, d) and double (a, c) slip systems. The CkNN algorithm, similar to double slip systems, can find cluster separation for sample with \( w \geq 1 \mu m \).
Figure 20: $w = 2\mu m$ -- 2D projection of PCA results for thin films -- Comparison between single and double-slip system: $h_0, h_0$ auto-correlation. The colors follow the definition of Fig. 10. Projection of data set on first two principal components. (a) Single slip system projection (b) Double slip system projection.

4.4. Accuracy & Validation

In order to consider a machine learning algorithm as successful, there needs to be a quantification of its’ accuracy, and it should also be validated with “unknown” data sets (testing data) which have the same features as the data set the algorithm was designed for (training data). In many cases, the availability of testing data sets is limited, especially when considering a previously unsolved/untested problem and so the whole known data set is split into two uneven parts (80%-20% being the most typical case), and the ML algorithm can be trained on part of the data set and its effectiveness tested on the rest. Other than this subsection, the results shown in the paper are an application of our ML work-flow on the whole data set (for a given $w$) and cannot be used to determine the validity of the classifier.

For validation purposes, we “trained” the algorithm by computing the PCA transformation from a randomly chosen half of the $w = 2\mu m$ samples and applying the CkNN algorithm. Then we applied the PCA transformation to the remaining half of the samples and examined whether they were projected into the correct clusters. The results are shown in Fig. 22. It is evident that the testing data (stars) perfectly matches the training set. Similarly, “training” the algorithm to samples of various sizes (i.e., half of the samples instead of all the samples), follows the results of section 4.1.1. For samples with $w \geq 1\mu m$ the “testing” data set is projected to the 3 classified clusters that have formed. In contrast, for smaller systems, the training data set is misclassified (as happens when examining the whole data set) and the testing data set falls within the misclassified results.
Figure 21: $w = 2 \mu m$ – 2D projection of PCA results for thin films – Comparison of principal components among double and single slip systems: Components shown correspond to the analysis of Fig. 18. (a) First principal component of double slip system. (b) First principal component of single slip system. (c) Second principal component of double slip system. (d) Second principal component of single slip system. For description of colormaps, see Fig. 10.

Figure 22: $w = 2 \mu m$ – 2D projection of PCA results for thin films – Double slip system – Validation: $h_0, h_0$ auto-correlation. Red blobs denote samples with 0.1 % strain (stage L), blue triangles samples with 1 % strain (stage L) and green squares denote samples with 10 % strain (stage L), respectively. Red stars depict testing samples of 0.1 % strain (stage L), blue stars testing samples of 1 % strain (stage L) and green stars testing samples of 10 % strain (stage L). Validated-split data set. Projection on first two principal components.
For the quantification of the degradation of the clustering process we use some of the tools provided in the scikit-learn metrics module (Pedregosa et al., 2011). We examine the accuracy score of the algorithm, as well as the $F_\beta$ score. Accuracy is the fraction of samples that were classified correctly. We apply the CkNN algorithm and generate clusters. Because we know the prior strain for each sample, we can immediately check whether the clusters correspond to the strain levels. Perfect clustering is when each cluster contains only samples with identical prior strains. The results are summarized in Fig. 23. For $w \geq 1 \mu m$ the accuracy score is 1 as seen in Fig. 23 (a); that is, all the samples are correctly classified. For smaller samples $w \leq 0.5 \mu m$ (or $w/w_0 \leq 2^2$ as in the figure), we have a 0.33 accuracy score, because only the samples of one cluster are correctly classified. The accuracy score is not affected by the wrongly classified samples and cannot provide a measure for the correct classification of individual clusters.

To quantify the performance of the classification process, we also use the $F_\beta$ score (Baeza-Yates and Ribeiro-Neto, 2011; Powers, 2011) which is computed separately for each cluster:

