

**Artificial Intelligence in Materials Science:
Applications of Machine Learning to Extraction of Physically
Meaningful Information from Atomic Resolution Microscopy Imaging**

**A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**Artem Borisovich Maksov
December 2018**

Copyright © 2018 by Artem B. Maksov
All rights reserved.

DEDICATION

To my family and my friends.

ACKNOWLEDGEMENTS

This work would not have been possible without the guidance and support I received from my advisors, colleagues, and friends at the Oak Ridge National Laboratory and the University of Tennessee. I would like to thank:

Bobby Sumpter, for believing in me and taking me in as a graduate student, for giving me the opportunities to succeed, and for supporting me throughout all the difficult times I encountered. I truly appreciate the ability to pursue my evolving interests in research.

Dr. Sergei Kalinin, for guiding me on the path of learning about Artificial Intelligence and academic research, for kindling my desire to succeed, for his patience, and for being a good friend.

Maxim Ziatdinov, for working closely with me to understand the results of our research, for exchanging ideas at odd hours leading to new discoveries, and for always looking out for new ideas and methods.

Ondrej Dyck, Rama Vasudevan, Stephen Jesse, Ilia Ivanov, Nouamane Iannait, Panchapakesan Ganesh, Miguel Fuentes-Cabrera, Peter Maksymovich, Anton Ievlev, for working together on various projects, whether those were successful or not, for sharing their knowledge with me, and for supporting my pursuit of the degree.

Hairong Qi, for serving on my committee, and for teaching me foundations of computer vision. I am grateful for your time and discussions.

Shaun Gleason, for serving on my committee, and for supporting the direction of my research.

Lee Riedinger, for accepting me into the Bredesen Center, and for ensuring my success.

Michael Barnes, my first research mentor, at University of Massachusetts, for inspiring me to pursue graduate degree, for teaching me how to do scientific research, and for making me aware of the existence of the Bredesen Center.

My family, for their continuous love and support: Pearl Howell, for supporting me even while working on her own graduate degree; my parents, Boris Maksov and Marina Arzumanova, and my grandmother, Aza Arzumanova, for giving me the strength to go on.

ABSTRACT

Materials science is the cornerstone for technological development of the modern world that has been largely shaped by the advances in fabrication of semiconductor materials and devices. However, the Moore's Law is expected to stop by 2025 due to reaching the limits of traditional transistor scaling. However, the classical approach has shown to be unable to keep up with the needs of materials manufacturing, requiring more than 20 years to move a material from discovery to market. To adapt materials fabrication to the needs of the 21st century, it is necessary to develop methods for much faster processing of experimental data and connecting the results to theory, with feedback flow in both directions. However, state-of-the-art analysis remains selective and manual, prone to human error and unable to handle large quantities of data generated by modern equipment. Recent advances in scanning transmission electron and scanning tunneling microscopies have allowed imaging and manipulation of materials on the atomic level, and these capabilities require development of automated, robust, reproducible methods.

Artificial intelligence and machine learning have dealt with similar issues in applications to image and speech recognition, autonomous vehicles, and other projects that are beginning to change the world around us. However, materials science faces significant challenges preventing direct application of the such models without taking physical constraints and domain expertise into account.

Atomic resolution imaging can generate data that can lead to better understanding of materials and their properties through using artificial intelligence methods. Machine learning, in particular combinations of deep learning and probabilistic modeling, can learn to recognize physical features in imaging, making this process automated and speeding up characterization. By incorporating the knowledge from theory and simulations with such frameworks, it is possible to create the foundation for the automated atomic scale manufacturing.

PREFACE

I not only use all the brains that I have, but all that I can borrow.

-Woodrow Wilson

TABLE OF CONTENTS

Chapter 1. INTRODUCTION.....	1
1.1 Materials Science and Technological Progress.....	2
1.2 Artificial Intelligence	3
1.2.1 History of AI.....	3
1.2.2 Machine Learning	4
1.2.3 Deep Learning.....	5
1.3 Atomic scale microscopies and their importance for Materials Science	5
1.3.1 Scanning transmission electron microscopy (STEM).....	6
1.3.2 Scanning tunneling microscopy and spectroscopy (STM/S)	6
1.4 The Role of Artificial Intelligence and Deep Learning in Materials Science.....	7
1.4.1 Previous applications	7
1.4.2 The goals of AI methods application to microscopy data.....	10
1.4.3 Current challenges of DL application for microscopy data	10
Chapter 2. STRUCTURE-PROPERTY RELATIONSHIP FROM STATISTICS.....	12
2.1 Introduction.....	13
2.2 Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl ₃	14
2.2.1 Publication information	14
2.2.2 My contribution	14
2.2.3 Abstract	15
2.2.4 Introduction.....	16
2.2.5 Results.....	18
2.2.6 Discussion	31
2.2.7 Methods.....	32
2.3 Knowledge Extraction from Atomically Resolved Images	34
2.3.1 Publication Information	34
2.3.2 My contribution	34
2.3.3 Abstract	35
2.3.4 Introduction.....	35
2.3.5 Results and Discussion	39
2.3.7 Methods.....	50
2.4 Deep data mining in a real space: separation of intertwined electronic responses in a lightly doped BaFe ₂ As ₂	51
2.4.1 Publication Information	51
2.4.2 My contribution	51
2.4.3 Abstract	52
2.4.4 Introduction.....	52
2.4.5 Results and discussion	54
2.4.6 Conclusions and outlook.....	66
2.4.7 Methods.....	67
Appendix.....	68

Appendix 2-1. Supplementary information for Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl ₃	68
Appendix 2-2. Supplementary information for Knowledge Extraction from Atomically Resolved Images	77
Appendix 2-3. Supplementary information for Deep data mining in a real space: Separation of intertwined electronic responses in a lightly-doped BaFe ₂ As ₂	80
Chapter 3. TEACHING PHYSICS TO DEEP LEARNING MODELS	104
3.1 Introduction.....	105
3.2 Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations.	106
3.2.1 Publication information	106
3.2.2 My contribution	106
3.2.3 Abstract	107
3.2.4 Introduction.....	107
3.2.5 Results and Discussion	110
3.2.6 Conclusions.....	126
3.2.7 Methods.....	127
3.3 Deep analytics of atomically-resolved images: manifest and latent features.	129
3.3.1 Publication Information	129
3.3.2 My contribution	129
3.3.3 Abstract	130
3.3.4 Introduction.....	131
3.3.5 Results and Discussion	133
3.3.6 Conclusions.....	147
3.3.7 Methods.....	147
3.4. Learning surface molecular structures via machine vision.	149
3.4.1. Publication Information	149
3.4.2. My contribution	149
3.4.3. Abstract	151
3.4.4. Introduction.....	152
3.4.5 Results.....	153
3.4.6 Discussion	168
3.5 Identifying phase transitions in first-principles simulations of relaxor ferroelectrics via variational autoencoders	171
3.5.1 Introduction.....	171
3.5.2 Results.....	172
Appendix.....	176
Appendix 3-1. Supplementary Information for Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations	176
Appendix 3-2. Supplementary Information for Learning surface molecular structures via machine vision	176
Chapter 4. TOWARDS ATOMIC SCALE MANUFACTURING	189
4.1 Introduction.....	190

4.2 Direct atomic fabrication and dopant positioning in Si using electron beams with active real time image-based feedback	191
4.2.1. Publication Information	191
4.2.2. My contribution	191
4.2.3. Abstract	192
4.2.4. Introduction	193
4.2.5. Electron matter interactions in STEM	195
4.2.6.. Directed crystallization	200
4.2.7. Direct growth of doped silicon	202
4.2.8. Modeling electron beam-induced transformations	204
4.2.9. Perspectives	208
4.2.10. Materials and methods	209
4.3. Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS ₂	211
4.3.1 Publication information	211
4.3.2 My contribution	211
4.3.3. Abstract	212
4.3.4. Introduction	213
4.3.5. Results and Discussion	215
4.3.6. Conclusions and Outlook	227
4.3.7. Methods	228
Appendix	230
Appendix 4-1. Supplementary Information for Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS ₂	230
Chapter 5. CONCLUSION	234
5.1 Summary	235
5.2 Future outlook	237
5.2.1 Automating electron microscope for atomic scale manufacturing	237
5.2.2 Quantifying the uncertainty of algorithms.	237
5.2.3 Expanding machine understanding from image to material.	238
5.3 AI in Materials Science	238
REFERENCES	239
VITA	264

LIST OF TABLES

Table 2-1. The structure parameters of α -RuCl ₃ measured at 200 K by single crystal neutron diffraction. The space group is $P3_1$, $a=5.956$ Å; $c=17.054$ Å, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$. U has units of Å ²	76
Table 3-1. Assignment of pairwise potentials based on our knowledge about the molecular system	160
Table 3-2. Assignment of pairwise potentials based on our knowledge about the molecular system	188

LIST OF FIGURES

Figure 1-1. Graph of authors and keywords in works on combination of machine learning and materials science from 2008 to 2018.	8
Figure 1-2. Timeline of clustering of works on combination of machine learning and materials science from 2008 to 2018.	9
Figure 2-1. STEM data on α -RuCl ₃ films.	19
Figure 2-2. Neutron diffraction data on single crystals of α -RuCl ₃	21
Figure 2-3. Analysis of structural distortions in α -RuCl ₃ films.	23
Figure 2-4. STM data on <i>in-situ</i> cleaved α -RuCl ₃ surface.	26
Figure 2-5. Analysis of STM data on α -RuCl ₃ surface.	28
Figure 2-6. Algorithm schematic.	38
Figure 2-7. Structural Information of FeSe _{0.45} Te _{0.55}	40
Figure 2-8. Statistics of local configurations.	44
Figure 2-9. Landscapes of squared statistical distance, s^2 , as a function of the three interaction parameters of the considered model.	47
Figure 2-10. Optimal configurations and results per particle.	49
Figure 2-11. Data of 1% Au-doped BaFe ₂ As ₂	56
Figure 2-12. Data of 1% Au-doped BaFe ₂ As ₂	57
Figure 2-13. Results of <i>k</i> -means clustering.	59
Figure 2-14. Results of PCA analysis performed within each of 3 clusters.	60
Figure 2-15. Results of Bayesian linear unmixing of STS dataset.	64
Figure 2-16. Comparison of room temperature STEM data with proposed structural models.	84
Figure 2-17. Schematics of octahedral distortion found in neutron diffraction data on single crystals (top view) that leads to enlargement of Cl-Cl intercolumn distances (schematically denoted by red arrows); this agrees with STEM observations on exfoliated films. See also Supplementary Note 2-1.2.	85
Figure 2-18. PCA scree plots.	86
Figure 2-19. PCA-derived first 6 normal modes of displacement for (a) center of mass of entire unit cell as origin; (b) center of mass of 6 Ru/Cl columns as origin.	87
Figure 2-20. The <i>k</i> -means clustering analysis.	88
Figure 2-21. Lattice distortion as a possible origin of observed charge order.	89
Figure 2-22. DFT STM simulations for bulk and surface models.	90
Figure 2-23. Comparison of average structural parameters obtained from the bulk and the top-layer of the slab.	91
Figure 2-24. Orbital resolved Density of States (DOS) from the bulk and the slab unit cell respectively.	92
Figure 2-25. Histogram of collected atomic intensities and best fit of 2 Gaussian distributions to represent dark and bright atoms.	93
Figure 2-26. Fraction of dark atoms identified within the image as a function of selected threshold value. At threshold of 0.61 fraction of dark atoms is 45%, so it was chosen for the further analysis.	94
Figure 2-27. Illustration of statistical distance method.	95

Figure 2-28. Top: Examples of six out of sixty local configurations, whose statistics was collected from the image and used as target information for model optimization. Bottom: The histograms of 60 local configurations in the target (blue) and model (yellow) systems.	96
Figure 2-29. Temperature-dependent magnetic susceptibility and resistance (inset) results of BaFe_2As_2 , lightly doped with $\sim 1\%$ gold ($x=0.009$). χ is measured along the two crystallographic axes; T_N is inferred from data; red arrows indicate STM/S temperature points at 4 K and 77 K.	97
Figure 2-30. STM topographic image recorded at $T=77$ K for 1% Au-doped BaFe_2As_2 . Tunneling conditions $U_{\text{bias}}=100$ mV, $I_{\text{setpoint}}=1$ nA. Two common types of point defects are marked with red and orange ovals.	98
Figure 2-31. (Left) STM topographic image recorded at $T=4$ K for 1% Au-doped BaFe_2As_2 . Tunneling conditions $U_{\text{bias}}=-50$ mV, $I_{\text{setpoint}}=200$ pA. (Right) FFT on the topographic image on the left. Two spots corresponding to 1D modulation of a charge density are denoted by dotted circles.	99
Figure 2-32. PCA scree plot describing variance in k -means derived cluster 1 (see main text)	100
Figure 2-33. Bayesian endmembers for different number of total endmembers: $R=5$ (a), $R=6$ (b), $R=7$ (c), $R=8$ (d).	101
Figure 2-34. Abundance maps associated with different number of total Bayesian endmembers in Fig. B2: $R=5$ (a), $R=6$ (b), $R=7$ (c), $R=8$ (d).	102
Figure 2-35. Endmember associated with a pseudogap-like state for BLU unmixing into $R=6$, $R=7$, $R=8$ components. Arrows and dashed line are guides for eye only.	103
Figure 3-1. Application of deep learning to a problem of finding atomic species and defects in a crystal lattice.	111
Figure 3-2. Finding atomic positions from raw experimental images.	114
Figure 3-3. Pixel-wise classification of atomic defects and lattice reconstructions.	116
Figure 3-4. From analysis of pixels in image to analysis of single defect chemistry.	118
Figure 3-5. Tracking a reversible switching between different coordinations of Si dopant.	121
Figure 3-6. Tracking complex defect transformations on the surface.	122
Figure 3-7. Tracking complex defect transformations on the surface during STEM atomic manipulations.	124
Figure 3-8. Analysis of “rotor spinning”.	125
Figure 3-9. Deep learning based identification of atoms, atomic columns and defects.	134
Figure 3-10. Using a single DL model to identify positions of atoms/atomic columns in the images of different materials obtained from two different experimental techniques.	136
Figure 3-11. Identification of anomalies in softmax layer of DL model.	139
Figure 3-12. Identification of lattice atoms and lattice disorder from STEM images using a single DL model.	141
Figure 3-13. Analysis of atomic columns distortion in 3-dimensional Si sample by tracking orientation of individual Si dumbbells.	142
Figure 3-14. Identification of anomalies in <i>softmax</i> layer of DL model for STM data. .	144

Figure 3-15. Identifying molecular rotations and detecting a switching between different molecular conformations via DL.	146
Figure 3-16. Description of the system and physical priors.	155
Figure 3-17. Workflow overview.	157
Figure 3-18. Generating and analyzing synthetic data.....	162
Figure 3-19. Analysis of MRF error rate.	163
Figure 3-20. Convolutional neural network (cNN) decoding.	165
Figure 3-21. Application of MRF and cNN to experimental data of buckybowls on Au(111).	166
Figure 3-22. Schematics of the proposed reaction path for bowl inversion and its effect on the neighboring molecules.	169
Figure 3-23. Schematic representation of the variational autoencoder model applied to relaxor ferroelectric simulations.	173
Figure 3-24. Latent Variables and reconstruction errors for different explored concentrations of [Pb – O] divacancies: (a) 0.035 (b) 0.055.	174
Figure 3-25. Calculated X[Pb–O] vs. T phase diagrams for the system $Pb_{1-x}(Sc_{0.5}Nb_{0.5})O_{3-x}$, with a nano-ordered Sc:Nb-cation configuration.	175
Figure 3-26. Illustration of data augmentation procedure for creating a sufficiently large training set.....	181
Figure 3-27. Plots of accuracy on the training and validation datasets over training epochs for FCN trained for finding atoms in the lattice. The accuracy on the test set is close to 97 %. The FCNs for finding vacancies and dopants typically achieve about 99 % accuracy.	182
Figure 3-28. Detection of vacancies and vacancy-like structures in graphene and WS_2	183
Figure 3-29. DFT calculations of projected density of states (PDOS) for U and D states (summed over all C atoms in the molecule). Strong difference in PDOS at negative energies are associated with enhancement of the STM tunneling current in the experiment at each D molecule.	184
Figure 3-30. First 16 eigenvectors of principal component analysis on the dataset containing all the 938 molecules extracted from Fig. 1c.	185
Figure 3-31. The STM image used in analysis in Fig. 1d without (a) and with (b) structure overlay.....	186
Figure 3-32. PDF functions for all considered rotational classes: (a) 1 st class, (b) 2 nd class, (c) 3 rd class, (d) 4 th class.....	187
Figure 4-1. Beam induced transformations in Si using a 200 kV beam.	197
Figure 4-2. Atomic scale tracking of the local crystallinity.....	199
Figure 4-3. Beam directed crystallization process.	201
Figure 4-4. Controlled crystallization of amorphous Si and dopant movement.	203
Figure 4-5. Timeline showing views of crystallization front in molecular dynamic simulation of Si crystallization. Main frame shows the side view, inset shows the front view.	207
Figure 4-6. Defects evolution under e-beam irradiation in Mo-doped WS_2	216
Figure 4-7. Locating atomic defects in dynamic STEM data using a cNN model trained on a single STEM movie frame.	219

Figure 4-8. Unsupervised classification of defects located by the deep convolutional neural network.	221
Figure 4-9. Defect flows.	223
Figure 4-10. Local crystallography analysis.	225
Figure 4-11. Analysis of transition probabilities between different defect states.....	226
Figure 4-12. (a) Raw experimental data (selected frames). (b) Same data with <i>point-like</i> <i>defects</i> localized. Note that here we do not consider extended defects.	231
Figure 4-13. (a) GMM splitting into 3 clusters. The class 3 structure is clearly distorted due to insufficient number of components (b) GMM splitting into 15 clusters.	232
Figure 4-14. (a) Projection of sub-windows associated with S vacancy (class 5). (b) Projection of sub-windows associated with Mo defects (class 1 and class 3).....	233

CHAPTER 1.

INTRODUCTION

1.1 Materials Science and Technological Progress

Modern civilization is heavily reliant on technological progress. According to the estimates, in 2017 about 9.2% of the U.S. economy was contributed by the technology industry, also employing over 11 million people (up from 7.1% and 6.7 million in 2015). [1] Total world research and development expenditure has grown from 1.956% of global GDP in 2007 to 2.228% in 2015. [2] The global battery market is expected to grow to \$100 billion by 2025. [3] Economic research shows that changes in technology are the only source of permanent increases in productivity. [2]

The explosion of the technological development has been powered mostly by advances in growth, theory, and fabrication of semiconductor materials and devices, allowing their permeation into everyday life, as well as running computations that would have been unfeasible. The rate of the technological change has been following the so-called Moore's Law, named after Gordon E. Moore, who in 1965 speculated that complexity for component cost would double every year until 1975 and double every 2 years afterwards. [4] However, it is expected to end by 2025, as physical limits to traditional transistor scaling are reached. [5] Already, Intel has been experiencing doubling in almost two and a half years. [6] This necessitates looking for non-traditional ways to continue the trend of technological growth. One of the ways is to scale design and manufacturing of materials and devices down to the atomic-level. [7] Another promising research direction has been to employ materials with complex functionalities, such as ferroelectric relaxors and morphotropic systems, [8, 9] multiferroics, [10] spin and cluster glasses, [11] nanoscale phase separated oxides. [12, 13]

However, the classical synthesis-characterization-theory approach has shown to be unable to keep up with the needs of materials manufacturing, requiring 20 or more years to move a material from discovery to market. [14, 15] The connection between structure of the materials and their desired functionalities (room-temperature superconductivity, giant electromechanical or magnetoelectric couplings, [16] efficient room-temperature oxygen reduction reactions [17]) is not well understood, and thus the pathways to synthesize such

materials are not obvious. Unfortunately, much of the literature available is lacking either in synthesis conditions information, or following characterization. [18]

In recognition of the above gaps, a grand effort termed the Materials Genome Initiative was launched to bridge efficiency in materials discovery and design. This multi-agency initiative was created to lead a cultural shift in materials design, to integrate experiment, computation, and theory, and to foster sharing of the digital data across academia, government, and industry. [19] One of the greatest achievements of the approach was development of high-throughput screening of candidate materials for experimental study. [20] However, a lot of interesting functionalities are defined on scales beyond what is accessible for first-principles simulations with current computational capabilities or are present in materials with structures with poor theoretical understanding. Additional difficulty for theory-based approaches is dealing with existence of defects and understanding dynamic states of the matter.

These limitations lead to a need of closer coupling of experiment and theory. To address this need, in my research, I utilized the synergy of imaging, simulations, and artificial Intelligence (AI) to achieve better understanding of materials on the nanoscale, and to pave the way for automated atomic scale manufacturing.

1.2 Artificial Intelligence

1.2.1 History of AI

Creation of beings capable of reasoning and performing tasks has fascinated humanity since antique times. According to the ancient Greek mythology, Hephaestus and Daedalus forged animated metal statues of humans and animals capable of thought. Early Judaism references golems, artificial beings created of mud who can serve humans. In modern times, Karel Capek created a word ‘robot’ to denote intelligent machines in 1920, and the notion has been a persistent topic in fiction since then. [21]

Alan Turing was the first researcher to suggest that a machine can potentially understand logic and perform tasks, leading to the idea of the Turing Test, which is purposed to

distinguish machine intelligence from human intelligence. Research of Artificial Intelligence started in 1956 with great enthusiasm, however, the achieved results did not live up to early promises, leading to significant setbacks in research in the 1970s and 1990s. Fueled by increasing computational power and accumulation of data, the field of AI saw resurgence in the early 2000s, which resulted first in data mining and machine learning algorithms, and finally, since 2012, an explosive growth of Deep Learning methods. In 2017, a MIT Sloan School survey has found that 23% of companies have adopted AI technologies, with another 23% running pilot projects. [22]

1.2.2 Machine Learning

In general, AI methods take in input data and provide resulting outputs. Early AI techniques required careful feature engineering by the domain experts in order to identify suitable representation of the data, as well as programmed instructions for its processing. Within AI, the field of Machine Learning (ML) has evolved to make analysis data-driven and remove the necessity to create explicit algorithms for individual datasets. While initially neglected due to requiring large amounts of data, ML started to flourish in the 1990s with popularization of the Internet and growing computational capabilities.

Generally, ML methods can be categorized into 2 types: supervised and unsupervised. In supervised learning scenarios, the inputs are matched with desired outputs within the training set, and the task of the ML algorithm is to figure out the function that represents the correspondence between inputs and outputs in a most optimal, generalizable way. Such algorithms are mostly trained for classification and regression tasks, resulting in a trained model that can work with new inputs. In unsupervised learning, outputs are not available in the training set, and the goal of the algorithm is to discover underlying structure in the data. This is mostly applied to dimensionality reduction, clustering, and other exploratory data analysis tasks.

1.2.3 Deep Learning

Finally, within the field of ML, representation learning has allowed using raw data without the necessity of constructing hand-picked features. One of the particularly promising methods was artificial neural networks (ANN), which is based on a structure of artificial neurons arranged in layers. These neurons can have different functions for processing of the signal, and different types of connectivity with the raw data and between each other. Such networks learn correspondence between the input data and the output through backpropagation of error and adjusting the synaptic weights, which creates an abstract representation of data at each layer. In the 2010s, growing computational power and utilization of specialized hardware such as GPU [23] and TPU [24] has allowed the researchers to work with raw data directly with high performance. This led to widespread utilization of deep neural networks (DNN).

Deep Learning is a multi-level representation learning method, allowing higher level abstraction features to be learned from the data, leading to higher complexity of functional representation of the data. [25, 26] Deep learning made significant advances possible in fields such as image [27] and speech recognition [28, 29], changing the world we live in. It can also change the way we can do science, as will be shown later in this dissertation.

.

1.3 Atomic scale microscopies and their importance for Materials Science

Macroscopic properties of technologically relevant materials can be significantly affected by the nanoscale inhomogeneity of their chemical, structural, and electronic orders. Major examples include formation of charge nanopuddles in graphene leading to reduced mobility of Dirac electrons [30, 31], piezoelectric response of relaxor ferroelectrics being increased due to the interaction between acoustic phonon mode and nanopolar domains [32], filamentary superconductivity [33], and fluctuating superconducting state in high-T_c cuprates due to electron-pairing regions. [34]

In the last decade, the progress in scanning tunneling microscopy (STM) and scanning transmission electron microscopy (STEM) has allowed real-space imaging of structural and electronic orders of materials on atomic scale. At the same time, atomistic simulations of both structural and electronic orders have been available for a while, using molecular dynamics (MD) and density-functional theory (DFT) approaches.

1.3.1 Scanning transmission electron microscopy (STEM)

STEM is a subtype of transmission electron microscopy (TEM). [35, 36] It produces images by transmitting a beam of electrons through a specimen, achieving higher resolution than optical microscopy due to smaller de Broglie wavelength of the electrons. STEM in particular focuses the beam on a smaller part of the sample which is then scanned over the entire sample. Using an aberration corrector, STEM can achieve picometer (pm) resolution. All STEM data analyzed in this dissertation was collected using annular dark-field imaging detectors, which collect scattered electrons outside of the electron beam. The intensity of atomic columns imaged in such a way is proportional to the square of atomic number of an element.

STEM is scanned in a raster over the sample, resulting in a 2D image. Such images can be collected consequently, allowing study of the temporal evolution of materials over the same spatial region.

1.3.2 Scanning tunneling microscopy and spectroscopy (STM/S)

Scanning tunneling microscopy and spectroscopy allows probing topographic and electronic properties of the surfaces with a sub nanometer-scale resolution. [37] A conducting tip is scanned over the surface of material, allowing tunneling of the electrons. The resulting data is a function of applied voltage, height of the sample, and local density of states (LDOS). At constant current it produces 2D (x, y) datasets of convolution of atomic heights and LDOS. In STS mode, it produces data in 3D, with conductance G proportional to LDOS at a specific energy E at each (x,y) point.

1.4 The Role of Artificial Intelligence and Deep Learning in Materials Science

1.4.1 Previous applications

While Artificial Intelligence methods have been used for materials science for a while [38], deep learning methods have mostly found applications in the related fields of high-energy physics [39-41], biological [42] and medical imaging [43]. What we can observe in those fields is the need to tailor artificial intelligence methods and deep learning models towards particular types of data.

To explore the previous work, I utilized a Citespace [44] analysis of references collected from the Web of Science [45] using a search query “machine learning OR deep learning OR artificial intelligence OR neural networks OR data mining OR computer vision” AND “physics OR chemistry OR materials science OR surface science OR microscopy OR transmission electron microscopy”. The applied algorithm uses both authors and keywords as nodes to create a graph of relationships between them. Graph of keywords can be seen in Figure 1-1.

Figure 1-2 shows the timeline of the clusters discovered in the data from 2008 to 2018. The majority of works in this data set is concerned with utilization of machine learning for analysis of simulations, informing experiment design, and prediction properties of materials.

At the start of the research described in this dissertation, there were no frameworks for utilizing deep learning for analysis of experimental microscopy data, and my research created state-of-the art approaches in this area.

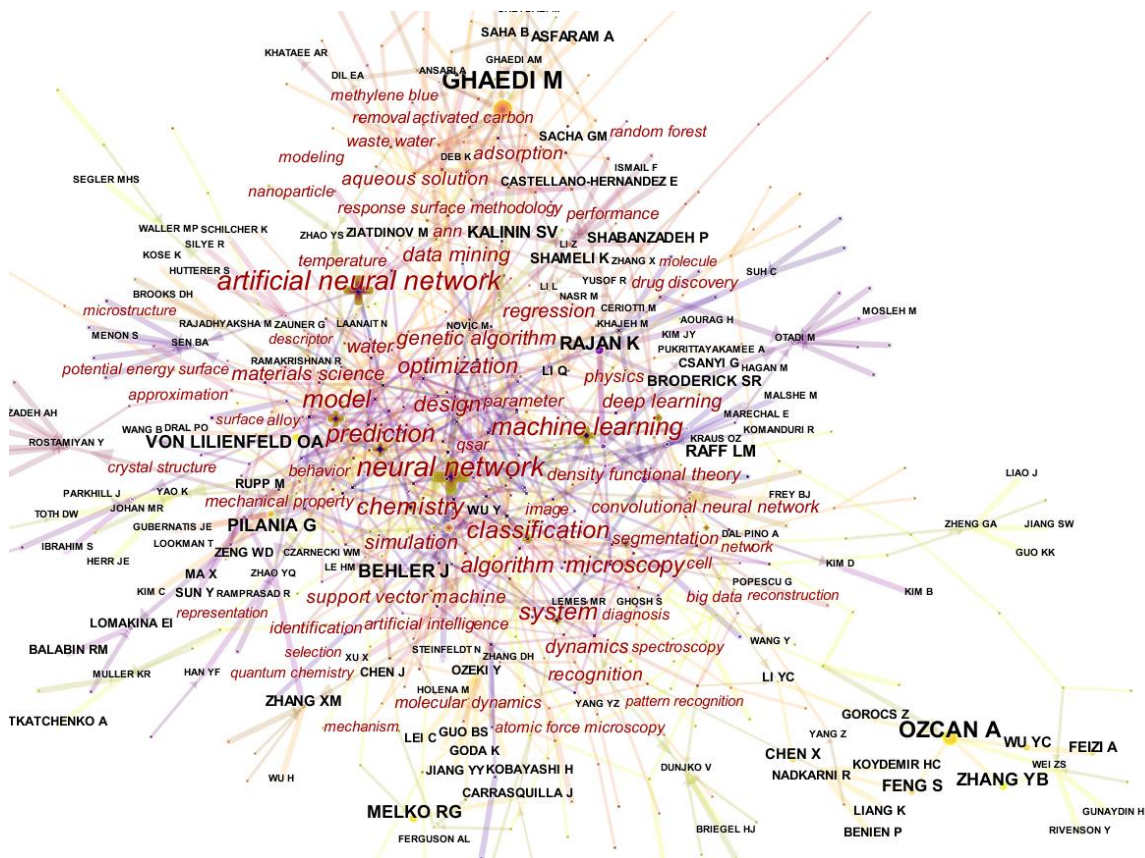


Figure 1-1. Graph of authors and keywords in works on combination of machine learning and materials science from 2008 to 2018.



1.4.2 The goals of AI methods application to microscopy data

Within the scope of this dissertation, I will explore the ways in which AI methods can help us extract the following statistical information from microscopy data:

1. **Localization of features:** this includes identification of atomic positions, as well as defining neighborhoods in a physically meaningful way. Where possible, transforming data from pixel-based to physical element-based (atom, molecule, domain) allows, where available, direct comparison with theory and simulations.
2. **Classification of features:** identifying states of the physical element brings us one step closer to comparison of experiment and theory. Additionally, it can inform theoretical modeling and improve simulations by incorporating knowledge extracted from experimental data.
3. **Grouping of features:** definition of neighborhoods based on physically-meaningful criteria allows application of graph algorithms, as well as correlative methods. This can allow us to identify local interactions, which is particularly important in spatially inhomogeneous materials studied with real-space imaging.
4. **Evolution of features in time:** once the first 3 goals of statistical analysis are achieved, we can study the results at each time step to gain insight into reaction pathways, as well as interactions with electron beam or scanning probes.

1.4.3 Current challenges of DL application for microscopy data

DL inherits the demand for large amount of data for proper training from ML, this poses a great challenge for creation of models applicable to materials science. While the datasets in the industry can be generated by providing technical infrastructure and allowing people

to provide the data for free (Google, Facebook), or by letting the AI learn by practice through reinforcement learning (Playing games, self-driving cars), in materials science, data is very expensive and lacks labels or distinct goals, which makes state-of-the art approaches useless as they rely on large amounts of data. Thus, with the scarce available datasets importance of prior knowledge and domain expertise increases significantly. If we already know analytical solutions to some of the questions we want to answer within the AI framework, it is important to be able to incorporate them, which is usually not a major concern for industrial applications.

AI frameworks mostly lack transferability. They can be taught to win at a particular game, but this trained model will not be able to formulate what are the rules of the game or what winning means.

Another issue that is not quite important in industrial applications is limited transparency of the frameworks. Understanding how a model works is an important consideration for scientific applications trying to extract physically-meaningful information. State-of-the-art architectures consist of large number of layers, obscuring the training. For example, AlexNet architecture uses 5 convolutional layers, and 3 fully-connected layers [46], while VGG family on architectures push the depth to 16-19 layers. [47] Architectures such as ResNet go even further, pushing into hundreds of layers [48] and expanding number of trained filters in the models [49]. Usually such structures are infeasible to study and unnecessary for complexity of the data in, for example, atomic resolution imaging of crystalline materials.

Additional concerns include addressing confounding effects in the data, causation vs correlation, overfitting and underfitting data, and are more generalizable to overall current approach to development of AI frameworks. [50, 51]

CHAPTER 2.
STRUCTURE-PROPERTY RELATIONSHIP FROM STATISTICS

2.1 Introduction

Before we explore applications of deep learning methods to materials science problems, it is important to understand what challenges we can address. In this chapter I demonstrate successful extraction of physical information from microscopy images and spectral data by utilizing AI methods combined with domain expert knowledge.

In particular, I concentrate on three such projects. In the first one, I utilized image processing and machine learning to find distortions in a periodic lattice imaged by STEM and STM. In the second project I again utilized image processing to identify and classify atoms in STM images. In both cases algorithms had to be tailored toward specific imaging types and materials due to differences in types of observed noise and lattice periodicity. These considerations will become important in Chapter 3 of this dissertation, where I will create a deep learning framework, alleviating the need to create image processing algorithms from scratch.

The third project is different, as it is concerned with spectroscopic imaging. I demonstrate how machine learning methods such as PCA and k-means clustering cannot fully address the complexity of the data and require utilization of more advanced methods such as Bayesian linear unmixing. Another important takeaway of this project is the importance of creation of databases of material properties. We show how electronic signals discovered by unsupervised learning can be directly related to ones calculated by theory. Such data could be used as a starting point for analysis of experimental results, yet there is currently no way to accomplish that.

2.2 Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl₃

2.2.1 Publication information

A version of this article was originally published by Maxim Ziatdinov, Arnab Banerjee, Artem Maksov, Tom Berlijn, Wu Zhou, HB Cao, J-Q Yan, Craig A Bridges, David. G Mandrus, Stephen E Nagler, Arthur P Baddorf, and Sergei V Kalinin:

Ziatdinov, M., Arnab Banerjee, A. Maksov, Tom Berlijn, Wu Zhou, H. B. Cao, J-Q. Yan et al. "Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl₃." *Nature communications* 7 (2016): 13774. [52]

The article content is reprinted as published with figure numbers modified and supplementary material added at the end of the chapter as an appendix.

2.2.2 My contribution

Most studies of α -RuCl₃ have concentrated on reciprocal space and bulk properties, but real-space imaging of structural and electronic orders can provide important information on the spin-orbit effects, electron-electron interactions, and lattice distortions, which all affect the poorly-understood phase diagram of the material. However, to produce quantifiable results, it was necessary to identify atomic positions and analyze their statistics.

This study concentrated on studying exfoliated/cleaved samples via in-situ high-resolution annular dark field STEM, generating 2D images which required further image and multivariate statistical analysis. Images were taken perpendicular to the c^* -axis at 295 K. Observed structure can be described as 2D unit cell containing 6 atomic columns with a mix of Ru and Cl atoms, and 3 atomic columns with only Cl. Additionally, there also was STM images.

In STEM and STM images, I identified positions of atoms in the image by using a combination of correlation surface between FFT-transformed average unit-cell image and real-space image and position refinement using local intensity maximum criteria.

For STEM images, I created histograms of distances to 6 nearest neighbors for each atom, revealing distribution deviation from normal for pure Cl atomic columns.

I further transformed positional data to be represented as a matrix with 90 identified unit cells with 18 (x,y) positions per cell. I performed principal component analysis (PCA) of normal modes of displacement, demonstrating direction and relative strength of displacement within unit cell, which confirmed much higher effect within pure Cl atomic columns, going beyond the predictions made by DFT.

In STM images, I identified atomic positions to construct a nearest neighbor lattice, extracted atomic intensities, and constructed a histogram, allowing classification of corresponding surface atoms within the charge ordered (CO) pattern. This classification has also revealed presence of dimer-like electronic superlattice admixed into the CO.

In summary, structural and electronic inhomogeneity of the material was quantified and compared to simulations and neutron scattering data, which already had methods to provide quantifiable results.

2.2.3 Abstract

A pseudospin-1/2 Mott phase on a honeycomb lattice is proposed to host the celebrated two-dimensional Kitaev model which has an elusive quantum spin liquid ground state, and fascinating physics relevant to the development of future templates towards topological quantum bits. Here we report a comprehensive, atomically resolved real-space study by scanning transmission electron and scanning tunnelling microscopies on a novel layered material displaying Kitaev physics, α -RuCl₃. Our local crystallography analysis reveals considerable variations in the geometry of the ligand sublattice in thin films of α -RuCl₃ that opens a way to realization of a spatially inhomogeneous magnetic ground state at the nanometre length scale. Using scanning tunnelling techniques, we observe the electronic

energy gap of ≈ 0.25 eV and intra-unit cell symmetry breaking of charge distribution in individual α -RuCl₃ surface layer. The corresponding charge-ordered pattern has a fine structure associated with two different types of charge disproportionation at Cl-terminated surface.

2.2.4 Introduction

The peculiar interplay between spin-orbit effects, electron–electron interaction, and subtle lattice distortions is expected to produce an elaborate phase diagram for the $4d^5$ transition metal layered compound α -RuCl₃, which to date remains inadequately understood. [53-56] The proposed state of an electron on the honeycomb lattice of edge-shared RuCl₆ octahedra is described by a spin-orbit entangled $J_{\text{eff}}=1/2$ Mott state [53, 56, 57] that can potentially host novel exotic quantum phases, such as a room temperature Quantum Spin Hall Effect [58] and Kitaev quantum spin liquid (QSL) behaviour [55, 56, 59, 60]. The elementary excitations associated with the Kitaev QSL model include Majorana fermions [61-64], which are relevant in the context of a topological quantum computer [65]. Recent identification of a magnetic phase proximate to the Kitaev QSL in neutron scattering experiments [55] have further fueled an excitement over a breakdown of the putative classical physics in the α -RuCl₃ system.

As the nearest-neighbour coupling between $J_{\text{eff}}=1/2$ moments in $4d^5$ and $5d^5$ honeycomb compounds is highly sensitive to a small distortion of the 90° metal-ligand-metal bond [66-68], a rigorous understanding of the local structural properties of the α -RuCl₃ compound is required for exploring a potential departure from the Kitaev QSL phase and realization of a spatially inhomogeneous magnetic ground state in the presence of lattice disorder. An equally important factor in establishing the atomic scale phenomenology for the α -RuCl₃ system is a study of its local electronic behaviour. Whereas most of the experiments have so far focused on bulk magnetic properties of α -RuCl₃, the experimental research on its nanoscale electronic properties, including a potential role of unit-cell scale charge (re)distribution in determining the final ground state and the associated exotic phenomena

in the Kitaev limit, is scarce. In addition, exploring properties of thin films, the surface effects in α -RuCl₃, and their relationship to the bulk parent compound can provide important clues for potential applications of ‘Kitaev materials’ in next generation of 2D nanoscale quantum electronic devices for the post-silicon era.

While, most of the studies in the material have concentrated on reciprocal space and bulk properties [54, 55, 69-71], to explore the many complexities of this system, it is important to be able to visualize directly its local structural and electronic orders in real space.

In the following, we employ a combination of *in-situ* scanning transmission electron microscopy (STEM) and *in-situ* scanning tunnelling microscopy (STM) on exfoliated/cleaved α -RuCl₃ samples as ideal tools for detailed evaluation of structural and electronic parameters of the system with a sub-nanometre precision. To date the only existing in literature atomically resolved STM study on α -RuCl₃ reported an apparently strong lattice distortions on the surface of the sample [72]. However, as those STM measurements were performed in the ambient environment and the samples were not cleaved *in situ*, it remains unclear whether the observed structure peculiarities were the intrinsic properties of the material or caused by some extrinsic effects, such of sample and/or STM tip contamination. The STM and STEM measurements performed here are further supported by neutron scattering, first principles calculations, and a multivariate statistical analysis. This combination of both reciprocal space and real space tools, together with a computational effort, provides an excellent starting point in establishing the structure-property relationship in α -RuCl₃ system. In this study, we find that the exfoliated thin films (thickness $\approx 15\sim 30$ layers) can be characterized by $P3$ -type space group, which is explained with a help of neutron diffraction data as due to transition from $C2/m$ to $P3_1$ type stacking order at above ~ 150 K. For each individual α -RuCl₃ layer we found that in addition to a uniform octahedral distortion reported earlier for bulk α -RuCl₃ crystals, there is a persistent local inhomogeneity in the ligand geometries. Such nanometre scale perturbations in structural order, inferred from the STEM data with a subpixel precision and not typically accounted for by theory, can have an important role for determining collective spin-orbital state of the system in Kitaev-Heisenberg model. The STM spectroscopic measurements in the paramagnetic phase of the *in-situ* cleaved α -RuCl₃

samples demonstrated a presence of ≈ 0.25 eV gap in density of states of the individual surface layer, which is explained by a realization of $J_{\text{eff}}=1/2$ Mott state. Furthermore, our STM probe into the atomic-scale surface electronic structure revealed a striking anisotropy in a charge distribution along the Ru–Cl–Ru hopping paths, indicating that the bond-dependent behaviour in this system may appear at the temperatures considerably higher than the estimated strength of the Kitaev coupling. In addition, we observed a fine structure of the corresponding charge ordered state characterized by a coexistence of trigonal-like and dimer-like patterns of charge disproportionation at the Cl-terminated surface.