$$F_\beta = (1 + \beta^2) \cdot \frac{p \cdot r}{\beta^2 \cdot p + r}$$

where precision $p$ is the number of correctly classified samples in the cluster divided by the number of all classified samples in the same cluster, and recall $r$ is the number of correctly classified samples in the cluster divided by the number of samples that should have been in that cluster. The $\beta$ number changes the weight of recall vs precision. For $\beta > 1$ recall is weighted more than precision, while for $\beta < 1$ precision is weighted more than recall. For $\beta = 1$, we have the $F_1$-score, with precision and recall having the same weight in the equation. Fig. 23 (b, c, d) shows the $F_1, F_2$ and $F_{0.5}$ scores for our results.
Figure 23: Measures of success for classification of samples – 0.1 % testing deformation: $h_0, h_0$ auto-correlations. The x-axis of each graph is the base 2 logarithm of the various sample widths we examine. (a) Accuracy score for the samples. Maximum value 1 means that all the samples have been correctly classified. (b) $F_1$-score of our 3 clusters that are formed. The line with the squares represents the cluster with samples at stage $L = 10$ % strain, while the line with the triangles is for the cluster with samples at stage $L = 1$ % strain. Finally, the line with the circles is for the cluster with samples at stage $L = 0.1$ % strain. For smaller sized systems we have observed that most of the samples are classified as belonging in the “square” cluster, hence the scored value for that cluster only. Since the algorithm correctly classifies the samples that were initially loaded to 10 % strain, but also classifies more samples as belonging to that cluster, then the score does not have the maximum value of 1 but lower. (c) $F_2$-score of our 3 clusters that have formed. The definition of the colored lines follows (b). Since for $F_2$-score we have increased weight of the recall, the 0.7 maximum value is expected for the “square” cluster. (d) $F_{0.5}$-score of our 3 clusters. The colors definitions follow (b). Since we have reduced weight of the precision, for lower sample widths it is expected to have lower score than $F_1$ for the “square” cluster.

For samples with $w \geq 0.5 \mu m$ (or $w/w_0 \geq 2^2$) we have value of 1 on all scores and all clusters, but for smaller $w$ we observe that the line with the squares, which corresponds to samples with 10% initial compressive loading returns non-zero values, varying as the $\beta$ value changes. For samples that are classified in the cluster, we do not obtain the highest possible result, because the number of correctly classified samples is smaller than the number of samples in the cluster (i.e., the precision is small). The line with the circles, which corresponds to samples with 0.1 % initial strain loading, has value 0 for $w \leq 0.5 \mu m$ because no samples have been classified as belonging to that cluster. The last line, with the triangles corresponding to samples with 1 % initial loading has non-
trivial values because in some cases there are some samples that are classified correctly (the recall and precision are very small). In summation: For the “square” cluster we have low precision but high recall, since we classify the samples that actually belong to that cluster correctly, but we also classify samples from other clusters; for the “triangle cluster” we have low recall and low precision, since we classify a small number of samples into that cluster.

We also tested the response of the algorithm with respect to accuracy and $F_1$-score while changing the number of tested samples. Fig. 24 shows the algorithm’s reduced effectiveness when the number of samples is less than 20% of our maximum. Fig. 24 (b) shows the average $F_1$-score across the three clusters instead of the score for each cluster individually.

![Figure 24: Measures of success for classification of samples – 0.1 % testing strain: $h_0, h_0$ autocorrelations. The x-axis of each graph is percentage of samples tested for classification. (a) Accuracy score for samples of $w = 2 \mu m$ (stars) and $w = 1 \mu m$ (disks). Maximum value 1 means that all the samples have been correctly classified. (b) Averaged $F_1$-score across the 3 clusters that have formed for samples of $w = 2 \mu m$ (stars) and $w = 1 \mu m$ (disks). It is obvious that we have good agreement for the classified samples even when we test less than 30 % of the total number of samples.](image)

4.5. Uncertainty sensitivity and quantification

Up to this point, the thesis focused on a standard set of parameters: 1) Characterization of microstructure by the $\phi$-scalar/ determinant of the deviatoric total strain, 2) Autocorrelations $h_0, h_0$ and 3) Dimensionality reduction by PCA. This set of parameters can change: For example, we can choose to use the $\phi$-scalar for the plastic strain instead of the total strain, or use another scalar, such as the 2$\text{nd}$ invariant of the strain tensor: $J_2 = \epsilon_{ik}\epsilon_{ki}$. Another change could be in the correlations ($h_0, h_1$ cross correlations or $h_1, h_1$ autocorrelations), and finally, we could choose to use other dimensionality reduction methods such as Incremental PCA (IPCA), which splits the data set into smaller batches and then transforms it as PCA. In IPCA, since features are left out from the dataset, the results may differ. In the following, we discuss some of these
changes for the samples with $w \geq 1 \mu m$, since those samples are reliably classified by CkNN.

In Figure 25, we show how our classification scheme produces similar results if we use the more common 2nd invariant of the strain deformation tensor, $J_2 = \epsilon_{ik} \epsilon_{kl}$, for $h_0, h_0$ (Fig. 25 (a)) and $h_1, h_1$ (Fig. 25 (b)) autocorrelations. The only change that can be observed is the change in variance of the PCA projected data.

![Image of Figure 25](image)

**Figure 25:** Effect of strain invariant type for pre-processing -- Examples of PCA projection maps -- $J_2$ strain tensor invariant: The colors follow the definition of Fig. 10. (a) $h_0, h_0$ autocorrelations, $w = 2 \mu m$. (b) $h_1, h_1$ autocorrelations, $w = 1 \mu m$.

In Figure 26, we show calculations with the plastic strain instead of the total strain as the value that is inputted in the scalar $\phi$. Fig. 26 shows that classification still works and there is an observable difference of the data variance in PCA coordinates. In Figure 27, we show results with a combination of plastic strain and the $J_2$ invariant of the strain tensor, and the results do not change, apart from the variance.

![Image of Figure 26](image)

**Figure 26:** Residual/Plastic strain for pre-processing: The colors follow the definition of Fig. 10. $h_0, h_0$ autocorrelations, $w = 2 \mu m$. 
Figure 27: Plastic strain calculations for $J_2$ and $\phi$ strain-tensor invariants – Examples of PCA projection maps: The colors follow the definition of Fig. 10. (a) $h_0, h_0$ autocorrelations, $w = 2 \mu m$. $J_2$ invariant. (b) $h_0, h_0$ autocorrelations, $w = 2 \mu m$. $\phi$ invariant. (a) $h_0, h_0$ autocorrelations, $w = 1 \mu m$. $\phi$ invariant. (a) $h_0, h_0$ autocorrelations, $w = 1 \mu m$. $J_2$ invariant.

Finally, an IPCA calculation is shown in Figure 28. The PCA projections are mirrors of the PCA calculation, but the CkNN algorithm correctly classifies the prior deformation histories of samples with $w \geq 1 \mu m$.

Figure 28: $w = 2 \mu m$, 2D projection of IPCA results for thin films -- Double slip system: $h_0, h_0$ autocorrelation. The colors follow the definition of Fig. 10.
5. Application of Machine Learning Algorithms on Spatial Strain Correlation Profiles: The Supervised Approach

In this section, we briefly discuss the aspects of this work if we had a supervised problem \textit{(i.e. the classification of the samples is already known)}. Supervised ML algorithms, as mentioned in Sec. 2.1, are trained on \textit{training} sets and their accuracy is tested on \textit{testing} sets, like the validation example shown in Fig.22 (Sec 4.4). Given that we do not have access to experimental data, we split the dataset of \textit{known} deformation histories into an 80% training set and 20% testing set. For this problem, the dataset now consists of the strain correlation matrices \( D \), for samples of varying width.

We train two types of supervised ML algorithms on the training set: Neural networks (Bishop et al., 1995) and Decision Trees (Quinlan, 1986). It is out of the scope of this thesis to discuss the particulars of the aforementioned ML algorithms, however Neural Networks are a set of algorithms, modeled loosely after the human brain, that are designed to recognize patterns in datasets and consist of \textit{neurons} from which the dataset passes through and activate through the input. Decision trees are a set of \textit{decisions} for the features of the input matrix, modeled after trees. The algorithm finds patterns in the features and creates leaves of a tree. When all possible patterns have been found, we have multiple leaves of a tree, hence the name decision tree. For the supervised problem, the result is the identification of the best ML algorithm by comparing the accuracy scores.

In the following tables, we show the results per width of the most accurate Neural Networks (Table 1) and Decision Trees (Table 2) algorithms applied on the dataset. The most accurate Neural Networks and Decision Trees are found through a parameter search using an algorithm for parameter optimization (GridSearchCV (Bergstra et al., 2011)). The GridSearchCV algorithm, allows as inputs multiple parameters of a given classifier (i.e. Neural Networks, Decision Tress in our case), and outputs the set of parameters that will provide the highest accuracy. In this thesis, the parameters we used in the GridSearchCV algorithm for Neural Networks were: Adaptive or constant learning rate, ranging from \( 10^{-5} \) to \( 10^{3} \). For Decision Trees, the parameters inputted on the GridSearchCV algorithm were: gini or entropy criteria with the maximum depth of the tree ranging from 12 to 16. With these parameters the highest accuracy was provided for adaptive learning rate of \( 10^{-5} \) for Neural Networks, while for Decision Trees, the best criterion was gini (Quinlan, 1986) with maximum depth (for \( w = 0.125 \, \mu m \)) set at 14 leaves.