2.2.5 Results

Structure (neutron and STEM)

We start with a real-space characterization of α -RuCl₃ films using high-resolution annular dark field STEM measurements. A representative Z-contrast STEM image (Z is atomic number) of α -RuCl₃ film (thickness ≈ 30 layers) taken perpendicular to the sample's c^* -axis at $T=295$ K is shown in Fig. 2-1a. The image consists of a periodic repetition of an effective unit cell containing six brighter columns and three less bright columns. The registration of higher and lower Z atom columns with the $P3_112$ structural model for α -RuCl₃ proposed in ref. [55] provides a good initial match to the entire image (see Supplementary Note 2-1.1 and Fig. 2-16). As intensity in these STEM images is proportional roughly to the square of Z, the six brightest spots in the image are assigned to overlapping atomic columns of Ru and Cl atoms (Ru/Cl), whereas the three less bright spots are the Cl atomic columns (see inset in Fig. 2-1a). It is worth noting that for the $C2/m$ phase, reported recently in bulk α -RuCl₃ crystals [73], one would expect a STEM image consisting of a well-defined line-like features with every third line substantially darker than the two others (see Fig. 2-16), which is not the case in Fig. 2-1a. In this regard, we point out that stacking order in quasi-2D exfoliated films may not necessarily resemble the stacking order of bulk as-grown crystals,

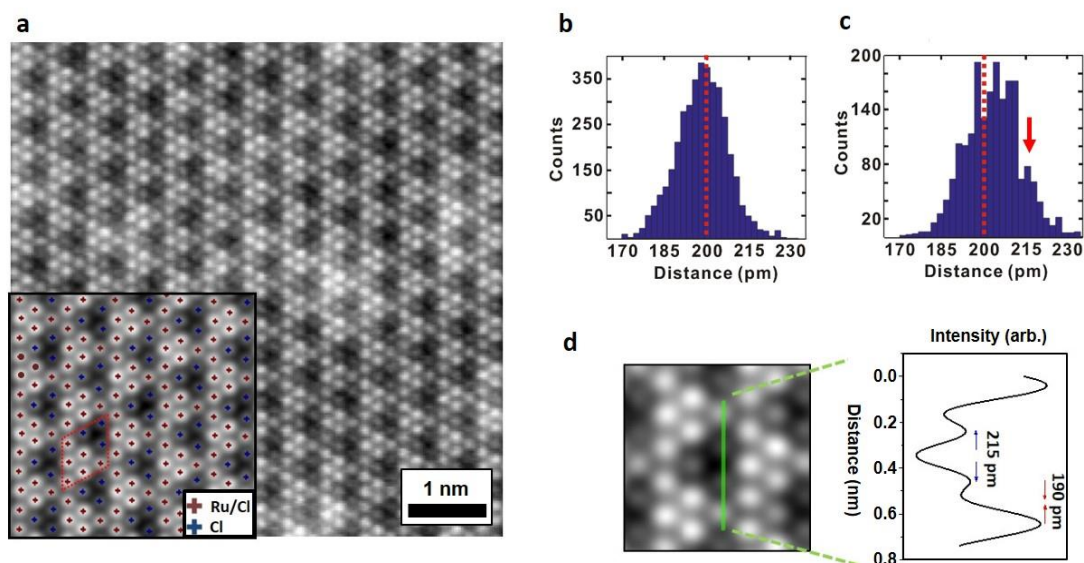


Figure 2-1. STEM data on α -RuCl₃ films.

(a) Z-contrast STEM image of α -RuCl₃ film (raw data). Inset shows filtered data, where brown and blue crosses correspond to the centers of Ru/Cl columns and Cl columns, respectively. (b, c) Histogram-based visualization of inter-column distances for 6 nearest neighbors of Ru/Cl columns (b) and Cl columns (c). Dotted red vertical lines show the value of inter-column distances expected for perfect trigonal symmetry. The red arrow in (c) indicates presence of a well-defined shoulder originating from an increase in nearest neighbor distances of Cl columns. (d) Line profile illustrating a modulation of inter-column distances for Cl and Ru/Cl columns with respect to the value of 198 pm for ideal trigonal symmetry.

which themselves may be subject to faults that modify the layer stacking under particular crystal growth conditions. However, our neutron measurements on bulk crystals does confirm a more general possibility of $P3$ -type space group scenario in α - RuCl_3 system as described below.

The larger single crystals which were exfoliated for this study are consistently $C2/m$ structure at low temperatures [73, 74], however, at high temperatures they comply with the $P3_1$ structure. The two structures are connected by a first-order phase transition at $T \sim 150$ K with a ~ 30 K hysteresis region (Fig. 2-2a), which explain the kinks observed in past susceptibility measurements [71]. This structural transition is characterized by the $(1,0,L=3n)$ and $(1,1,L \neq 3n)$ peaks appearing, while $(1,1,L=3n)$ ($n=\text{integer}$) peaks disappearing above 150 K, as shown in Fig. 2-2b (see Table 2-1 for the refined parameters, and methods section for details of the measurement). Indeed, the $C2/m$, $P3_1$ and $P3_1$ type space groups are polytypes [74] of each other with similar in-plane honeycomb morphology. The difference is the out-of-plane stacking arrangement. All the analysis in the following sections will be based on the $P3_1$ -type space group assuming that the in-layer structural properties discussed in the text are applicable to both $C2$ and $P3$ stacks of weakly coupled α - RuCl_3 layers.

The statistical analysis of the inter-column distances distribution performed for six nearest neighbours of Ru/Cl and Cl columns (Fig. 2-1b,c, respectively) shows a significant distortion from ideal trigonal symmetry (that is, the symmetry expected from a 2D-projection of atomic columns in ABC-stacked α - RuCl_3 with undistorted RuCl_6 octahedra), in which all inter-column distances are expected to be 0.198 nm. This is particularly clear from the non-symmetric nearest neighbours distance distribution for Cl columns that has a well-defined shoulder at around 215 pm (red arrow in Fig. 2-1c). The pairwise analysis of inter-column distances shows that the shoulder is associated with an increase in the Cl–Cl inter-column distances by $\Delta D_{\text{Cl-Cl}} \gtrsim 15$ pm. (Fig. 2-1d). This indicates a presence of a lattice distortion in the film system that is addressed below.

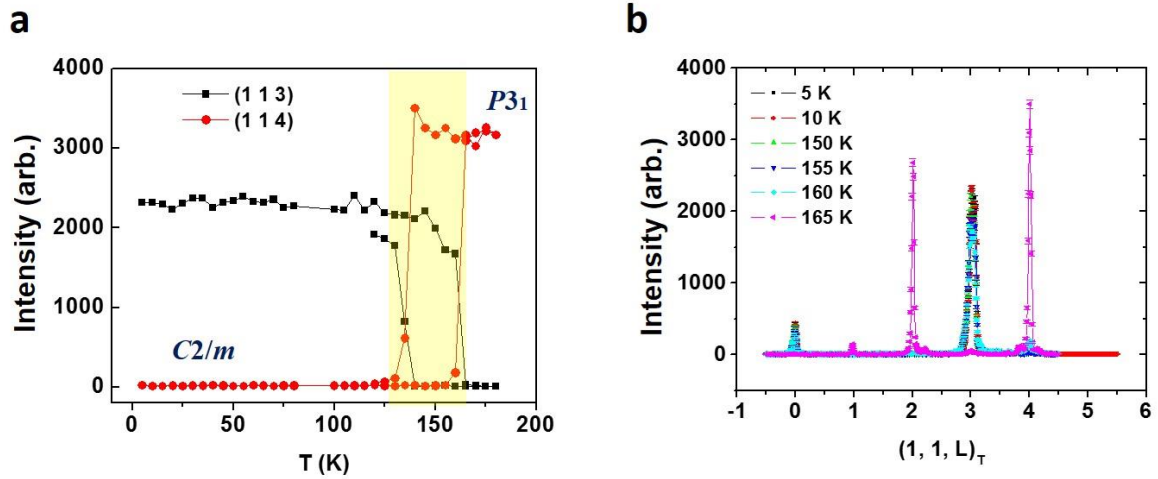


Figure 2-2. Neutron diffraction data on single crystals of α -RuCl₃.

(a) Temperature dependence of the (1,1,3) and (1,1,4) peak intensities. The hysteresis of the (1,1,3) and (1,1,4) orders is clearly observed (shaded area). The structural transition temperature between monoclinic and trigonal lattices is $T_s=165$ K for warming up and $T_s=140$ K for cooling down. (b) L-scans at several selected values of temperature.

To gain further insight into the details of structural distortions found above we performed density functional theory (DFT)-based first-principles calculations on α -RuCl₃ (see Methods section for computational details) and compared them to the STEM results. Our starting point is the $P3_112$ unit cell with nearly perfect local cubic symmetry in which all Cl–Ru–Cl bonds in the Ru-centered Cl octahedra are equal to $(90\pm1)^\circ$ (refs [69, 75]). The DFT calculated relaxed structure shows that Cl atoms from opposite sublayers in each RuCl₃ ‘sandwich’ become displaced towards each other (Fig. 2-3a). This result is consistent with the findings reported in ref. [76]. From additional simulations we have found that it is robust against changing the stacking and the inclusion of spin-orbit coupling, interactions and magnetism. The resultant compression of the Cl ligand cage along Ru–Ru ‘bonds’ (C_3 symmetry axis for threefold rotation) leads to a deviation from its original perfectly octahedral symmetry. Notably the lattice distortion is limited mainly to the Cl sublattice, leaving the Ru honeycomb lattice close to a distortion-free state within the resolution of available experimental structure methods. In Fig. 2-3c we show the DFT calculated structure of atomic columns superimposed on to the STEM experimental image of α -RuCl₃ film (top-view). One can immediately see that the direction and magnitude of a lateral component of Cl cage deformation matches well with the Cl columns displacement found from the STEM experimental image. Particularly, we found that a distance between centers of mass of Cl columns in theoretical structure is expanded by 12 pm, which is very close to an average experimental value of 15 pm (see Fig. 2-3b). While the presence of distortion in the bulk α -RuCl₃ crystals was known from earlier papers (see also Supplementary Note 2-1.2 and Fig. 2-17), the STEM measurements presented here provide the first real-space evidence for octahedral distortion in thin films of α -RuCl₃.

The average STEM unit cell of α -RuCl₃ films derived above agrees well with neutron data and DFT calculations providing us a good understanding of a long range atomic order in the α -RuCl₃ system. We now use information contained in microscopic degrees of freedom available from STEM images to characterize atomic structure on a local scale [77-80]. Particularly, we employ a principal component analysis to search for statistically significant distortions of an average unit cell structure of STEM observations in Fig. 2-1a

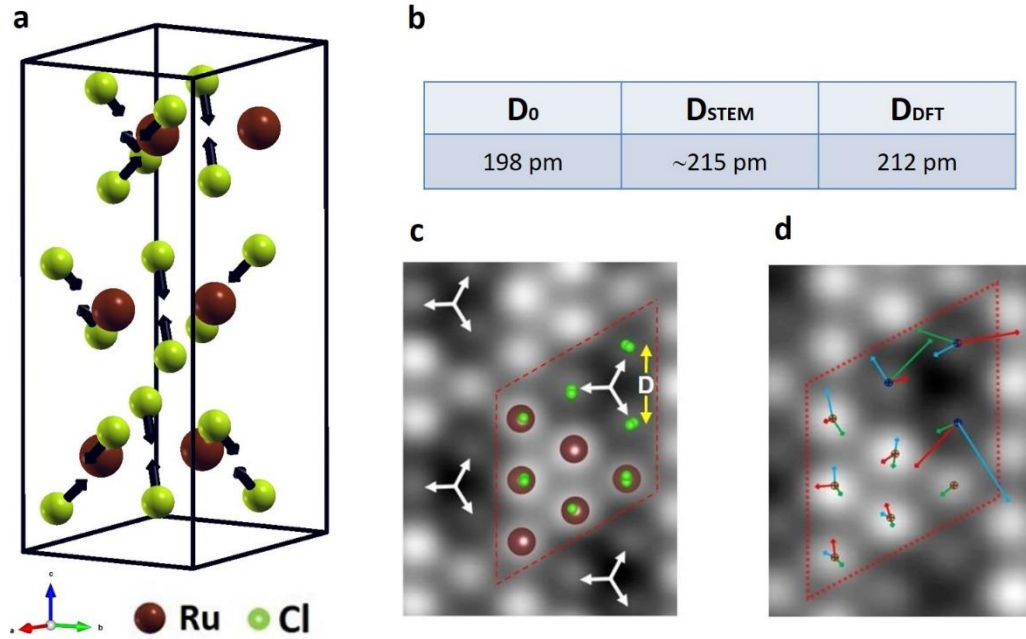


Figure 2-3. Analysis of structural distortions in α - RuCl_3 films.

(a) DFT-calculated relaxed structure of α - RuCl_3 unit cell. The equilibrium length of Ru-Cl bonds found in DFT is 236 pm. The Cl-Ru-Cl bond angle at the shared octahedral edges in the relaxed structure is $\approx 86.3^\circ$, whereas the rest of Cl-Ru-Cl angles are approximately 91.3° . (b) Cl-Cl inter-column distances in unperturbed lattice (D_0) [69, 75], DFT calculations (D_{DFT}), and experimental STEM data (D_{STEM}). (c) DFT-based coordinates of ABC-stacked α - RuCl_3 cell superimposed on experimental STEM image. The white arrows schematically show the displacements of Cl columns. (d) PCA analysis of normal modes of displacement for $M=90$ unit cells (810 atomic columns); the center of the mass of the entire effective unit cell is chosen as origin. The PCA-derived first three eigenmodes are presented as vectors of deformation from columns position in the averaged cell. Coloring scheme: first three displacement modes are described by blue, red, and green arrows, respectively. The length of the arrows is magnified by a factor of 20. The dashed rhombus in (c) and (d) schematically denotes the effective unit cell.

[78]. This issue is important because unit-cell scale variations in order parameters, usually not seen in spatially averaged, or reciprocal space measurements of bulk crystals, may have powerful and non-random effects on the macroscopic observables in strongly-correlated systems [81, 82]. Interestingly, we found only relatively small displacements from average structure in the Ru/Cl columns (see also Fig. 2-18 and Fig. 2-19). As the location of Ru/Cl column centers is determined largely by a contribution from the Ru core electrons (due to the large difference in atomic numbers (Z) between Cl and Ru), the results in Fig. 3-3d suggests that the positions of Ru atoms in a honeycomb lattice remain nearly intact in a symmetric fashion. On the other hand, relatively large variations from average were observed in the positions of Cl columns. The detected local inhomogeneity in a ligand sublattice can originate from a presence of distortion nanodomains ('patches'), whose local structure, such as bond length and bond angles, deviates from the average lattice structure (see also Supplementary Note 2-1.3). A resultant interplay between local inhomogeneity and 'regular' lattice structure may lead to alternation of relative strength (and, in principle, signs) of terms in the spin-1/2 Hamiltonian. It will be interesting to see its implications for the quantum Heisenberg-Kitaev model, where a large number of phases can be generated from the pure Kitaev limit [83] by using perturbations that are highly sensitive to the local ligand environment [84, 85]. One intriguing consequence of such local perturbations is a potential realization of a spatially inhomogeneous ground magnetic state in the system. Furthermore, we suggest that our observation of the local variations in lattice structure can provide a lacking ingredient for matching theory models to the results of magnetic measurements on α -RuCl₃ (refs [86, 87]).

Local density of states

Having analyzed in detail the local structural properties of α -RuCl₃ crystals, we turn to a real-space characterization of its electronic structure by virtue of ultra-high-vacuum STM imaging and spectroscopic tools. The representative STM topographic image of *in situ* cleaved α -RuCl₃ surface shown in Fig. 2-4a reveals a hexagonal lattice associated with Cl-terminated surface layer. A typical scanning tunneling spectroscopy (STS) probe of the density of states on such surface over a wide energy region at $T=295$ K (Fig. 2-4b) shows

a particle-hole asymmetric dI/dU curve with two well-defined peaks centered at about ± 0.7 eV. The presence of a pronounced STS peak at ≈ -0.7 eV is in a good agreement with early angle-resolved photoemission spectroscopy (ARPES) measurements of α -RuCl₃, in which the author observed a quasi-flat band at 0.7 eV below the Fermi level associated with localized Ru 4d states [88]. Zooming in to a narrower energy range in Fig. 2-4c, we found persistent peaks in LDOS at about ± 0.12 eV, as well as a strong suppression of density of states in the energy window of ≈ 0.1 eV around the Fermi level. These observations demonstrate a presence of a charge excitation gap at the Fermi level. The formation of an energy gap well above the temperature of a magnetically ordered phase indicates that α -RuCl₃ is a Mott-type insulator [89]. Recent DFT calculations [90] have shown that the α -RuCl₃ single layer remains metallic if the spin-orbit coupling and electron–electron interactions are not accounted for. The inclusion of spin-orbit coupling and electron correlations leads to an insulating character of density of states with unambiguous charge gap due to a Mott-Hubbard splitting of $J_{\text{eff}}=1/2$ states. Therefore, our STS observations support a theory view of a monolayer α -RuCl₃ as a spin-orbit assisted Mott insulator [90]. To the best of our knowledge, this is the first observation of $J_{\text{eff}}=1/2$ state in an individual surface layer of α -RuCl₃.

Interestingly, the ≈ 0.25 eV surface charge gap found in our STS experiment (measured as an inter-peak separation in Fig. 2-4c) is much smaller than the gap of ≈ 1.2 eV observed in recent ARPES measurements [91], which was also explained as the correlation induced Mott gap. On the other hand, the presence of ≈ 0.25 eV gap in the STS experiment agrees well with observations of a peak at about 0.3 eV in X-ray adsorption measurements [69, 92] and of a 0.2 eV gap in neutron measurements of bulk α -RuCl₃ crystals [55]. The origin of the current discrepancy between STM and ARPES measurements of the gapped state of α -RuCl₃ must be carefully addressed in future studies of this material.

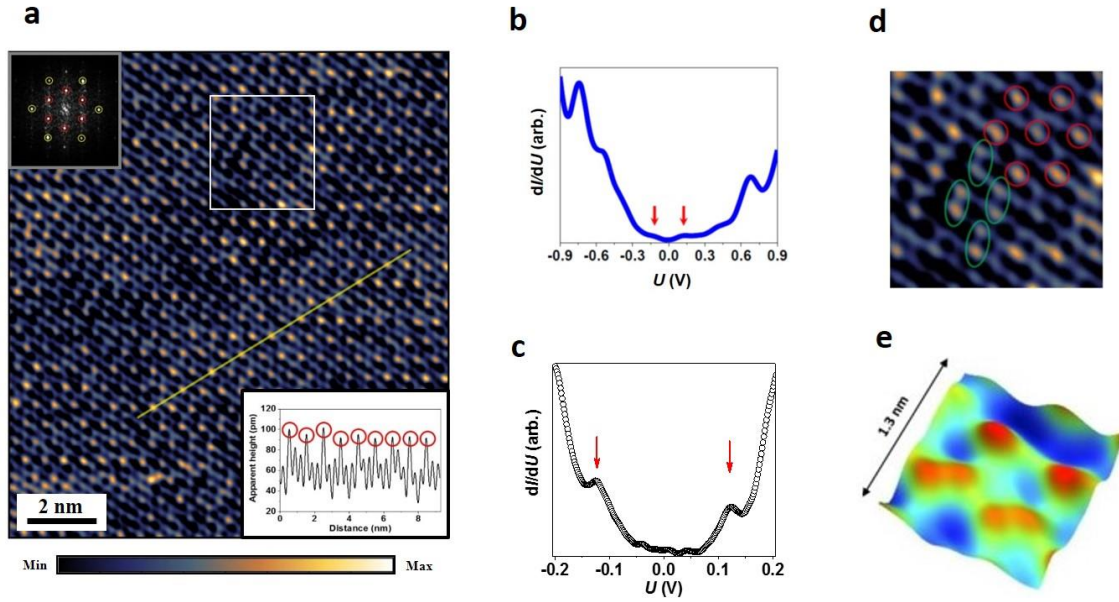


Figure 2-4. STM data on *in-situ* cleaved α - RuCl_3 surface.

(a) Representative STM constant current image of the RuCl_3 surface layer (tunneling conditions: $U_{\text{tip}}=300\text{mV}$, $I_{\text{setpoint}}=60\text{pA}$). Insets: 2D FFT of the image (top left), STM cross-sectional profile along the yellow line (bottom right). (b) Averaged STS spectra over a wide energy window; $U=-U_{\text{tip}}$ (c) STS spectra in the narrower energy window around the Fermi level showing peaks at $\approx\pm 0.12\text{V}$ (denoted by red arrows). The spectra were averaged over 40 individual dI/dU curves. Setpoint parameters: $U_{\text{tip}}=-500\text{ mV}$ (b), $U_{\text{tip}}=-400\text{ mV}$ (c); junction resistance $R_j=25\text{G}\Omega$ (b), $20\text{G}\Omega$ (c). (d) Zoomed-in area from white rectangular in (a); coexistence of trigonal CO and dimer-like CO shown by red circles and green ovals respectively. (e) 3D rendered high-resolution STM image highlighting a crossover between 2 different phases of CO at $U_{\text{tip}}=-1.0\text{ V}$, $I_{\text{setpoint}}=100\text{ pA}$.

Charge density modulation

To further explore details of the real-space electronic behaviour on a local scale it is crucial to be able to identify positions of electron density peaks associated with an underlying Cl atomic sublattice in the STM image in automated fashion. For this purpose, we first create a template unit cell from fast Fourier transform-filtered image data, and then perform normalized cross-correlation analysis coupled with position based and intensity based refinement in order to extract maximal electron densities associated with Cl sublattice (see Supplementary Note 2-1.4 for details). The resultant lattice of electronic densities is shown in Fig. 2-5a (see also Supplementary Note 2-1.5 and Fig. 2-20 for analysis of the nearest neighbour distances in the constructed lattice).