The scores for the supervised problem, exceed the scores of the unsupervised problem: A classification algorithm is not needed to classify the prior deformation histories; In the supervised problem, the deformation histories are known for the training
set. If we had access to experimental micrographs, we would have only needed to calculate the spatial strain correlations of the experimental data, and then project the corresponding matrices, on the Decision Tree or Neural network classifier of known prior deformation histories, and immediately have as a result the deformation history of the experimental dataset.

<table>
<thead>
<tr>
<th>Width</th>
<th>Training Set Accuracy</th>
<th>Testing Set Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125 μm</td>
<td>90.38 %</td>
<td>83.3 %</td>
</tr>
<tr>
<td>0.25 μm</td>
<td>91.35 %</td>
<td>100 %</td>
</tr>
<tr>
<td>0.5 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>1 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>2 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Table 1: Neural Networks accuracy scores on supervised machine learning for identification of prior deformation histories via spatial strain correlations.

<table>
<thead>
<tr>
<th>Width</th>
<th>Training Set Accuracy</th>
<th>Testing Set Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.125 μm</td>
<td>100 %</td>
<td>83.3 %</td>
</tr>
<tr>
<td>0.25 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>0.5 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>1 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
<tr>
<td>2 μm</td>
<td>100 %</td>
<td>100 %</td>
</tr>
</tbody>
</table>

Table 2: Decision Tree accuracy scores on supervised machine learning for identification of prior deformation histories via spatial strain correlations.

6. Prediction of mechanical responses and material properties

As we discussed in the introduction, and shown in this thesis, the mechanical response of a specimen, is largely controlled by the prior deformation. Moreover, we briefly discussed the application of supervised approaches (Sec. 5) on the dataset presented in Sec. 3, for the small testing deformation. In other words, Sec. 5 shows that supervised algorithms are able to find a relationship between the known prior deformation histories (3 classes of uniaxial compressive strain) and spatial strain correlations in training samples and use it for the classification of testing samples with high accuracy. The result is that we have a way to classify the prior deformation history, and as such there should also be a way to predict the mechanical behavior of samples. We assume that the mapping of training samples is an average behavior pattern of the samples, or that the samples which belong to each class are “similar” in a mechanical sense. In this section, we use the classified samples as averages for the prediction of the mechanical response upon further compression. In Fig.29, a schematic for the prediction of the mechanical response is shown, and we discuss the detailed process of calculating the average response based on prior deformation.
Figure 29: Schematic for obtaining the 1% strain mechanical response of unknown samples: A new test sample is mechanically loaded to 0.1 % strain, spatial strain correlations are captured and used to find the classification of the sample based on the previously tested samples (whose classification is known by Supervised ML algorithms). Depending on the exported classification, the future mechanical response at 1% can be predicted, and is the average mechanical response of the previously tested samples of the same class.

Since the classified samples are thought of as “similar”, we create 3 separate datasets, one for each deformation class (irrespective of the accuracy of the algorithm). For samples in each class, we also know the future deformation (1 % testing deformation), since there is a one-to-one correspondence between testing deformation levels (i.e. the same sample that is loaded to 0.1% testing deformation to capture the strain correlation patterns is also loaded to 1% testing deformation). While the classification of the “large reload” dataset is not possible, that does not matter; the samples share the same deformation history, which can only be found for small reload strain (see Sec 4.3). For each dataset, we collect the average reload response (1 % strain) per width. The results can be seen in Figs. 30-35.

In Figs. 30-34, we observe, for decreasing width, whether the prior deformation history controls the mechanical response and the hardening behavior of the material. In particular, we see a transition for samples of $w \geq 1 \mu m$ and $w \leq 0.5 \mu m$ where the maximum stress response in further deformation, changes from being high deformation history dominated (prior deformation history = 10%, Figs. 30, 31) to low deformation history dominated (prior deformation history = 0.1%, Figs. 32-34). However, the problem is one of crystal plasticity, and the mechanical response is described by the movement and multiplication of dislocations inside the crystal. Accordingly, high deformation history is high dislocation density in the crystal, which implies that we have an accurate method to describe and predict non-equilibria mechanical phenomena: Given a sample of
unknown deformation history, and a known database of prior deformation histories (as provided in Sec. 5), one only needs to apply a small load mechanical test, capture spatial strain correlation features, use them as part of the test set in the supervised ML problem and obtain a prediction of future mechanical response, and the prior deformation history/dislocation density of the crystal.