Once the construction of electron densities lattice is completed (that is, all the positions of electron density peaks are identified), one can clearly see a modulation of STM intensity associated with an intra-unit cell symmetry breaking in a charge density distribution at three ‘top’ Cl atoms in the α -RuCl₃ layer unit cell. This can be confirmed by looking at the histogram-based visualization of the STM intensities distribution (Fig. 2-5b) for all the ‘lattice’ points associated with Fig. 2-4a.

The two well-defined peaks in the histogram, whose height differs by roughly a factor of 2, reflect the enhancement of the STM intensity on every third Cl atom. On a larger scale, this leads to a formation of a charge ordered pattern (CO) with a $\sqrt{3} \times \sqrt{3}R30^\circ$ ($R3$) surface symmetry (R is a translational vector of an unperturbed Cl surface lattice). Noteworthy, an emergence of a qualitatively similar $R3$ electronic superlattice has been recently reported in the STM experiment on the honeycomb iridates, where it was explained as purely due to the structural reconstructions [93]. Since STM cannot always distinguish between electronic and structural contributions, one may argue that the non-uniform distribution of STM intensity on α -RuCl₃ surface is in fact due to the geometrical tilts/rotations of Ru-centered Cl octahedral cage. However, a presence of such octahedral tilts with long range order does not follow from the detailed structural analysis described in the first part of the paper. Indeed, all Cl surface atoms remained confined to a two-dimensional plane in our structural analysis. This confirms that the formation of the $R3$ superlattice is due to the

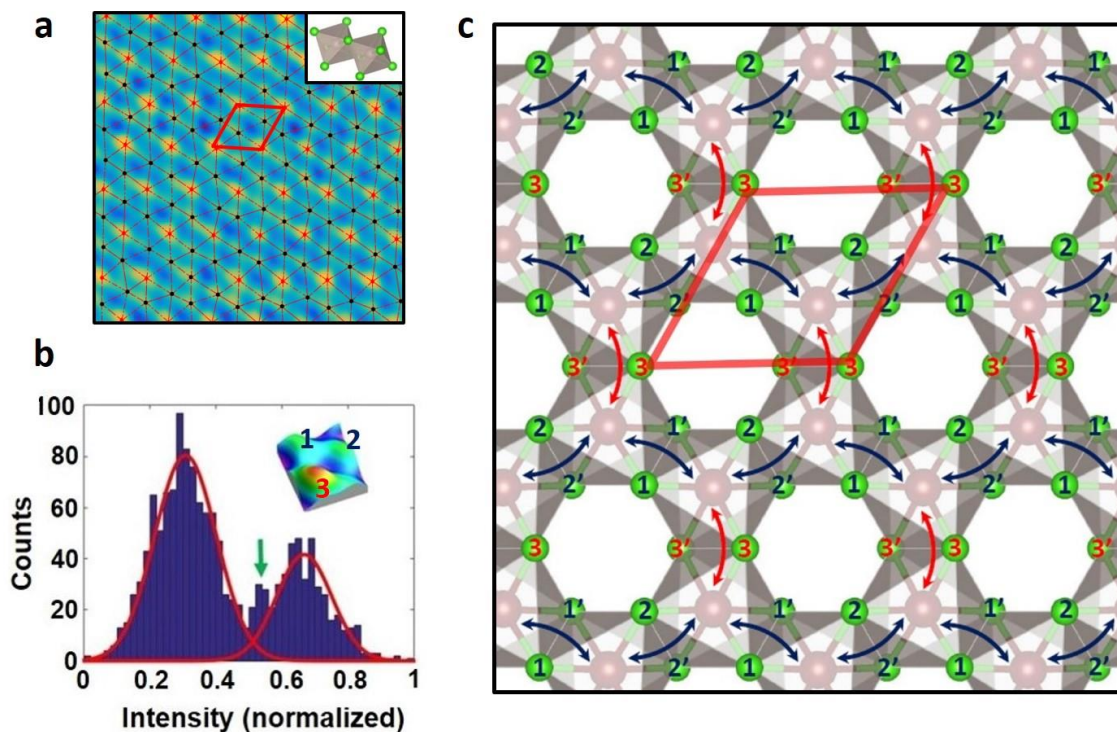


Figure 2-5. Analysis of STM data on α -RuCl₃ surface.

(a) Color-coded binary lattice overlay. The red lattice spots correspond to the normalized intensity above the threshold of 0.5 and form a superlattice of a $\sqrt{3} \times \sqrt{3}R30^\circ$ periodicity. (b) Histogram-based visualization of distribution of STM intensities at Cl atoms in (a); the two peaks are approximated by Gaussian curves. The green arrow indicates admixture of a competing CO into charge density distribution. Inset in (b) shows a 3D rendered experimental image of Cl triangular unit with a non-uniform distribution of STM intensities. (c) Schematic model showing anisotropy of charge density at Cl atoms along Ru-Cl-Ru electron hopping pathways (double-end blue and red arrows). The 3 “top” Cl atoms in the unit cell are labeled in the same way in (b) and (c).

electronic symmetry breaking at Cl-terminated surface of α -RuCl₃ crystals. The superlattice and regular Cl lattice form the inner and outer hexagons, respectively, rotated by 30° with respect to each other in the 2D fast Fourier transform of the STM image (inset in Fig. 2-4a). The *R3* pattern was observed up to relatively high energies with respect to Fermi level ($U_{\text{tip}} \approx |1.5 \text{ V}|$), at both positive and negative bias polarities.

To elucidate a possible physical origin of the CO we start with discussing a potential relation between CO and lattice degrees of freedom in α -RuCl₃. The trigonal component of RuCl₆ octahedral distortion found in DFT results in a well-defined periodic modulation of Cl interatomic distances that breaks the hexagonal symmetry of 2D surface layer. The resultant distorted hexagonal pattern features alternating short and long Cl–Cl interatomic distances of 337 pm and 359 pm, respectively (see Supplementary Note 2-1.6 and Fig. 2-21). The symmetry of atomic distortions at the surface plane matches well with the symmetry of the *R3* superlattice formed by the brightest spots in the STM image, implying a direct relation between Cl atomic displacements and CO. However, our DFT-based STM simulations did not reveal any imbalance in the distribution of charge density despite finding the distortion of Cl cage (see Fig. 2-21), indicating that the Cl displacements are not the primary cause of the observed CO. We have also investigated the possible role of surface relaxation effects (see Supplementary Note 2-1.7, as well as Figs 2-22 to 2-24) and found these to be negligible reflecting the fact that α -RuCl₃ is a quasi-2D system.

Another argument explaining the emergence of a CO phase stems from the possible non-equivalence of the three ‘bond axes’ (that is, the nearest neighbour Ru–Ru distances within a layer, which hereafter are referred as ‘bonds’) in the sample. We note that the intra-unit cell symmetry breaking in the STM experiments persists at the energies as high as +1.5 eV above the Fermi level suggesting involvement of both t_{2g} -derived and e_g -derived electronic states in a formation of the CO pattern [69]. This observation is important because a hopping process between t_{2g} and e_g orbital via ligand atom is vital in determining a ground state of Kitaev-Heisenberg model on the honeycomb lattice [94] (see also Supplementary Note 2-1.8).

The experimentally observed anisotropy in a charge density distribution along Ru–Cl–Ru hopping pathways, schematically denoted by red and blue arrows in Fig. 2-5c, indicates

that the bond-dependent behaviour in α - RuCl_3 may appear at temperatures ~ 3 times higher than an estimated strength of the Kitaev coupling (~ 100 K) (refs [55, 71]). Since the STEM and neutron measurements suggest that the high-temperature lattice structure primarily complies with a trigonal $P3_1$ type space group and that the Ru sublattice is isotropic (see for example, Fig. 2-1b), which agrees with DFT calculations, we argue that it is the anisotropic exchange pathways via the Cl sublattice that imparts additional bond-directional anisotropy along the Ru–Ru bond axes. It results in an overall anisotropic energy distribution, that could explain why the RuCl_3 layers rearrange and form a $C2/m$ structure at temperatures below 154 K (ref. [74]). Arguably, this would also determine the direction of the ‘a’ and ‘b’ axes of the $C2/m$ structure. This, in turn, would determine the direction of the zig-zag ordering, where the blue ‘bonds’ in Fig. 2-5c have aligned spins, and the red ‘bonds’ have anti-aligned spins. This anisotropy in the chlorine position will directly affect the Kitaev Hamiltonian and has to be included in the calculations to simulate both the anisotropic susceptibility and the spin-wave gap of 2 meV observe in this material³.

We next discuss a fine structure of the CO state at the Cl-terminated surface, in which the clusters of dimer-like electronic superlattice are admixed into the trigonal $R3$ CO electronic superlattice discussed above (see Fig. 2-4d,e). These two types of superlattices were persistently found in various areas of the sample at different bias voltages. The admixture of the dimer-like superlattice into CO produces an additional weaker peak in the histogram in Fig. 2-5b (denoted by a green arrow), as confirmed by the inspection of absolute values of the normalized STM intensities extracted from Fig. 2-4a. This effectively turns the surface of α - RuCl_3 into a competing CO heterogeneous system. The peculiar switching between two surface electronic superlattices at nanometre length scales can be related to a subtle crystalline disorder, such as occurrence of vacancies in the bottom Cl layer of RuCl_3 ‘sandwich’ which cannot be directly seen in the STM, as well as due to the presence of defects in the second RuCl_3 ‘sandwich’ layer. Furthermore, the variations of an average unit cell structure in the Cl sublattice have been routinely seen in our analysis of the STEM data. The nanoscale strain fields associated with competing structural orders can induce a transition of a purely electronic nature between different CO phases as was recently

reported for NbSe₂ (ref. [95]). Another interesting possibility is that the observed fine structure of CO is a reflection of magnetic short-range correlation persisting above T_N (ref. [87]). It remains to be seen in the future, if magnetic patches could couple via the strong spin-orbit coupling to the orbital occupation in the Ru atoms, which in turn would alter the hybridization with the Cl- p orbitals and thus the tunnelling current in the STM tips. In addition, it will be educational to measure the anisotropy of the in-plane susceptibility signal, as such will also be affected by these magnetic patches. Finally, the spatial fluctuations of CO may produce an additional scattering potential for electrons in α -RuCl₃ and associated contribution to the resistivity of the individual α -RuCl₃ layers.

2.2.6 Discussion

To summarize, we have presented the first comprehensive study of the local electronic and lattice degrees of freedom in α -RuCl₃. We expect that the findings reported here will have significant ramifications in several areas of research on spin-orbit coupled Mott insulators with possible Kitaev QSL state. First, our finding of a charge gap at E_F in a surface monolayer of α -RuCl₃ well-above a temperature of a magnetically ordered phase supports the theoretical view of a $J_{\text{eff}}=1/2$ Mott insulator produced by an interplay of electron correlations and spin-orbit coupling. Second, we observed the emergence of a charge-ordered pattern originating from anisotropy in the charge distribution along Ru–Cl–Ru hopping pathways. The charge order appears to have a fine structure, in which two types of the electronic superlattices, dimer-like and $R3$ trigonal, coexist at the nanometre length scale. One of the intriguing questions for future studies is: What are the exact implications of the charge distribution at the Cl- ligands on the spin degrees of freedom in the Kitaev limit? Third, our finding of a structural inhomogeneity in the ligand sublattice is significant as it may result in a unit-cell scale inhomogeneity of the magnetic ground state. Finally, our discussion on potential interplay between charge, lattice and spin degrees of freedom hinted at the possibility of nanoscale strain engineering of magnetic and electronic properties in α -RuCl₃, which would enable applications in hybrid spintronics and

straintronics devices. In the future, we expect that extensions of this work using heterojunction techniques, as well as thin films with mis-oriented RuCl_3 layers [96], may provide a pathway to drive $\alpha\text{-RuCl}_3$ or related materials towards the pure Kitaev limit where the true quantum-spin-liquid ground state flourishes.

2.2.7 Methods

Sample preparation (bulk crystals)

Commercial- RuCl_3 powder was purified in-house to a mixture of $\alpha\text{-RuCl}_3$ and $\beta\text{-RuCl}_3$ and converted to 99.9% phase pure $\alpha\text{-RuCl}_3$ after annealing at 500 °C. Single crystals of $\alpha\text{-RuCl}_3$ were grown by vapour transport with TeCl_4 as the transport agent.

Neutron scattering

Neutron scattering experiments and the data in Fig. 2-2 were obtained using the HB-3A Four-Circle neutron diffractometer at High Flux Isotope Reactor at ORNL using an incident energy $E_i=14.7$ meV on a large single crystal ($1\text{ cm} \times 1\text{ cm} \times 0.3\text{ mm}$) mounted on the top of the Goniometer head on a 4 K displacer. The data has been analysed using FullProf structural refinement routine, and further confirmed using X-ray diffraction measurements refined using ShelX program. Above 200 K, the structure is consistent with the $P3_1$ space group, while at low temperatures the structures are consistent with $C2/m$ proposed in ref. [73]. Interestingly enough, the phase transition was not observed in powder and smaller samples which remain in the same space group at all temperatures.

STEM experiment

The films of $\alpha\text{-RuCl}_3$ for STEM measurements were exfoliated by sonication of bulk samples. STEM-annular dark field imaging was performed on an aberration-corrected Nion UltraSTEM-100 microscope operated at 60 kV. Only the images without noticeable electron beam induced structural damage were used for analysis. The STEM images used in analysis were summed over 30 frames after cross-correlation.

STM experiment

The STM experiments were performed using an Omicron VT-STM equipped with a Nanonis controller in an operating pressure of 2×10^{-10} Torr. The samples for the STM experiments were cleaved in the ultra-high vacuum ($\sim 10^{-10}$ Torr). The STM images were obtained with mechanically cut Pt/Ir alloy tips with electrically biased tips and grounded samples. Due to the insulating nature of α -RuCl₃, the STS measurements were performed at $T=295$ K, using a standard lock-in technique, with high-tunneling junction resistance $R_J=20\text{--}25\ \Omega$. To ensure a reproducibility of the high-temperature STM spectroscopic measurements, the dI/dU curves were recorded at clean surface areas (that is, with no apparent defects) in several macroscopically different regions of each sample and the resultant averaged data was compared across the samples from three different batches.

DFT calculations

The DFT calculations were performed within the generalized gradient approximation using the Perdew–Burke–Ernzerhof exchange correlation scheme [97] and projector augmented wave potentials^[98] as implemented in VASP [99, 100]. We used a 400 eV kinetic energy cutoff and $7 \times 7 \times 3$ Monkhorst-Pack type k-point grid [101] to calculate the relaxed atomic geometries. The internal forces were relaxed to $<10\text{ meV \AA}^{-1}$ without enforcing any symmetry. The external lattice parameters were fixed to the experimental values reported in ref. [75].

Statistical and image analysis

The principal component analysis of STEM data, construction of electron densities lattice from STM image, and analysis of atomic distances were performed using MATLAB software.

Data availability

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

2.3 Knowledge Extraction from Atomically Resolved Images

2.3.1 *Publication Information*

A version of this article was originally published by Lukas Vlcek, Artem Maksov, Minghu Pan, Rama K Vasudevan, Sergei V Kalinin in:

Vlcek, Lukas, Artem Maksov, Minghu Pan, Rama K. Vasudevan, and Sergei V. Kalinin. "Knowledge Extraction from Atomically Resolved Images." *ACS nano* 11, no. 10 (2017): 10313-10320. [102]

The article content is reprinted as published with figure numbers modified and supplementary material added at the end of the chapter as an appendix.

2.3.2 *My contribution*

The goal of this study was to pave the way for incorporation of generative theoretical models into the analysis of imaging data, which can reveal information only implicitly contained within the images. I identified locations of all atoms within the image by extending the lattice (under the assumption of square lattice) and applying corrections based on local intensities in order to account for image drift and local disorder/artifacts. I further collected atomic intensities and classified the atoms by optimizing intensity threshold to correspond to the best match to the expected stoichiometry. This allowed for full information extraction from the image and further usage of this data with 3D model simulations and statistical distance minimization to identify the best parameters for the model and predicting 3D structures of the observed data based on them.

2.3.3 Abstract

Tremendous strides in experimental capabilities of scanning transmission electron microscopy and scanning tunneling microscopy (STM) over the past 30 years made atomically resolved imaging routine. However, consistent integration and use of atomically resolved data with generative models is unavailable, so information on local thermodynamics and other microscopic driving forces encoded in the observed atomic configurations remains hidden. Here, we present a framework based on statistical distance minimization to consistently utilize the information available from atomic configurations obtained from an atomically resolved image and extract meaningful physical interaction parameters. We illustrate the applicability of the framework on an STM image of a $\text{FeSe}_x\text{Te}_{1-x}$ superconductor, with the segregation of the chalcogen atoms investigated using a nonideal interacting solid solution model. This universal method makes full use of the microscopic degrees of freedom sampled in an atomically resolved image and can be extended via Bayesian inference toward unbiased model selection with uncertainty quantification.

2.3.4 Introduction

Direct imaging of atomic structures has remained the scientific dream for millennia. Ever since Democritus proposed the existence of atoms as the smallest indivisible units of matter, this idea and its consequences have captured the attention of the scientific community. In the 19th century, the broad adoption of atomistic theory led to remarkable progress in physics and chemistry and eventually to the dawn of quantum physics in the early 20th century. Soon after, the work by Ruska and others laid the foundation for electron microscopy, which by the 1970s provided reliable images of atomic structures under certain special conditions [103-105]. Finally, in the last 30 years, the progress in scanning transmission electron microscopy (STEM) and scanning tunneling microscopy (STM) has rendered observations of atomic structures routine [106].

Despite this progress, the amount of information extracted from the atomically resolved images remains extremely small. In virtually all cases until a decade ago, the images were

interpreted on a purely qualitative level—in terms of observed structural defects, atomic arrangement on grain boundaries, dislocation cores, or interfaces, presence of grain and phase boundaries or topological defects, while more quantitative interpretation of the observed features, was confined to comparisons with the results of simulations with pre-existing atomistic force fields [107]. The advent of aberration correction in STEM has allowed rapid, high veracity imaging when not only compositions of atomic columns but also small deviations from high symmetry positions and changes in bond angles and length become quantifiable. Similar progress is being achieved in scanning tunneling microscopy, visualizing the atomic structure of metal and semiconductor surfaces. However, by now, the preponderant analysis framework remains the mapping of the microscopic degrees of freedom on the mesoscopic order parameter fields such as polarization [108-111], chemical expansion [112, 113], and octahedral tilts [114-117]. However, these methods do not integrate knowledge from theoretical frameworks and are not data-driven, limiting their potential utility. In addition to static imaging, dynamic information on vacancy and atomic diffusion on surfaces and in the bulk for many material systems are readily available; however, typically, the main method of analysis of these data is to observe the presence or absence of specific configurations (ordering) or to estimate diffusion barriers [118-120]. It should be noted that, in recent times, both machine learning and the use of Bayesian inference have been proposed to extract more meaningful information from atomically resolved data sets [121], such as in the example of Ishikawa *et al.* [122], who used Bayesian statistical approaches in conjunction with simulations to identify the location of Ce dopant atoms in AlN, and Cossio *et al.* [123], who used Bayesian inference to rank structural models from cryo-electron microscopy data. Because these approaches are intended for identification of likely atomic structures that give rise to the observed imaging data, they can be considered a part of the image analysis. However, it is rarely the case that a particular instance of an imaged structure is the ultimate target of experimental research. More often, the goal is to understand the properties and behavior of the underlying physical system giving rise to the observed structures. Such knowledge in the form of a suitable computationally efficient model can be then used to extract information only implicitly contained in the imaging data and to make quantitative predictions facilitating materials

design. Even though large-scale atomistic simulations [124] and simplified statistical mechanical models [125] have been recently used to explain the basic mechanisms behind structures observed in images, the experimental data were used as a target to be interpreted rather than as a direct source of new information to further improve our predictive abilities beyond that of the existing models. Also, no metrics evaluating the model quality were provided. To this date, the rich statistical information contained in the spatially resolved atomic configurations remains largely unused. Hence, presently the scientific community faces the unique situation when the capability to map the atomic structures and physical functionalities on the atomic level exists; however, the framework to convert quantitatively and rigorously these observations to materials-specific knowledge is absent. That is, most analysis is limited to simple qualitative comparisons between imaging data and models, and efforts to improve models based on available imaging data have not been reported.

Outline of the General Framework

Here, we develop a general framework for the incorporation of imaging and other complementary experimental data into predictive statistical mechanical models of solid-state materials and demonstrate it on a nonideal solid solution model as applied to the $\text{FeSe}_x\text{Te}_{1-x}$ superconductors. However, the proposed approach, which is schematically presented in Figure 2-6 and described in the following text, is general and can be applied to any imaging data and generative physical model, as long as statistical information can be collected. We apply the method toward analyzing segregation of chalcogen atoms in an $\text{FeSe}_x\text{Te}_{1-x}$ superconductor, with imaging data from low-temperature atomically resolved scanning tunneling microscopy. By utilizing the data present in the images, *i.e.*, the local atomic configurations, and building a histogram, we are able to use this information toward optimization of a microscopic model through a combination of statistical distance minimization and perturbation technique. This study shows a general framework for going beyond standard image analysis and qualitative matching of imaging data with theoretical simulations and toward quantitative use of all available statistical information from the imaging data to develop microscopic models with predictive capability.

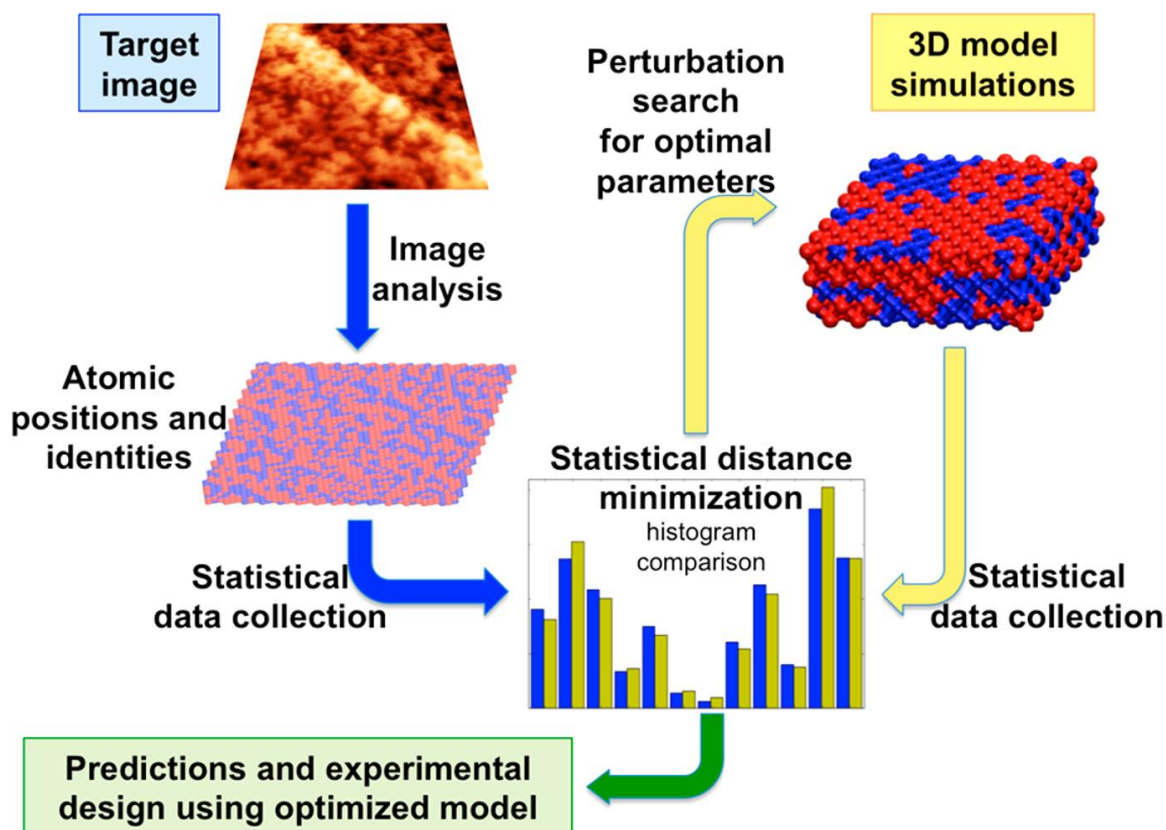


Figure 2-6. Algorithm schematic.

Schematic representation of the proposed framework for model derivation from imaging data, with statistical distance forming the interface between the experimental data and their theoretical counterpart.

2.3.5 Results and Discussion

As a target for this study, we have chosen $\text{FeSe}_x\text{Te}_{1-x}$ materials (with $x = 0.45$), extensively studied in the context of its superconductive behaviors due to its cleavability. Shown in Figure 2-7 is the high-resolution STM image of the surface obtained at 4 K. Direct examination of the image reveals clearly visible bright and dark atoms, corresponding to Se and Te atoms, respectively. The elemental distribution in the system is clearly nonrandom, with the atoms tending to form clusters of the same type. This behavior is well-known in these and similar systems, and it has been hypothesized that it may be the result of the spinodal decomposition or nucleation and growth [126, 127], but the actual mechanism can be expected to vary depending on the conditions of preparation. In the present case, the material preparation involved slow cooling from melt and long equilibration at elevated temperatures (4 days at 400 °C), which suggests that at least the statistics of local structures will reflect the equilibrium Boltzmann distribution. Here, we aim to extract the corresponding thermodynamic parameters from an observed image and infer system properties not directly visible in the data.

As the first step of the analysis, we identify the location of all atoms in the image by extending the lattice and applying corrections based on visual identification under the assumption of a perfect square lattice [128]. Next, for each identified position, we collect the total intensity of a region around it and perform the thresholding of atomic intensities to classify the atoms in accordance with the expected stoichiometry of the material (Figures 2-25 and 2-26). Finally, knowing the positions and types of all atoms, we transform the data from the coordinate system of the image to that of the $\text{FeSe}_{0.45}\text{Te}_{0.55}$ crystal lattice, *i.e.*, associate each atom with a specific position on a square lattice.

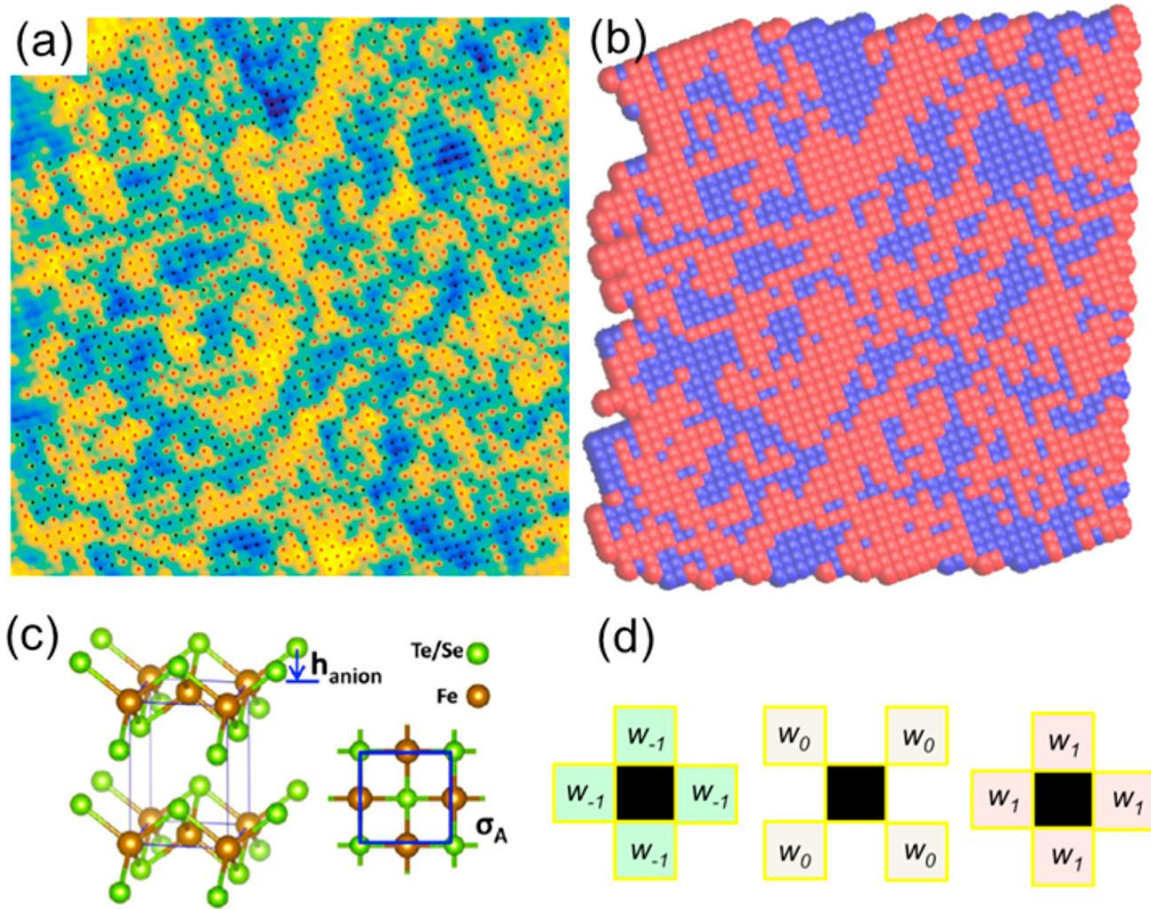


Figure 2-7. Structural Information of $\text{FeSe}_{0.45}\text{Te}_{0.55}$.

(a) STM image of the $\text{FeSe}_{0.45}\text{Te}_{0.55}$ surface exposing the chalcogen layer. (b) Reconstruction of atomic types and positions from image analysis covering approximately 50×50 atoms; Se (blue), Te (red). (c) Structure of the bulk $\text{FeSe}_x\text{Te}_{1-x}$ crystal lattice. (d) Geometry of model interactions defined by eq 2.1, between Se/Te sites (black) and the neighboring Te/Se sites within the same layer, w_0 , and the layers below, w_{-1} , and above, w_1 .

Model Definition

To understand the origin of the concentration fluctuations observed in the image, we adopt a statistical mechanical model of microscopic interactions between Se and Te atoms that would allow us to extract information encoded in the image in a condensed form suitable for making predictions. These fluctuations, which the model aims to capture, also implicitly determine the system's response to external conditions as follows from the linear response theory [129]. The experimental information is supplemented with our *a priori* knowledge of the $\text{FeSe}_x\text{Te}_{1-x}$ structure, which is built into the geometric arrangement of interaction sites mimicking the spatial distribution of the chalcogen atoms within the crystal lattice (Figure 2-7c). We also assume that the driving forces behind the elemental segregation can be represented by short-range interactions. This is supported by the correlation length $\xi \approx 1$ in the units of the nearest-neighbor distance, which was determined from the image by approximating the Te–Te correlation function with $\exp[-|i-j|/\xi]$, where i and j are the positions of Te sites. We note that these effective interactions do not only represent the energetics of individual atom–atom interactions but also include entropic contributions that may be related to lattice distortions and other larger scale effects. As can be seen from Figure 2-7c, each chalcogen (Se,Te) layer has two types of neighboring chalcogen layers: one that is separated by an Fe layer and one that is directly adjacent. As we do not have previous knowledge of the relative strength of interactions between the different layers, we define three independent adjustable parameters controlling the strength of cross interactions between unlike atoms: one for the same layer, w_0 , and two for the distinct neighboring layers, w_1 and w_{-1} (Figure 2-7d). The energy of configuration i in the model system is given as

$$u_i = w_0^p a_i^0 + w_{-1}^p a_i^{-1} + w_1^p a_i^1, \quad (2.1)$$

where w_x^p are the interaction parameters in system p , and quantities a_i^x are defined as

$$a_i^0 = \sum_{NN} \delta^0; a_i^{-1} = \sum_{NN} \delta^{-1}; a_i^1 = \sum_{NN} \delta^1, \quad (2.2)$$

with the summations running through all nearest-neighbor (NN) pairs of atomic sites; the quantity δ^x is equal to either one if the interaction between particles is defined by parameter w_x or zero otherwise. Monte Carlo simulations at constant temperature and the relative concentration of Se and Te atoms fixed at the bulk stoichiometry ($x = 0.45$) are then used to generate equilibrium configurations that are to be matched with the observed data. We assume that, while some nonequilibrium features may persist in the material at larger scales, thermodynamic forces drive the local ordering of atomic configurations. Note that the model chosen here represents a basic lattice model; more complex sets of interactions or models incorporating additional known information can be considered.

Structure Analysis

To establish the interaction parameters controlling segregation, we analyzed the statistics of atomic configurations. Even though an image represents a single state of the experimental system, which by itself is not sufficient to evaluate thermodynamic properties and construct statistical mechanical models, it is possible to evaluate the range of correlations within the observed configuration and collect the statistics of uncorrelated local structural descriptors as if they originated from many independent states. The interaction model can be then optimized to reproduce this statistical information, which is collected in the form of histograms of different atomic configurations. In the present case, the target information consisted of the relative frequency of 12 local surface configurations encompassing all possible arrangements of the nearest atomic neighbors, as shown in Figure 2-8. The standard deviation of the relative frequency values for each histogram bin, indicated by error bars, was estimated from the average count variations among 9 blocks into which the image was divided for this purpose. To test whether new information can be extracted from the image, we also tried to match 60 local configurations of the nearest and next-nearest neighbors. However, the extra information did not provide significant constraints on the model; therefore, we can conclude that the formation of the observed structure is predominantly driven by nearest-neighbor interactions.

When comparing two systems (*i.e.*, the experimental target and its model), using statistical data from measurements on these systems, special care must be taken to use a proper

measure of their similarity to avoid biased results. Here, we base the model optimization loss function on a metric known in physics as statistical distance [130-132], a similarity measure quantifying the distinguishability of two physical systems based on measurements collected as sets of individual samples. It can be written as

$$s = \arccos \left(\sum_{i=1}^k \sqrt{p_i} \sqrt{q_i} \right), \quad (2.3)$$

where p_i and q_i are the probabilities of finding outcome i in the measurement of systems p and q , respectively, with the total of k possible outcomes. The distance takes into account that both systems are the source of sampling uncertainty and builds in a variance equalizing transformation. As a result, different sources of data can be treated on equal footing [133, 134]. In the present case, the two systems correspond to $\text{FeSe}_{0.45}\text{Te}_{0.55}$ probed by imaging and simulations. The probability distributions are determined as the maximum likelihood estimates equal to the relative frequencies of the histogram of k local configurations in the two systems (Figure 2-8).

Parameter Optimization

The optimal model parameters were found using the computationally efficient perturbation technique [132, 135], which takes advantage of the thermodynamic perturbation relations to predict the changes of arbitrary properties A of system p as a function of interaction parameters. This is achieved by sampling a reference system q and reweighting contributions A_i to property A from states i according to

$$\langle A_i \rangle_p = \sum_{i=1}^k A_i p_i = \sum_{i=1}^k A_i \left(\frac{p_i}{q_i} \right) q_i = \langle A_i \exp[-(\beta \Delta u_i - \Delta F)] \rangle_q, \quad (2.4)$$

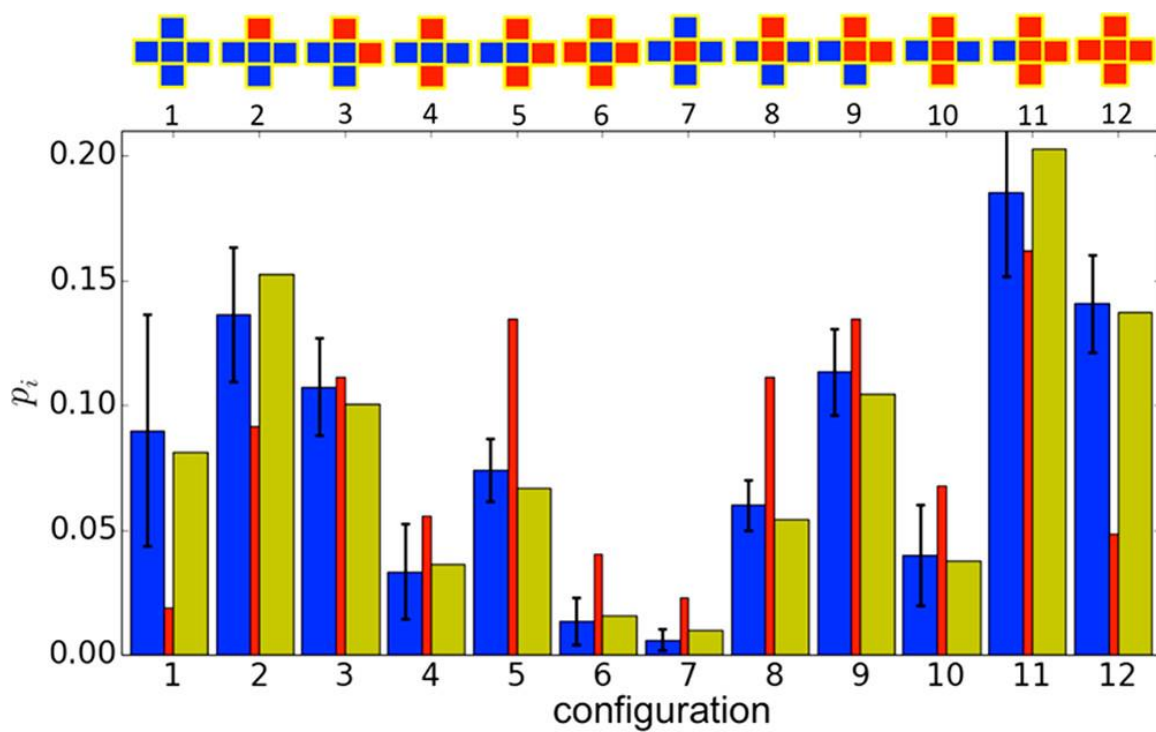


Figure 2-8. Statistics of local configurations.

Top: Local configurations collected from the image whose statistics were to be matched by the model, where Se atoms are blue and Te are red. Bottom: Histograms of the target (blue) and model (yellow) local configurations. The statistics of local configurations generated by a noninteracting model are shown as red.

where the angle brackets denote ensemble averaging, Δu_i denotes the energy difference of state i in systems p and q , respectively, ΔF is the Helmholtz free energy difference between the two systems, and $\beta = 1/k_B T$, where k_B is the Boltzmann constant and T is absolute temperature. In the present case, A_i corresponds to a contribution to the histogram of local configurations from state i generated in the simulation of the reference system q . The predicted histograms are then compared with the target data using the statistical distance, and the best match is selected as the optimal model.

For lattice models with nearest-neighbor interactions defined by eq 2.1, the energy difference Δu_i can be expressed as

$$\Delta u_i = (w_0^p - w_0^q)a_i^0 + (w_{-1}^p - w_{-1}^q)a_i^{-1} + (w_1^p - w_1^q)a_i^1, \quad (2.5)$$

where w_x^p and w_x^q are the interaction parameters in systems p and q , respectively. Once quantities a_i^x defined by eq 2.2 are collected from a reference simulation of system q , the perturbation scan of the parameter space is executed by multiplying the arrays of a_i^x values by different parameter combinations, and then eqs 2.4 and 2.5 are used to predict the histograms of local configurations without the need for repeated simulations. The search for a parameter set minimizing the statistical distance, eq 2.3, between the target and model histograms can be initiated by a matrix search and followed by local searches starting from the most promising candidates using the conjugate gradient method [136]. In the present case, we performed a Monte Carlo simulation of a reference system with all interaction parameters set to zero and used it to determine the parameters of a trial model, which was in turn used as a new reference system. This procedure was repeated in three iterations before reaching the optimal values: $w_0 = 0.529$, $w_1 = 0.023$, and $w_{-1} = -0.040$ in the units of $k_B T$. The agreement between the histograms of local configurations collected from the image and from modeling is within the standard deviation indicated by error bars in Figure 2-8.

Model Predictions and Interpretation

The resulting model sheds more light on the elemental distribution in the material and its origins. The fact that the statistics can be well reproduced by an equilibrium model strongly suggests that the experimental structures are, indeed, the features of an equilibrated system. The energetically unfavorable positive values of the optimized parameters for the unlike pair Se–Te interactions reflect the observed segregation. The larger magnitude of the same layer interactions, w_0 , relative to the interlayer interactions suggests that the segregation is driven mostly by forces acting between atoms within individual layers, implying that correlations across multiple layers are substantially weaker. In fact, a simple two-dimensional model with a single interaction parameter w_0 can provide nearly the same quality of match between the histograms and capture the main driving force behind the observed data. Given this result, we can hypothesize that most segregation occurs only in the newly built layers during crystallization. If we assume that the crystallization occurred at 400 °C, at which the system was equilibrated for 4 days, the energy of the effective same-layer interaction determined by w_0 is approximately 3.0 kJ/mol, which represents a penalty for the formation of a Se–Te nearest-neighbor pair relative to Se–Se and Te–Te pairs. To gain better insight into the constraints on individual interaction parameters provided by the imaging data, we show the landscapes of statistical distance as a function of the parameters in Figure 2-9, which provide a more comprehensive view of the parameter space than a simple sensitivity analysis. The contour plots, generated using the perturbation technique based on a single reference simulation, show that the valleys near the minimum distance are narrow along the horizontal w_0 direction, while elongated along the vertical w_{-1} and w_1 directions. This indicates better constraints on the same-layer interaction parameter w_0 . Such an outcome could be expected given that the target imaging data provide direct information about intralayer correlations. The model also allows us to estimate the system's thermodynamic and structural properties as a function of temperature and stoichiometry, formulate quantitative predictions of segregation, clustering, or percolation, which may relate to material's properties,

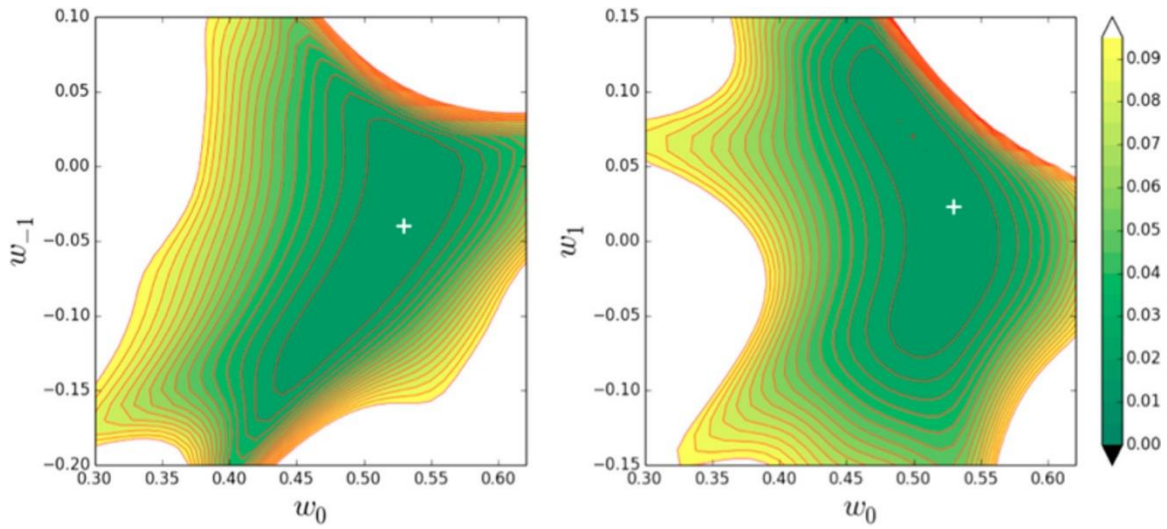


Figure 2-9. Landscapes of squared statistical distance, s^2 , as a function of the three interaction parameters of the considered model.