This transition becomes even more clear in Fig. 35, where the maximum stress of the predicted response is plotted against the thin film width. Furthermore, fitting functions are used to describe the relationship of stress and width: $\sigma = w^{-a}$, where $a$ is shown in the legend, taking values previously shown in the literature (Papanikolaou et al., 2017; Song et al., 2018), but an explanation as to the variance has not yet been accurately provided. The results and the discussion of this section however, provide an explanation as to the spread; There is a dependence on the prior deformation history.

![Figure 30: Prediction of the average mechanical response of 2 µm samples with a known prior deformation history](image)

*Figure 30: Prediction of the average mechanical response of 2 µm samples with a known prior deformation history*
Figure 31: Prediction of the average mechanical response of 1 $\mu$m samples with a known prior deformation history.

Figure 32: Prediction of the average mechanical response of 0.5 $\mu$m samples with a known prior deformation history.
Figure 33: Prediction of the average mechanical response of 0.25 μm samples with a known prior deformation history.

Figure 34: Prediction of the average mechanical response of 0.125 μm samples with a known prior deformation history.
7. Ongoing and future work

In Sec. 5 we briefly showed results of the supervised problem, for 3 classes of deformation material history. Material scientists rarely deal with a distinct set of deformed samples \( (i.e. \ 0.1, 1 \ and \ 10\% \ strain) \), and specimens are normally deformed over a continuous range of elastoplastic responses. Future projects on thin films, as an extension to the present thesis, deal with a continuous deformation range \( (0.1 \ - \ 10\% \ strain, \ for \ 0.1\% \ steps \ of \ strain) \). We have collected the required samples from 2D-DDD simulations, and we plan on calculating the spatial strain correlations of each sample. When the dataset is complete, the goal is to apply the supervised ML algorithms shown in Sec. 5, to the continuous set of deformation. The result would be a database of known strain correlations, corresponding to specific prior deformation histories – any experimental data set of thin films \( (Al) \) would be validated with the high-throughput ML approach we demonstrated. Moreover, the work of Sec. 6 would be further validated and accurately described.

Applications of spatial correlations need not be limited to thin films (or strain information); If surface information is available, spatial correlations can be calculated, and other material properties can be found. For example, an ongoing project of supervised ML, is focused on nanoindentation. 3D-DDD simulations are done for single crystal Cu, of varying orientations: \( \{100\}, \{011\}, \{111\} \). The supervised ML aspect of this project is to identify the crystal orientation of the samples, based on the spatial correlations of the top
surface deformation field (experimentally accessible through AFM (Rugar and Hansma, 1990). The datasets consist of initial dislocation densities $\rho = 1.0e + 11 \, [m^{-2}]$, $\rho = 1.0e + 12 \, [m^{-2}]$ and $\rho = 5.0e + 12 \, [m^{-2}]$. Preliminary results are shown in Figs. 35-37, for increasing dislocation density, with the application of a Support Vector Machine (SVM, (Hearst et al., 1998)) classifier for supervised learning problems.

In Fig. 36, samples of initial dislocation density $\rho = 1.0e + 11 \, [m^{-2}]$, indented at 15 nm, are classified in 3 crystal orientations (called 1,2,3 in the legend). Samples are classified as having orientation 1,2 or 3, or \{100\}, \{111\}, \{011\} respectively. We can also see that the test data projected on the PCA derived subspace, fall within the specified domains of the test data. The scores for the SVM classifier are of 100% accuracy for dislocation densities $\rho = 1.0e + 11 \, [m^{-2}]$ (Fig. 36), $\rho = 1.0e + 12 \, [m^{-2}]$ (Fig. 37). The accuracy for the higher dislocation density $\rho = 5.0e + 12 \, [m^{-2}]$ is a little lower; As seen in Fig. 38, a sample of \{011\} orientation is misclassified as having crystal orientation \{111\}. It is obvious that a lot of information can be extracted for even this small data-set (10 samples/orientation/density). The ML work on nanoindentation data is set to be submitted in 2019.

![Figure 36: Supervised learning of nanoindentation samples with $\rho = 1.0e + 11 \, [m^{-2}]$](image)

An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples, with 100% accuracy.
Figure 37: Supervised learning of nanoindentation samples with $\rho = 1.0e + 12 \text{ [m}^{-2}\text{]}$: An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples, with 100% accuracy.