Left: Space of parameters w_0 and w_{-1} for the optimal $w_1 = 0.023$. Right: Space of parameters w_0 and w_1 for the optimal $w_{-1} = -0.040$. The white cross indicates the combination of optimal parameters.

such as conductivity, and identify new targets for further experimental study. Figure 2-10 illustrates the changes in atomic clustering and segregation as a function of temperature and the profiles of the free energy and entropy contributions from elemental segregation as a function of temperature and composition. The latter plots indicate the possibility of phase separation suggested by the abrupt decrease in the entropic contribution seen at lower temperatures and stoichiometry around $x \sim 0.5$.

2.3.6 Conclusions

Here, we have developed an approach for consistent integration of the knowledge contained in atomic configurations derived from microscopic images into a model of a nonideal solid solution and extracted interaction parameters between the two types of different constituent atoms. Although for this system our previous study has shown that the impact of segregation on the conductance spectra is relatively small, the presented framework is general and can be applied to any atomically resolved imaging data. These may include direct information on interactions between vacancies, adatoms, and dopants which can be otherwise difficult to compute from first-principles theories (especially in, *e.g.*, technologically relevant oxides [137, 138]), or measured by experimental techniques (*e.g.*, diffraction) [139, 140]. Thus, the framework presented here offers a method to learn information contained in the microscopic degrees of freedom sampled during experiments and, in principle, enables exhaustive utilization and exploitation of the available statistical information. As long as a generative model is known (or can be postulated based on previous knowledge), the method shown in this communication can be universally applied. This also includes the incorporation of data from first-principles simulations, which can be regarded as pseudoexperiments providing series of configuration–energy samples at the microstate resolution. As shown earlier [132], this quantum chemical input can be used in a form fully consistent with the type of experimental data discussed here, with these diverse complementary sources of information providing joint constraints on the model. While we have used an example of an equilibrium system, the framework is equally applicable to nonequilibrium systems as long as enough statistical

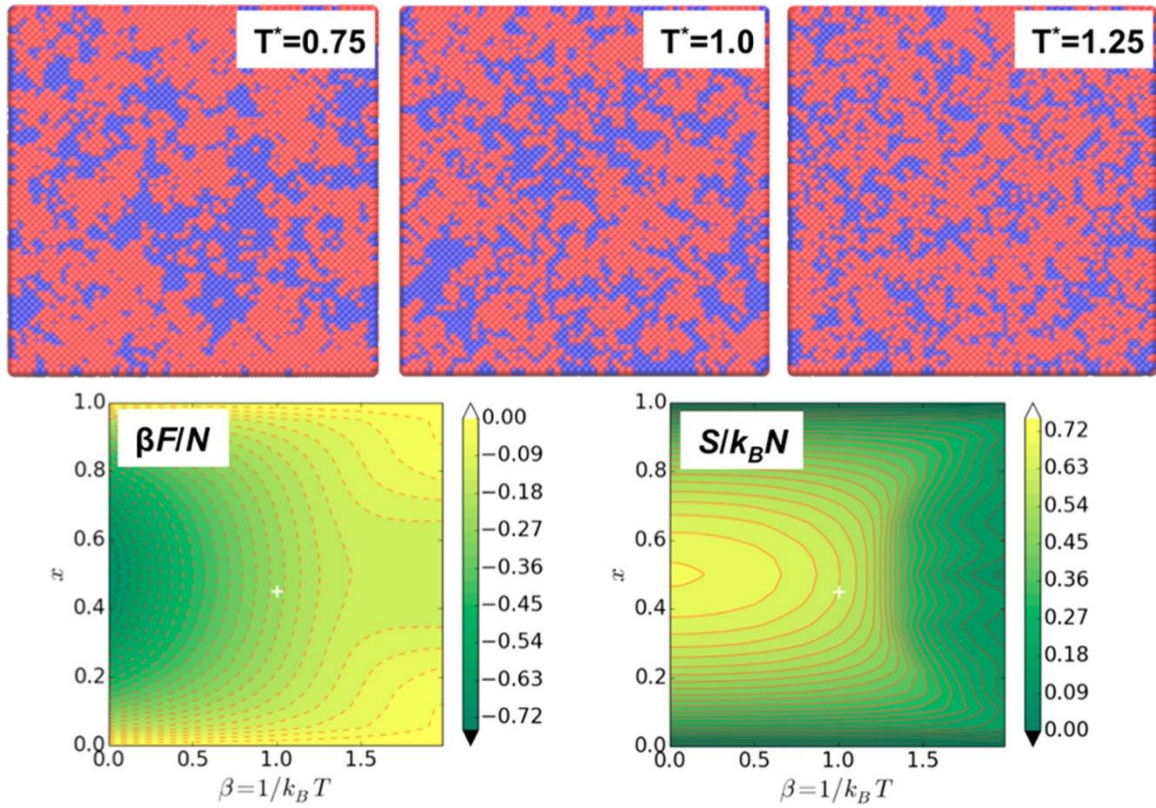


Figure 2-10. Optimal configurations and results per particle.

Top: Typical surface atom configurations generated by the optimal model for Se concentration $x = 0.45$ at reduced temperatures $T^* = 1$ (consistent with the experimental data) and at temperatures $T^* = 0.75$ and 1.25 . The area covers 100×100 surface atoms. Bottom: Free energy $\beta F/N$ and entropy $S/k_B N$ contributions per particle originating from elemental segregation as a function of temperature and composition. The set of experimental conditions is indicated by a white cross.

data as a function of time are collected. Further improvements can be made based on combination with Bayesian inference [141] to determine the likelihood of particular models with quantified uncertainty in a nonbiased manner. Such techniques are vital to guide researchers to help design future experiments that can reduce the model uncertainty further, thereby increasing efficiency and creating the feedback between theory and experiment that is highly data-driven.

2.3.7 Methods

The STM data on the $\text{FeSe}_x\text{Te}_{1-x}$ system was acquired on a home-built STM system, on a crystal of composition $\text{FeSe}_{0.45}\text{Te}_{0.55}$ cleaved in UHV conditions at room temperature and imaged in constant current mode at 77 and 4.2 K. The crystals are tetragonal and cleave perpendicular to the c -axis, exposing the ab -plane for imaging. More details about the experimental procedure are available in a previous publication [128].

Simulation Details

The Monte Carlo simulations of the model systems were performed on a simple cubic lattice. To conform to the geometry of the $\text{FeSe}_x\text{Te}_{1-x}$ structure, the Se and Te atoms were placed on the same color sites of a three-dimensional checkerboard; that is, the 12 nearest neighbors of Se or Te atoms shared the cubic edges (Figure 2-7d). Simulations of the reference systems intended for the perturbation technique parameter optimization were performed in a simulation box with 16 lattice sites in the directions parallel with the Se/Te layers and six lattice sites in the perpendicular direction. Additional simulations, which exceed the area covered by the image, were performed in a box with 100×100 lattice sites in the parallel direction and six sites in the perpendicular one. At each Monte Carlo step, the positions of two randomly chosen particles were exchanged, and the move was accepted or rejected following the Metropolis acceptance condition. The simulations were performed at reduced temperature $T^* = 1.0$ over 10^6 sweeps, where each sweep consisted of N individual steps, with N equal to the total number of particles in the system. Samples were collected every 10 sweeps, resulting in the collection of 10^5 system configurations,

which were further analyzed to obtain the statistics of local configurations used for optimization. The plots of free energy and entropy surfaces were created using the multistate Bennett's acceptance ratio method [142] based on a set of simulations at 10 different values of Se concentration x between 0.05 and 0.5 and eight different values of inverse reduced temperature β between 0 and 2.0.

2.4 Deep data mining in a real space: separation of intertwined electronic responses in a lightly doped BaFe₂As₂

2.4.1 Publication Information

A version of this article was originally published by Maxim Ziatdinov, Artem Maksov, Li Li, Athena S Sefat, Petro Maksymovych and Sergei V Kalinin in:

Ziatdinov, Maxim, Artem Maksov, Li Li, Athena S. Sefat, Petro Maksymovych, and Sergei V. Kalinin. "Deep data mining in a real space: separation of intertwined electronic responses in a lightly doped BaFe₂As₂." *Nanotechnology* 27, no. 47 (2016): 475706. [143]

The article content is reprinted as published with figure numbers modified and footnote referencing changed. Supplementary material added at the end of the chapter as an appendix.

2.4.2 My contribution

Strongly inhomogeneous electronic response of BaFe₂As₂ doped with 1% gold prevents recovering atomic positions from topographic data and results in a very difficult to analyze spectral data. In order to study multidimensional STM/S data reliably, a reliable and unbiased identification and spatial mapping of statistically significant electronic behaviors is necessary.

The subject STS dataset has dimensions of 80 by 80 by 768, containing a stack of 768 conductance maps. I preprocessed the data to convert it into 3D matrix. I applied PCA and k-means clustering algorithms to estimate number of significant sources of signal within the data and explore the data. I used Bayesian linear unmixing (BLU) on the data to unmix the signals under non-negativity and full-additivity physical constraints. Compared to PCA and k-means clustering, this produces signal sources that are more physically meaningful, as well as addresses the issue of high spatial inhomogeneity of data.

In additional work, presented at MRS and APS conferences, but not included in this publication, I have also utilized local indicators of spatial association (LISA) to identify correlation between unmixed electronic signal sources and topography of the system.

2.4.3 Abstract

Electronic interactions present in material compositions close to the superconducting dome play a key role in the manifestation of high- T_c superconductivity. In many correlated electron systems, however, the parent or underdoped states exhibit strongly inhomogeneous electronic landscape at the nanoscale that may be associated with competing, coexisting, or intertwined chemical disorder, strain, magnetic, and structural order parameters. Here we demonstrate an approach based on a combination of scanning tunneling microscopy/spectroscopy and advanced statistical learning for an automatic separation and extraction of statistically significant electronic behaviors in the spin density wave regime of a lightly ($\sim 1\%$) gold-doped BaFe_2As_2 . We show that the decomposed STS spectral features have a direct relevance to fundamental physical properties of the system, such as SDW-induced gap, pseudogap-like state, and impurity resonance states.

2.4.4 Introduction

Nanoscale inhomogeneity of chemical, structural and electronic orders in a crystalline matter is expected to have a profound and non-random effect on the macroscopic properties

of technologically relevant materials. Celebrated examples include reduced mobility of Dirac electrons in graphene transistor devices due to formation of charge nanopuddles [31, 144], ultra-high piezoelectric response of relaxor ferroelectrics due to interaction between nanopolar domains and acoustic phonon mode [32], filamentary superconductivity [145], and fluctuating superconducting (SC) state above a transition temperature (T_c) in high- T_c cuprates associated with emergence of nanometre-sized electron pairing regions [146].

Scanning tunneling microscopy and spectroscopy (STM/S), which probes topographic and electronic properties of the surfaces with a nanometer-scale resolution, constitutes an ideal experimental tool for exploring local inhomogeneity in materials. The STM topographic images are typically recorded in a constant current regime [147], resulting in a two-dimensional (2D) $Z(X, Y)$ dataset, where Z represents a convolution of actual height variation and electronic local density of states (LDOS) at each point (X, Y) on the surface. Meanwhile, the STS mode allows to acquire 3D $G(X, Y, V)$ datasets, where $G = dI/dV$ corresponds to a value of differential conductance proportional to LDOS at specific energy $E = eV$ at each (X, Y) point. In the simplest realizations of the bi-phase or multi-phase nanoscale systems, a separation between two or more phases is clearly visible in the STM topography, and comparison of STS spectra associated with different topographic features allows a straightforward analysis of electronic properties in these phases. Examples include STM/S measured on 2D lateral heterostructures sufficiently far from the boundary [148] or STM/S experiments on an isolated impurity embedded in otherwise ideal lattice [149].

For many strongly correlated materials, however, a complex local inhomogeneity patterns in conductance maps do not have a direct and simple connection to topographic features [150-152]. To complicate things even further, the morphology and chemical composition of the top-most layer of a cleaved surface in many complex compounds is usually itself a subject of controversy [153] which makes it nearly impossible to predict electronic properties in a characteristic field of view of STM/S experiment from the first principles. Given an ever-growing amount of multidimensional STM/S data on strongly correlated materials [153-155], there is an urgent need for developing a deep data driven analysis that would allow reliable and un-biased identification and spatial mapping of statistically significant different electronic behaviors without *a priori* knowledge about the details of

surface structure. Here we present a physics-robust statistical learning approach based on k -means clustering, principal component analysis (PCA) and Bayesian linear unmixing (BLU) to uncover a wealth of 'hidden' information from the STS measurements in a lightly doped, 'precursor', magnetic regime of iron-based superconductor [156]. We show how the features extracted from multivariate statistics-based decomposition of STS signal have a direct relevance to fundamental physical properties of the system, which we illustrate by uncovering a 'buried' pseudogap-like phase and impurity induced double resonance states. As a model system, we have chosen a lightly Au-doped BaFe_2As_2 single crystal, $\text{Ba}(\text{Fe}_{1-x}\text{Au}_x)_2\text{As}_2$ with $x = 0.009$ ($\sim 1\%$). This compound shows a coupled structural and antiferromagnetic (AF) transition, from the tetragonal non-magnetic state into the orthorhombic striped SDW phase at $T_N \approx 110$ K [157]. Upon increased Au-dopants, the AF interactions becomes suppressed and the system develops into a superconductor ($T_c \approx 4$ K) at $\sim 3\%$. It has been recognized that interactions present in such SDW states of the FeAs-based compounds play a crucial role in understanding unconventional superconductivity [158, 159]. However, the details of local electronic structure at low temperatures in the non-SC phase of FeAs compounds, including the role of lattice strain, presence and origin of a pseudogap-like state, and character of impurity-induced quasiparticle states, remain a subject of a debate [153, 159].

2.4.5 Results and discussion

We first present the STM topographic image over a relatively large FOV on a cleaved surface of 1% Au-doped BaFe_2As_2 in the SDW phase recorded at $T = 77$ K (figure 2-11(a)). The typical surface area at 77 K appears to be peppered with dark nanoscale regions. Upon cooling down to $T = 4$ K, we found a dramatic increase in the density of the dark nanoregions as can clearly be seen from the representative STM topographic image for the same Au-doped BaFe_2As_2 surface in figure 2-11(b). In general, the variations in apparent topographic height associated with dark and bright regions can be of both topographic and electronic origin. However, we do not expect any extensive surface damage or profound

changes in nanoscale chemical composition as we cool down the sample from 77 to 4 K. Instead, the observed change in STM topographic patterns in figures 2-6(a) and (b) is likely related to the enhanced nanoscale electronic inhomogeneity as we approach towards a phase region with competing normal and SC orders [160] or with admixture of another form of magnetic order within the SDW phase (see supplemental material). Such inhomogeneity shows the necessity for applying data mining tools based on multivariate statistical analysis for extracting relevant electronic behaviors in this system [161].

Zooming into a smaller FOV reveals a stripe-like reconstruction at the surface with a periodicity across the stripes of ≈ 0.7 nm (inset in figure 2-11(b)). Similar unidirectional modulation of charge density has been also reported for SDW phase of SrFe_2As_2 [162]. While the exact origin of these charge stripes and their relation to SDW, if any, is not clear at present moment, it is worth to note that we were not able to observe similar 1D modulations at 77 K on the same cleaved surface. This suggests that the reconstruction is not cleavage-induced.

In figure 2-11(c) we show the STM topography at $T = 4$ K measured in a region with extended quasi-1D defect which appears as a bright 'diagonal' feature in the topography. The spatial extension of this defect typically exceeds $\approx 1 \mu\text{m}$ and we were able to reproducibly observe it in several areas of the sample. Furthermore, a similar structure was reported in another iron-based superconductor compound [128] suggesting that this defect can be a common feature of iron pnictides. High-resolution spatial maps of differential conductance G (figures 2-11(d)–(g)) recorded at the area shown in figure 2-11(c) at several selective energies confirm a highly inhomogeneous electronic structure of the surface, with no one-to-one correspondence to underlying topographic data. Accordingly, averaging over even relatively small surface area can lead to a loss of significant physical information contained in individual G curves. Such a lossy compression of the original data is illustrated in figure 2-12(b), where the STS curve averaged from 182 individual line spectra inside the box in figure 2-12(a) fails to reproduce physically important features at the Fermi level seen in the 4 selected single-pixel (point) spectra recorded within the box area. It therefore becomes clear that the surface electronic behavior cannot be characterized reliably by a simple visual assessment of the topographic image and individual inspection

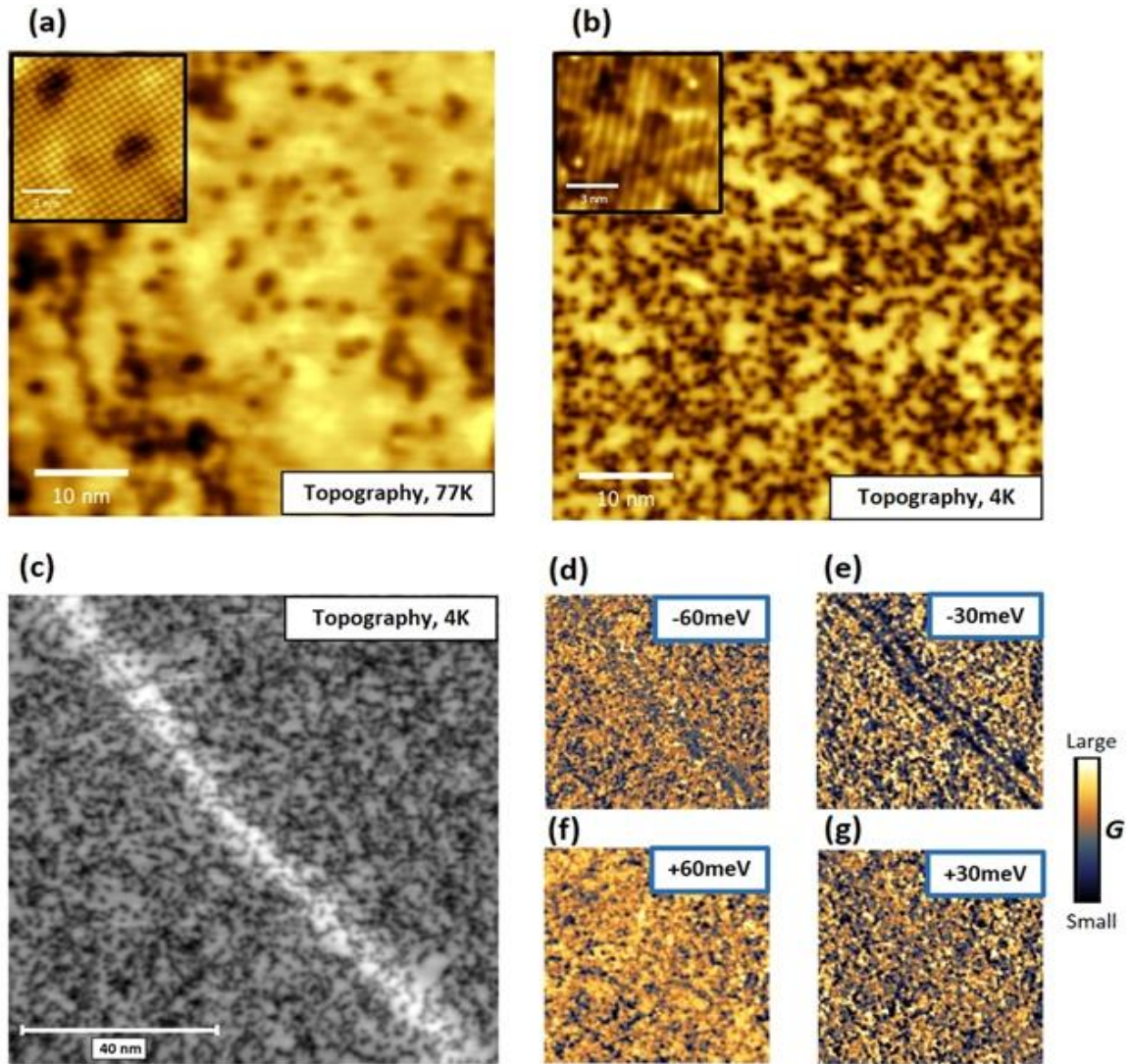


Figure 2-11. Data of 1% Au-doped BaFe₂As₂.

(a) and (b) Representative experimental STM topographic images at 77 K (a) and 4 K (b). $U_{\text{bias}} = -60$ mV, $I_{\text{setpoint}} = 100$ pA. Insets in (a) and (b) show a zoomed-in area of the surface which displays atomic lattice contrast ($U_{\text{bias}} = -100$ mV) and stripe-like features ($U_{\text{bias}} = 250$ mV). (c) STM topographic image of quasi-1D defect at 4 K. $U_{\text{bias}} = -110$ mV, $I_{\text{setpoint}} = 150$ pA. (d)–(g) Conductance maps $G(r, E)$ in the same area as in (c) at several selected energies.

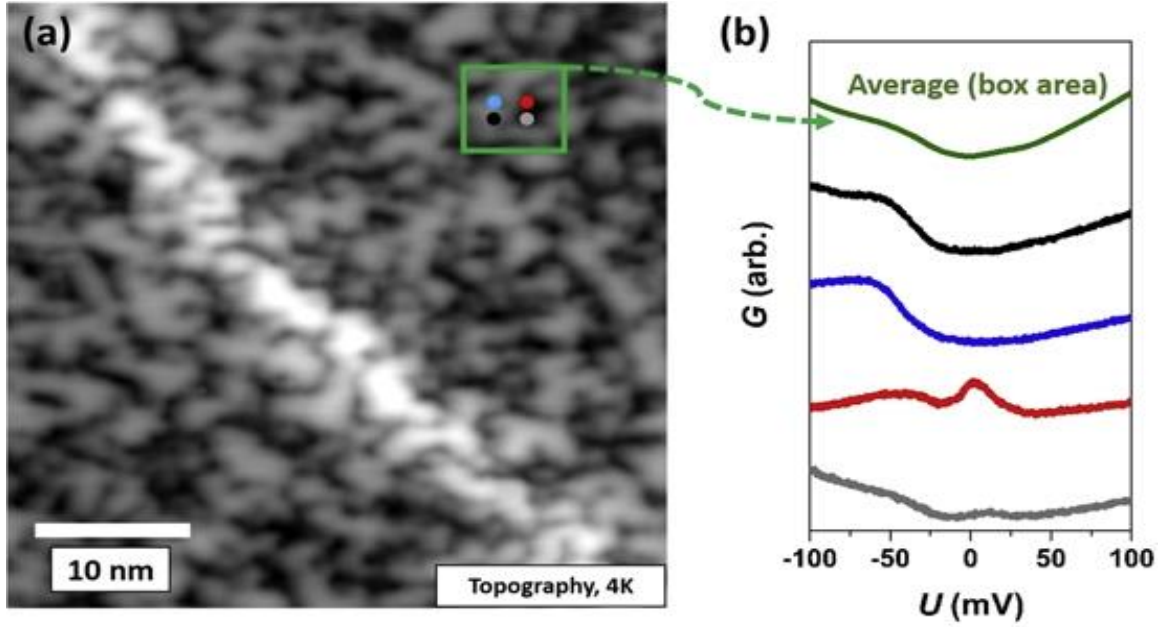


Figure 2-12. Data of 1% Au-doped BaFe₂As₂.

(a) STM topographic image of the area on which STS grid measurements are performed. (b) Green STS curve is averaged over the 182 individual STS spectra inside the box in (a). Gray, red, blue, and black curve are single-pixel (point) STS spectra recorded at the locations schematically denoted by colored dots in (a). Note that gray and red curves were extracted from locations displaying similar topographic features.

of STS curves from $G(X, Y, V)$ dataset.

We now proceed to the accurate extraction of statistically significant information associated with surface electronic structure from a deep data style analysis [161]. We use the STS dataset recorded on the topographic area shown in figure 2-12(a). The dataset has dimensions of $X \times Y \times V = 80 \times 80 \times 768$, that is, it contains a stack of 768 conductance maps with a spatial resolution $80 \text{ px} \times 80 \text{ px}$ ($1 \text{ px} \approx 0.56 \text{ nm}$). To decorrelate the STS data in a statistically meaningful way we start with imposing a lower bound limit on the number of relevant electronic behaviors within the dataset.

The smallest reasonable number of statistically significant behaviors can be estimated using k -means algorithm [163]. The k -means algorithm divides the dataset in a specified number of optimally selected clusters of curves that have similar behavior so that the within-cluster sum of squares is minimized [163, 164]:

$$\arg \min \sum_{l=1}^k \sum_{x_j \in S_l} \|x_j - \mu_l\|, \quad (2.6)$$

where μ_l is the mean of points in S_l . The selection of the number of clusters is based on the analysis of dendrogram in figure 2-13(b), in which larger vertical drops in the binary branches indicate a better cluster organization scheme in the data [163]. Based on the results shown in figure 2-13(b), we used 3 clusters as an input in our k -means analysis. The resultant spatial distribution of the 3 clusters is shown in figure 2-13(a), and the mean curves associated with each of 3 clusters are displayed in figures 2-13(c)–(e).

We further analyzed a variance in the STS curves distributed over each nanoregion (cluster) by means of PCA [165-167]. The deviation from the mean curve within each cluster associated with first 5 eigenvectors in PCA is shown by dashed lines in figures 2-14(a)–(c), and the corresponding scree plots are depicted in figures 2-14(d)–(f). The PCA analysis indicates that the cluster 2 and cluster 3 show a relatively moderate variance in the shape of the mean STS curve, allowing us to extract physical information from the curves. The STS curve from cluster 2 displays a metallic behavior and a dip at about 5–10 meV below

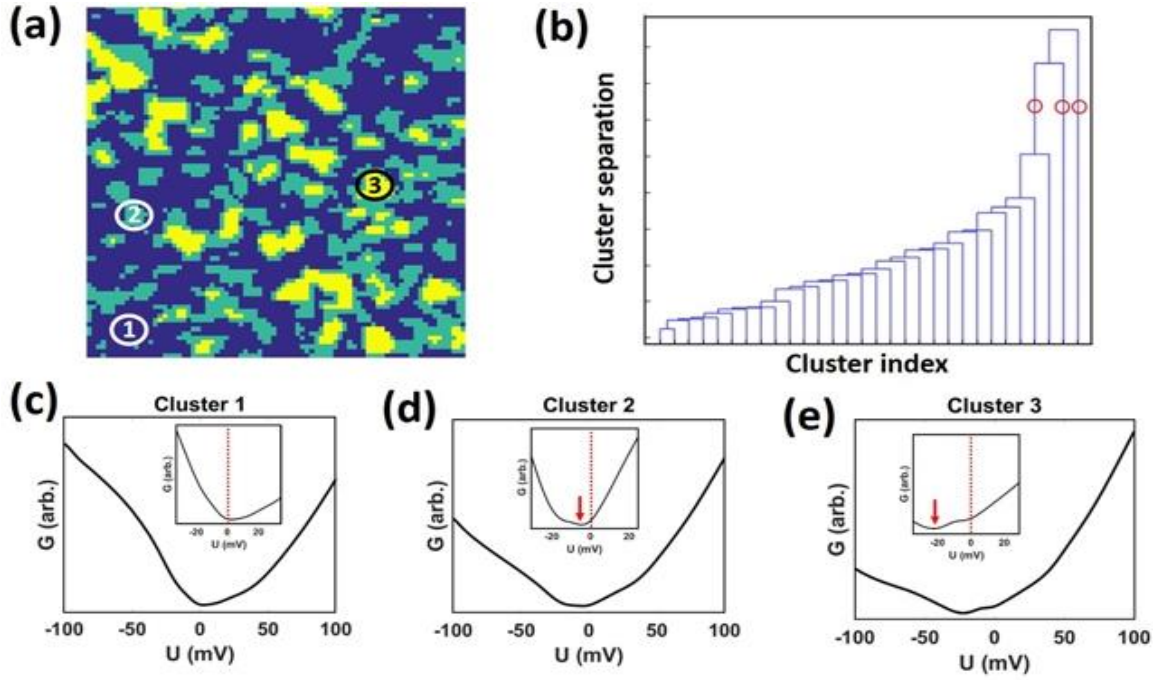


Figure 2-13. Results of k -means clustering.

(a) k -means cluster algorithm resultant map with 3 clusters specified in the image. The surface area is identical to the one in figure 2-12(a). (b) Dendrogram plot of hierarchical binary cluster tree (circles illustrate the optimal number of clusters). (c)–(e) Mean STS curves for each of 3 clusters. Insets show zoomed-in area near the Fermi level (dotted red line).

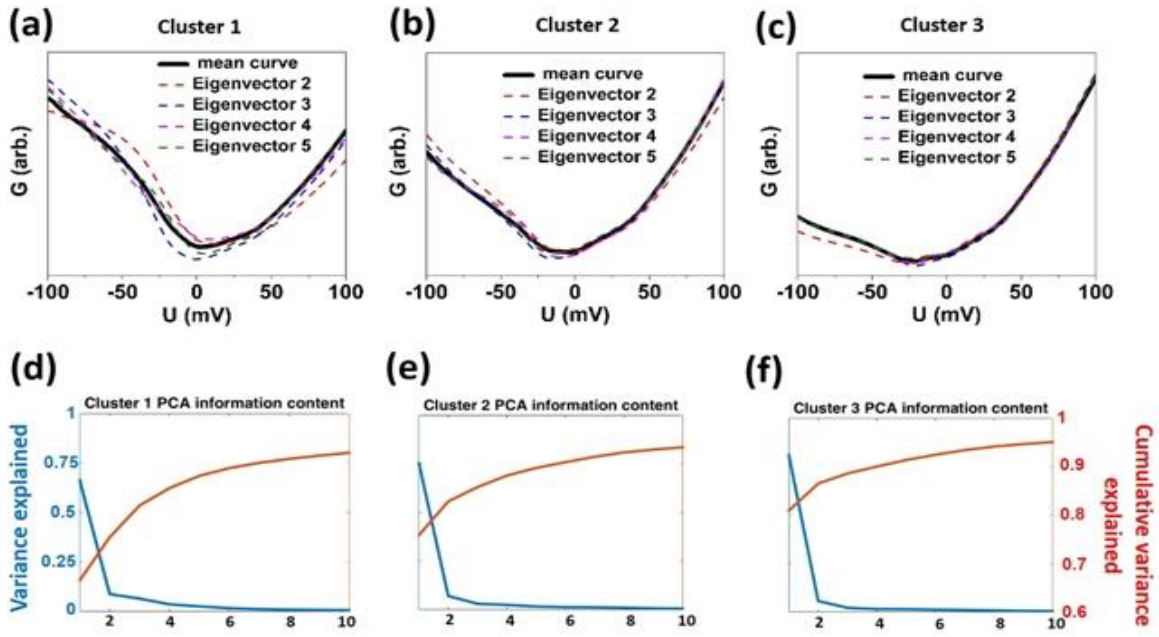


Figure 2-14. Results of PCA analysis performed within each of 3 clusters.

(a)–(c) Mean STS curves for each of 3 clusters are plotted with black solid line. The PCA-derived deviation from mean curve within each cluster is shown by dotted color lines. (d)–(f) PCA scree plots showing variance within each of 3 clusters (the horizontal axis starts with 1 in all three plots).

the Fermi level. This is in a good agreement with an observation of the SDW gap centered at around -10 meV in the ARPES measurements of BaFe_2As_2 [168]. We note that the theoretical modeling in [168] and [169] also showed that the SDW phase features a finite density of states at the Fermi level in the absence of the (coexisting) superconducting state, which is supported by our results. The mean STS curve from cluster 3 shows a metallic behavior and is somewhat similar to the curve from cluster 2, but with the center of the dip shifted to about -25 meV. We tentatively assign this behavior to the SDW phase whose characteristics were altered locally due to the presence of distortion nano-domains in crystalline lattice induced by the quasi-1D defect. Noteworthy, we did not observe similar lineshape in the regions far (>100 nm) from the defect in our experiments. The PCA-derived variance within cluster 2 and cluster 3 can be understood as relatively minor fluctuations of electronic response within a defined phase. The situation, however, is quite different for cluster 1. Here, a stronger variance in the shape of STS curves, especially in the regions close to the Fermi level, does not allow assigning any physically defined phase. This suggests that the total number of relevant electronic behaviors is larger than estimated by the k -means method. However, additional, 'hidden', electronic responses cannot be accurately revealed from the PCA eigenvectors, as they are constructed to be orthogonal and hence may not have a well-defined physical meaning.

To perform a more thorough and detailed separation of electronic behaviors in the STS dataset we adopt BLU technique. This algorithm, developed by Dobigeon and co-workers [170], is used for separating linear mixtures of spectral sources under non-negativity and full additivity constraints that allows assignment of physical meaning to the shape of the end-member curves [171]. We assume that the registered STM current signal g at each pixel p can be approximated as a linear combination of the currents flowing through each of the available 'channels', so that the latter can be represented by the endmembers [172],

$$g_p = \sum_{r=1}^R m_r a_{p,r} + n_p, \quad (2.7)$$

$$G = MA + N, \quad (2.8)$$

where m_r is the spectral measurement of an endmember r , R is the total number of endmembers, $a_{p,r}$ is the abundance of endmember r at pixel p , n_p is a zero-mean Gaussian noise, and $G = [g_1, \dots, g_p]$, $M = [m_1, \dots, m_R]$, $A = [a_1, \dots, a_p]$, $N = [n_1, \dots, n_p]$. The model is imposed with non-negativity and full-additivity constraints of the abundance coefficients:

$$a_{p,r} \geq 0, r = 1, \dots, R, \quad (2.9)$$

$$\sum_{r=1}^R a_{p,r} = 1. \quad (2.10)$$

Additionally, the model is constrained to non-negativity of the endmember spectra,

$$m_r \geq 0, r = 1, \dots, R \quad (2.11)$$

and no assumption of the presence of pure pixels is made. The estimation of endmembers and abundances is performed under a hierarchical Bayesian model [170]. For the estimation of the prior for the Bayesian model, the data $X = MA$ dimensionality is reduced to K ($R - 1 \leq K \leq V$), by an assumption that without the noise data can be represented by $(R-1)$ -dimensional convex polytope of R^V , where vertices represent pure endmember spectra m_r . For the next step, PCA projection is obtained, which allows to recover a simplex in a reduced dimensionality space via NFINDR [173, 174]. Using these results, the endmember abundance priors as well as noise variance priors are estimated from the conjugate multivariate Gaussian distribution, where the posterior distribution is calculated based on the endmember independence using Markov chain Monte Carlo (MCMC), which generates asymptotically distributed samples probed via Gibbs sampling strategy. Unmixing was run for 100 MCMC iterations for each attempt.

The number of endmembers R in the BLU algorithm must be postulated by a researcher. The lower bound limit for a total number of endmembers has already been set by the results of k -means method. To set up the upper bound limit for possible number of relevant electronic behaviors, we refer to a general underlying physics of the problem. Here, in

addition to the states associated with the SDW phase discussed in the *k*-means calculations, we must add states associated with (i) unidirectional modulation of surface charge density; (ii) possible presence of a different magnetic order 'admixed' into the SDW phase; (iii) 2 common types of point defects on the cleaved surface; (iv) diluted concentration of Au dopants; (v) randomly scattered atoms on the surface¹ [175]. Using these constraints, coupled with the results of PCA analysis, and by performing over- and under-sampling of BLU *R*-components, we found that the most relevant description of electronic behavior is achieved for *R* = 6 endmembers. The BLU results with *R* = 6, for both endmembers and abundance maps, are shown in figure 2-15. One can immediately see that endmember 4 and endmember 5 (figures 2-15(d) and (e)) corroborate the results on SDW-associated phase found earlier from *k*-means algorithm. In addition to phases already seen in the *k*-means, the BLU analysis revealed new features in electronic behavior that can be linked to the fundamental physical properties of the material, as described below.

The endmember 2 shows a well-defined signature of a spectral gap of $2\Delta \approx 40$ meV centered near the Fermi level (figure 2-15(b)). We note that the gap of a similar behavior and magnitude (~ 30 – 40 meV) was observed in STM experiments on a closely related compound from AFe_2As_2 family, SrFe_2As_2 , in which it was explained as the SDW-originated gap [162]. However, recent photoemission spectroscopy measurements and theoretical modeling [168] on the AFe_2As_2 type compound revealed that the SDW and SC orders must coexist spatially in order to produce a gap at the Fermi level. Otherwise, the SDW opens a gap below the Fermi level [168, 169], in agreement with behavior observed in the endmembers 4 and 5. If the SDW and SC orders indeed coexist on a local scale in our sample, the formation of the electron-pairing 'islands' associated with the SC order is expected to produce a continuous drop in bulk resistivity measurements [145, 176]. However, this is clearly not the case for our compound, in which the resistance showed a

¹ See online supplemental material stacks.iop.org/NANO/27/475706/mmedia for bulk magnetic susceptibility and resistivity measurements, additional STM data at 4 and 77 K, and the details of multivariate statistical data analysis.

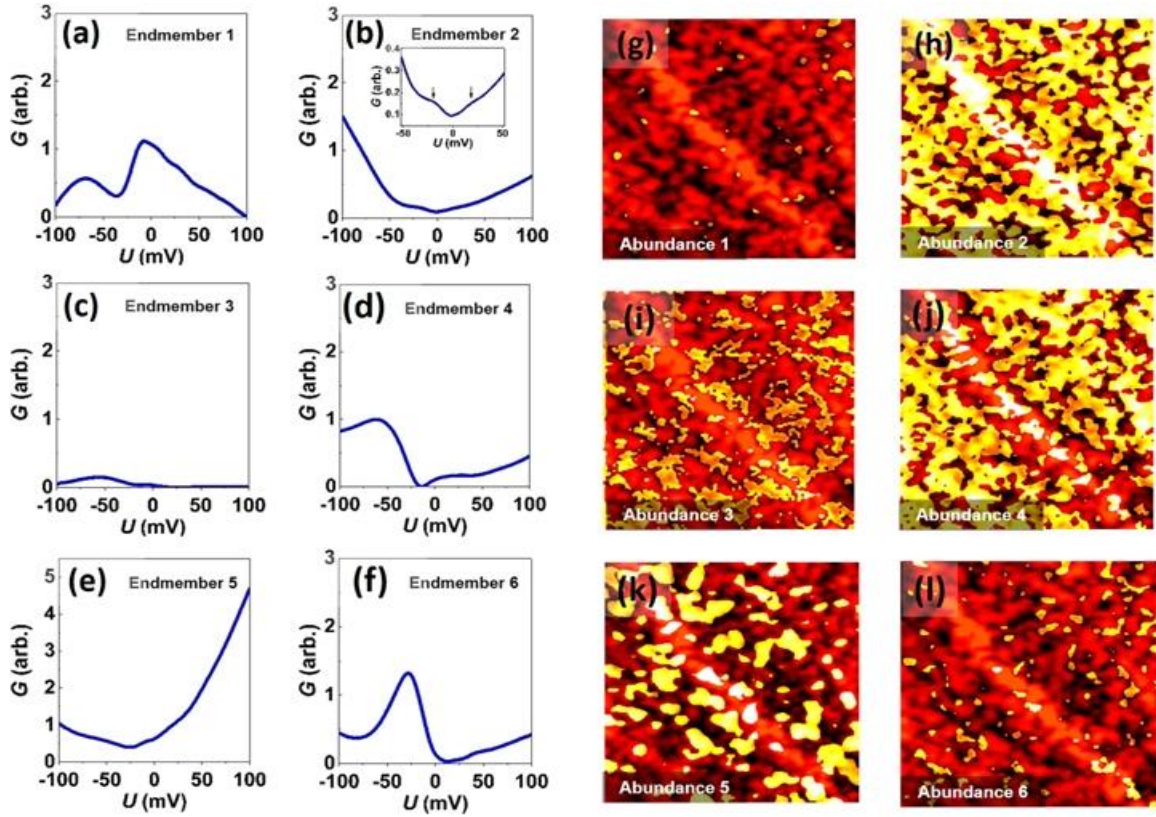


Figure 2-15. Results of Bayesian linear unmixing of STS dataset.

(a)–(f) 6 Bayesian endmembers (see text for details). Inset in (b) zooms in the spectral gap features at the Fermi level. (g)–(l) Corresponding abundance maps (yellow) overlaid on the topographic image (red). The intensity contributions in the abundance maps below 0.2 were cut off for a better visualization (see SM for a full (0, 1) intensity maps).

small *upward* trend in the temperature range of interest (see footnote 5). This rules out a scenario in which the 'admixture' of SC order leads to the gap feature seen in the endmember 2. We therefore describe the spectral weight loss at the Fermi level observed in the endmember 2 in terms of a pseudogap-like state, which is defined here as a state outside the 'superconducting dome' and not directly associated with either SDW-induced gap or 'SDW+SC'-induced gap.