Figure 38: Supervised learning of nanoindentation samples with $\rho = 5.0e + 12 \text{ [m}^{-2}\text{]}$: An SVM classifier is applied on a deformation field spatial correlations dataset from samples with \{100\}, \{111\} and \{011\} crystal orientations. The SVM classifier identifies 3 different orientations -- (1,2,3) correctly corresponding to the crystal orientations of the samples. A misclassification of a sample is seen, and the expected accuracy is lowered.
8. Conclusion

Our results could be generalized in several ways. Our work is applicable to thin films (Nicola et al., 2006; Papanikolaou et al., 2017; Song et al., 2016), but may also apply to more general families of materials. Firstly, the material parameters can change to correspond to composites and/or polymers, and their associated mechanical behavior during testing. For composites it is important to model and study ductile fracture while for polymers creep phenomena might be of interest. Secondly, an expansion can be made to the experimental protocol. Instead of examining uniaxial compression of thin films and their spatially resolved strain correlations, we could have included multi-cycle loading-unloading tests, multiaxial compression or nanoindentation. Thirdly, the data matrix \( D \) (see Eq. 8, Sec. 4.1) can be defined in different ways. While the protocol would have remained the same within the algorithm, we could have used geometrically necessary dislocations or local misorientations to calculate spatial correlations. In this case, the required data for the correlations would have been obtained, for example, by Electron Back-Scatter Diffraction (EBSD). In future studies, we will examine data from theoretical solutions and aim to compare them with experimental data sets for dislocation-density related problems. A natural next step in our approach is a development of a regression method which can provide a continuous assessment of clustering and classification, and naturally provide error bars. Instead of using only 3 values of the applied strain at Stage L, we can use a continuous set of values, and apply regression-based methods (e.g. decision trees (Quinlan, 1986)) to identify features at each load. Moreover, accessing a continuous set of values of applied strain, will help elucidate the predictive capabilities of our approach for the mechanical response of unknown samples (see Sec. 6), and well define boundaries for prior deformation history which influence future mechanical responses.

In addition, there are some caveats of the approach that one has to be careful with: When samples used for unsupervised ML have either been reloaded to high strain (1%) or exhibit large noise due to their nano size \((w \leq 0.5 \, \mu m)\), our classification method does not work. There are many possible reasons that the algorithm occasionally fails to identify these samples. For example, in the case of smaller \(w\), short-range correlations may not be enough to distinguish the deformation history. Moreover, we use a simple ML workflow, that may not distinguish features of the data matrix \( D \) (see Sec. 4.1, 4.2). Advanced ML protocols such as neural networks (Bishop et al., 1995) or deep learning algorithms (LeCun et al., 2015) could capture more information than correlation based approaches, which is briefly discussed in a supervised setting in Sec.5. The occasional failure of our unsupervised methodology to distinguish prior deformation could also stem from basic aspects of the physical phenomenon of crystal plasticity at small scales: The data shows a substantial amount of noise at smaller widths (Nicola et al., 2006; Papanikolaou et al., 2017; Song et al., 2016) making classification occasionally unsuccessful (see Sec. 4.3.1),
and at larger reload strain (see Sec. 4.3.1) the prior deformation history is overwritten (Asaro and Lubarda, 2006) and becomes undetectable by the algorithm.

In summary, we examined the applicability of spatial correlations to practical and relatively inexpensive experimental methods for the detection of the degree of prior plastic deformation of thin films, especially when they display significant plasticity size effects. Our overall conclusion is that ML algorithms can achieve our objective with varying levels of success. Through mimicking experimental protocols with two-dimensional discrete dislocation plasticity simulations, we identified realistic cases (single and double slip thin films with widths larger than 1 \( \mu m \)) where data clustering and classification is possible, based on the degree of prior plastic deformation. When size effects come into play, we found that clustering and classification becomes gradually more difficult, since the intrinsic, plasticity-induced crackling noise causes large variance in smaller systems. In general, for the success of our unsupervised methodology for thin films, the physical size of the samples should exceed 500 \( nm \) in the lateral direction (see Sec. 4.3), while the data set should consist of more than 50 samples (See Sec. 4.4).

Furthermore, we uncovered a crucial parameter for the applicability of our methods, namely the testing total strain during reloading. The stage T reload strain should be small enough that it does not overwrite the prior deformation history of the samples; Reload strains less than 0.4% could be applicable for detecting deformation history. While for a small-reload level of 0.1% (half of the commonly defined engineering yield stress, found at engineering strain 0.2%), our methods are highly successful (see Sec. 4.3.1, 4.4, 4.5), they are clearly not successful one order of magnitude higher, at 1% (see Sec. 4.3.1).

The main part of this work is set to appear on Physical Review E. (Papanikolaou et al., n.d.).
Bibliography