We next discuss a possible physical origin of the pseudogap-like state associated with endmember 2. At first glance, it is tempting to link the pseudogap-like state to the 1D striped charge order seen in the STM topographic image. However, a possibility of such direct correlation quickly falls apart as we were able to find the 1D stripes even in the areas without spectral features of a pseudogap. Another explanation of a pseudogap-like state is based on the possible formation of a different, short-range, magnetic order admixed into the SDW phase, which is consistent with the presence of upturn in a magnetic susceptibility data below T_N . The formation of a pseudogap-like state may also explain a peculiar upswing in the resistivity in the SDW phase below ≈ 20 K. Noteworthy, our finding of a $2\Delta \approx 40$ meV pseudogap-like feature, which is not directly related to SDW, correlates with photoemission spectroscopy results of Xu *et al* [177] that showed emergence of $2\Delta \approx 36$ meV pseudogap state at the Fermi level of the underdoped $\text{Ba}_{1-x}\text{K}_x\text{Fe}_2\text{As}_2$ in both SC phase and non-SC phase without long-range SDW order.

The endmembers 1 and 6 in figures 2-15(a) and (f) show a clear resonance peak features associated with impurity induced bound states in the SDW phase of this compound. The impurity-induced nature of these peaks is further confirmed by the inspection of the corresponding abundance maps that show the peak features are generally constrained to a point-like areas on the surface (figures 2-15(g) and (l)). Of particular interest is the endmember 1 which can be described by a non-magnetic impurity-induced double resonance peak model studied in [169]. We tentatively ascribe the origin of this spatially diluted double peak state to the Au dopants. Finally, the origin of endmember 3 is likely related to the minor instabilities ('noise') of the tunneling junction during the grid acquisition.

2.4.6 Conclusions and outlook

Our results on the identification of a surface nanoscale electronic structure in the underdoped state of FeAs-based superconductor, by means of a deep data style analysis are important for providing clues to understand how the high-temperature superconductivity may emerge in these systems. First, while there is a growing evidence of the pseudogap state formation in FeAs compounds [177-180] there is still an open debate on the relation of the pseudogap to the superconducting state and on the role of magnetic correlations in the formation of the pseudogap. Our revelation of 'buried' pseudogap-like spectral features in the SDW phase, combined with results of magnetic susceptibility and resistivity measurements, suggests a potential link between a pseudogap state and weak or short-range magnetism within SDW phase. In future, the temperature-dependent STS measurements coupled with the presented here deep data analysis could provide additional details on a physical nature of this phase. Second, the real-space analysis of the electronic character of impurity-induced quasiparticle states found in the spectral unmixing of our data can be further used as a probe into the details of the strong correlations in the system. In this sense, it is natural to extend the deep data approach to the reciprocal space, which is commonly used to study quasiparticle interference pattern [181]. We expect that a nanoscale inhomogeneity in the electronic structure of the surface would produce spatially different scattering patterns at the same value of energy. The application of techniques such as sliding FFT combined with multivariate analysis [182] would allow hidden scattering patterns to be uncovered. Finally, we note that the presence of 1D charge modulation at the surface did not allow us to measure the atomic lattice constant in the regions close to the defects that showed peculiar changes in electronic behavior within the SDW phase. We do expect, however, that for the systems in which the atomic lattice can be resolved (i.e., no surface 'reconstructions' occur), one can perform a direct data mining to correlate minute variations in atomic positions with the changes in spectral characteristics, such as the magnitude of SDW and/or SC gaps. As the ever-increasing amount of STM/S data on strongly correlated systems makes the individual inspection of datasets highly impractical and, in many cases, nearly impossible, the approach outlined here will present an ideal tool

for an accurate mapping of locally inhomogeneous electronic structure on the surfaces in an automated fashion of a full information extraction.

2.4.7 Methods

Single crystals of lightly Au-doped BaFe_2As_2 ($x = 0.009$, $\sim 1\%$) were grown out of self-flux using a high-temperature solution-growth technique. STM/S measurements presented in the manuscript were carried using a Joule–Thomson scanning tunneling microscope (JT-STM, Specs, Berlin). Tungsten (W) STM tip was prepared by gentle field emission at a clean Ag(111) sample. The samples were cleaved *in situ* in the STM machine chamber at approximately 110 K. The samples were always measured first at 77 K, followed by cooling the system further down and performing measurements at 4 K. The representative figures of the surface at 77 K and at 4 K were selected from data obtained by routine shift of the scan frame over $(2 \times 2) \mu\text{m}^2$ surface area, which ensured the overlap between areas imaged at different temperatures. The estimated lateral 'drift' upon cooling in the current experimental set-up is less than $1 \mu\text{m}$. STS measurements were performed using a standard lock-in amplifier techniques, with a bias modulation between 2 and 5 mV. Matlab codes were used for data processing and statistical analysis.

Appendix

Appendix 2-1. Supplementary information for Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl₃

Supplementary Note 2-1.1. Comparison of STEM results with $P3$ and $C2/m$ models.

In $P3_112$ structure, each Ru/Cl column contains one Ru atom ($Z_1=44$) and two Cl atoms ($Z_2=Z_3=17$) per unit cell, whereas each Cl column contains two Cl atoms ($Z_1=Z_2=17$) per effective unit cell. As the intensity of 2-dimensional STEM image is roughly proportional to Z^2 and scales linear with a number of atoms in the column [183-185], the projection of the $P3_112$ structure on to the plane perpendicular to c^* -axis would lead to a well-defined difference in contrast between three Cl columns and six Ru/Cl columns producing pattern with a periodic repetition of 6 brighter and 3 darker columns [see Figure 2-16 (b)]. This agrees well with experimental observations. The situation, however, is different for $C2/m$ structure. Here, when considering the projection on to the plane perpendicular to c^* -axis, the Z^2 dependence of the intensity would produce a pattern with well-defined line features, in which every third line containing only Cl atomic columns appears substantially darker than the two other lines originating from mixed Ru/Cl columns [see Supplementary Figure 2-16 (c)]. This does not fit the experimental observation.

Supplementary Note 2-1.2. Neutron diffraction measurements on single crystals.

The single crystals neutron diffraction data reproduces well a lateral component of Cl cage distortion found in DFT and STEM (schematically shown with red arrows in Figure 2-17; see also Table 2-1 for details). It is noteworthy, however, that several other distortion components seen in neutron data, including a distortion in Ru sublattice, were not directly evident from STEM and DFT results. This can be explained by a relaxation of the structure upon exfoliation.

Supplementary Note 2-1.3. Computation of normal displacement modes from STEM image via principal component analysis (PCA).

PCA normal modes are computed from a set of atomic centers segmented into M periodic unit cells and represented by the matrix $\mathbf{N}$. The rows of matrix $\mathbf{N}$ correspond to the unit cells and columns correspond to (x, y) coordinates of each of atoms within the studied system. The rows and columns represent “observations” and “variables”, respectively. In the first analysis, (x, y) coordinates are computed for each atom within each unit cell with origin $(0, 0)$ corresponding to the center of mass of all 9 atoms, resulting in 18 variables. In the second part of the analysis, we calculate the center of mass of Ru/Cl mixed columns, and (x, y) coordinates are only computed for 3 pure Cl columns, resulting in only 6 variables. We run PCA to find statistical normal modes of displacements. PCA is used to convert the matrix $\mathbf{N}$ into a product of expansion coefficients a_{ik} and eigenmodes w_k [78, 186].

$$N_i(\omega_j) = a_{ik}w_k(\omega_j) \quad (s.2.1)$$

The eigenvectors are orthogonal and are arranged in descending order by variance. We do not perform mean subtraction, which makes the first eigenvector the average coordinates of atomic columns. Therefore in the main text, we start count from the second eigenvector. We note that PCA is a dimensionless type of analysis, which results in coordinates not corresponding to the coordinates we see in original data. In order to explore the variance in the atomic column positions we explore the 17 or 5 other eigenvectors in each case. Scree plots and percentage of the variance explained for both parts of the analysis are shown below in Figure 2-18.

Selection of the number of eigenvectors explored is usually based upon the presence of a bend in the Scree plot. However, in the first case the estimation is somewhat obfuscated by eigenvectors 16 and 17 containing very little information, while in the second we have too few variables to disregard any. For the Ru/Cl center of mass analysis we have shown the plot for all 5 eigenvectors. For the unit cell center of mass we have shown the first 6

eigenvectors, as further exploration had revealed that eigenvectors 7 through 15 show mostly slight variations in the mixed Ru/Cl atomic positions, while the first 6 are most prominent in the pure Cl sublattice.

In Figure 2-19a we show the data for center of mass of entire cell as origin for the first six eigenmodes. It is clear that the same trend is preserved, i.e. variation in geometry of Cl sublattice stronger than those in Ru sublattice. In Figure 2-19b we show principal normal modes only for three Cl columns in the unit cell whose coordinates were recalculated with respect to the center of the mass of six Ru/Cl columns in each cell. Such analysis can provide greater details on the potential forms of distortion variation in Cl sublattice with respect to Ru-contained columns. The first eigenmode (blue arrow) shown in Figure 2-19b can be described as a polar displacement mode of Cl columns, whereas the rest four eigenmodes can be assigned to a combination of shear-, stretch-, and rotation-like modes. Note that a small mismatch between the centers of the atomic cores of Cl atoms (less than 1.5 pm) calculated by DFT (and also found in neutron diffraction data) results in a fine splitting of each Cl column into 2 sub-columns. This splitting is beyond our experimental resolution and therefore the STEM views them as a single column, in which the center of the column corresponds to a “center of mass” of 2 sub-columns. Due to this fine structure of Cl columns, the presence of atomic vacancies in one of Cl sub-columns can shift a location of the center of mass. However, this alone cannot explain the complex distortion variation patterns observed in Figure 2-19a and Figure 2-19b. One possible additional contribution may originate from strain associated with a presence of Cl vacancies, which may induce tilts and rotations of RuCl_6 octahedral. Another component of distortional variation may arise from an interplay between several alternative pathways of the RuCl_6 octahedra compression. Overall, our multivariate analysis implies a presence of complex local distortions in Cl ligand cages.

Supplementary Note 2-1.4. Identification of positions of electron density peaks in STM image.

We first used FFT-filtered experimental STM images in order to acquire unit cell template and identify the coordinates of the centers of electron densities maxima within it. We then perform normalized cross correlation between the original image and the template and apply threshold of 0.70 to the resulting correlation surface in order to identify positions of probable centers of the unit cells.

Based on the coordinates in the template, for each potential center of the unit cell we first make a guess of one position. Using the proposed coordinate we perform intensity analysis within the 7-pixel radius of initial guess. We perform intensity threshold (with value of mean intensity of pixels in the selected area) and identify center of the largest resulting binary region. We assign this as the center of first (out of three) electronic density maxima for the unit cell.

We perform the same analysis for the 2 other protrusions associated with electronic density maxima, except at the final step we now check if newly found center is within 10 pixel distance to any other previously found center, in which case we do not record it. This alleviates the issue with identifying same centers twice due to low correlation surface threshold chosen to maximize number of identified unit cells.

After the refinement, we extract the sum of pixel intensities in a 9 by 9 pixel patch centered at each identified density center. We normalize the distribution to lie on the 0 to 1 interval.

Supplementary Note 2-1.5. Analysis of structural inhomogeneity in Cl sublattice from STM image of a surface monolayer.

The STM measurements reflect mainly the spatial distribution of the electron charge density near the Fermi level and are generally considered not as accurate as, for example, STEM measurements, for a detailed analysis of lattice structure. For example, the shift of charge density maxima with respect to Cl atomic cores may produce an apparent uniform distortion of individual hexagons of Cl atoms in the STM image of α -RuCl₃ [187]. On the

other hand, as long as the electron density lattice is commensurate with the underlying atomic lattice, the *relative* distortions of the atomic positions in crystallographic lattice can be inferred from the STM data using statistical learning methods. Below we show a result of application of k-means clustering algorithm to high resolution STM image of Cl-terminated RuCl₃ surface. The k-means algorithm divides the dataset in a specified number of optimally selected clusters, which define the groups of atoms with specific structure neighborhoods, so that the within-cluster sum of squares is minimized [78]:

$$\arg \min \sum_{i=1}^k \sum_{x_j \in S_i} \|x_j - \mu_i\|, \quad (s. 2.2)$$

where μ_i is the mean of points in S_i . The dendrogram and the resultant spatial distribution of nearest neighbor distances for the case of 3 and 4 clusters are shown in Figure 2-20.

This analysis clearly reflects a non-uniform (inhomogeneous) distribution of the nearest-neighbor distances of electron density lattice points thus implying the distortion of the underlying Cl lattice itself. The latter is in the agreement with the structure inhomogeneity found in the STEM analysis discussed in the main text, but now it is demonstrated for a single surface layer of α -RuCl₃.

Note that several small areas in Figure 2-20 were not identified in this analysis as their structure shows a larger deviation from average FFT unit cell than the rest of the structure.

Supplementary Note 2-1.6. (Uniform) lattice distortion as a possible origin of charge ordered state.

In Figure 2-21 we showed a projection of Cl atom distortion found in DFT onto the 2D surface plane. The symmetry of atomic distortions at the surface plane matches well with the symmetry of the $R\bar{3}$ superlattice formed by the brightest spots in the STM image as illustrated in Figure 2-21a. However, our DFT-simulated STM images performed on the structure with Cl distortion did not reproduce a charge imbalance at 3 Cl atoms in the unit cell, as is clearly seen from Figure 2-21b and Figure 2-21c.

Supplementary Note 2-1.7. Theoretical investigation of potential surface relaxation effects with PBE.

To theoretically investigate the possibility of surface relaxations in α -RuCl₃ we perform PBE calculations of a bulk and slab unit cell illustrated in Figure 2-22. The bulk unit cell is taken to be the $P3_112$ unit cell (same as in the manuscript). The three layer slab is constructed from the bulk $P3_112$ unit cell by inserting 10 Å of vacuum. Contrary to the calculations described in the manuscript, the lattice constants in the simulation described in this section are relaxed. Also approximate van der Waals corrections have been added in order to take into account the interactions between the RuCl₃ “sandwiches”, according to the DFT-D2 method of Grimme [188]. The other computational details are the same as described in the Method section of the manuscript.

As can be seen in Figure 2-22 the integrated Local Density of States (LDOS) from the slab and the bulk calculation are very similar and neither of them shows a reordering of the charge. Therefore surface relaxations do not seem to be a possible explanation of the experimentally observed charge order.

Figure 2-23 compares the Ru-Cl bond distance ($d_{\text{Ru-Cl}}$) and three bond angles obtained from the bulk and the top-layer in the slab respectively. The small differences between the bulk and the surface results reflect the quasi 2D-ness of α -RuCl₃. The difference with results in the caption of Figure 3 of the main text stems from the fact that in the simulations described in this section the lattice parameters are relaxed instead being set to the experimental values.

Finally Figure 2-24 compares the Density of States (DOS) from the Ru-d orbitals in bulk α -RuCl₃ and from the Ru-d orbitals in the surface layer of the three RuCl₃ layer slab with 10 Å of vacuum described above. In particular we note that the Ru- t_{2g} band width in both cases are very similar (roughly within [-1.0, 0.2] eV) indicating that the overlaps between the Ru- t_{2g} (Wannier) orbitals in the bulk are very similar to those in the surface again reflecting the quasi 2D-ness of α -RuCl₃.

Supplementary Note 2-1.8. Relevance of our data to quantum magnetism (extended discussion)

As explained below, the measurements presented in our manuscript should be valid in bulk samples and wide temperature regimes. This automatically leads to the relevance of our findings to the Kitaev physics. These specific points are as elucidated below:

Relevance of thin films and surface measurements to bulk α -RuCl₃ samples.

Based on the current experimental [55] and theoretical [76] understanding, it is reasonable to assume that the α -RuCl₃ can be treated as a quasi-2D material, in which each layer is nearly independent of the others. Taking into account the absence of dangling bonds, as well as the absence of alkali atomic layers above and below each RuCl₃ layer (in sharp contrast with iridates), we do not have any reason to believe that such quasi-2D system should develop any additional disorder *specific* to exfoliation/cleavage and reduction of the number of layers. This line of thinking is further supported by reports on similar layered compounds TiCl₃ and VCl₃ [189], as well as by conclusions of Weber *et al.* [96] that an in-plane structure of α -RuCl₃ is maintained in the exfoliation process. As shown in the Supplementary Note 2-1.7, the surface relaxation effects can also be negated. Thus the modulations of structural and electronic orders reported in our study are very likely intrinsic to the bulk of the sample (i.e. our findings should be applicable to each of the quasi-isolated RuCl₃ layer inside a bulk sample).

Relevance to ‘Kitaev physics’.

- a. We first note that the Kitaev interactions should be present within Mott ground state [60], an existence of which in our study was confirmed *via* STS measurements of local density of states in α -RuCl₃ surface monolayer. The persistence of Mott state to high temperature also agrees with proposals put forward by Plumb *et al.* [69] and Zhou *et al.* [91].
- b. In a 2D Kitaev QSL there is no phase transition into the paramagnetic state [190]. The spins are short-range correlated at low temperatures and as the temperature increases

paramagnetic fluctuations overwhelm, however, there is no order parameter associated with this phase transition. This means that the same spin fluctuations which persist at low temperature should persist to, theoretically speaking, infinite temperature, although thermal fluctuations will make it harder to access them. This is already demonstrated in several papers and results, most notably *i)* Nasu et al [191] where the bi-fermionic distribution is a fundamentally continuous evolution with temperature, up to very high temperatures close to room temperature, *ii)* ARPES measurements [91] that indicate the qualitatively same conclusions, and *iii)* the susceptibility measurements are definitely anisotropic to much above T_N [53, 71, 87] which shows that the magnetic interactions (K and J terms) responsible for the magnetic ground state and the magnetic Hamiltonian have formed at (and possibly above) room temperature, at which the current atomic-scale measurements were undertaken.

c. Finally, the peculiar charge order(s) we have found in our study will be a direct input for first-principle calculations to understand the role of such in determining the actual Hamiltonian of this material which has so far not been determined to any precision, especially since we believe it persists to low temperature and also in the bulk.

For all of the above reasons, it is natural to expect that our data captures much important information necessary to understand the quantum magnetism at low temperatures.

Table 2-1. The structure parameters of α -RuCl₃ measured at 200 K by single crystal neutron diffraction. The space group is $P3_1$, $a=5.956$ Å; $c=17.054$ Å, $\alpha=\beta=90^\circ$, $\gamma=120^\circ$. U has units of Å².

Atom	Type	x	y	z	U
Ru1	Ru	0.44445	0.55555	0.03333	0.005
Ru2	Ru	0.07567	0.20533	0.36667	0.005
Cl3	Cl	0.13267	0.56633	0.00000	0.009
Cl4	Cl	0.12067	0.54833	0.46667	0.009
Cl5	Cl	0.08067	0.21233	0.66667	0.009
Cl6	Cl	0.11967	0.21333	0.13333	0.009
Cl7	Cl	0.40667	0.20233	0.00000	0.009
Cl8	Cl	0.42967	0.20533	0.46667	0.009

Appendix 2-2. Supplementary information for Knowledge Extraction from Atomically Resolved Images

Supplementary Note 2-2.1. Atom identification

For each image, we extend the lattice under the assumption of underlying perfect square lattice with potential distortions due to either image processing specifics or inhomogeneity of the crystalline order. This leads to the necessity of enforcing the uniformity of local translational vectors, while introducing some corrections in the long-range scale by visually identifying mismatching results. After identifying positions, we collect total intensities within a patch of a normalized image. The size of such patch is determined by the resolution of the image. We normalize collected intensities to be in the range from 0 to 1, and create a fit of two Gaussian distributions to the resulting histogram of intensities. (Figure 2-25.) We select the threshold for classifying the atoms by matching resulting distribution to the expected stoichiometry of the material. (Figure 2-26) Finally, we construct a graph with nodes corresponding to atoms and their type and edges connecting first nearest neighbors. Based on this representation we transform the data from image coordinate system to that of the underlying square crystal lattice.

Supplementary Note 2-2.2. Notes on statistical distance.

Statistical distance s defined by Eq. (2.1) has been introduced to physics by Wootters as a measure of distinguishability,[130] while Braunstein and Caves have shown its relation to the relations of uncertainty (quantum and classical).[131] It can be understood as the shortest distance (geodesic) between two points representing probability distributions in the space of all k -state probability distributions with the metric tensor given by the Fisher information matrix (Figure 2-17a).[192]

The distance can be related to the better-known χ^2 -distance between probability distributions P and Q ,

$$\chi^2 = \sum_{i=1}^k \frac{(q_i - p_i)^2}{p_i}, \quad (\text{s. 2.3})$$

which is typically used in estimation. In Eq. (s. 2.3) p_i and q_i are the probabilities of states i in systems P and Q , which determine the average number of Poisson distributed counts in histogram bins i . For Poisson distributions the mean value of counts is equal to its variance, i.e., $p_i = S_i^2$. In the task of approximation, such as model development, where we seek a symmetric similarity metric, $\mathcal{C}^2$ -distance can be symmetrized by replacing p_i with $(p_i + q_i)$ in the denominator. However, a more appropriate distance for the current application can be constructed by the following symmetrization,

$$\begin{aligned} D_H^2 &= \sum_{i=1}^k \frac{(q_i - p_i)^2}{(\sqrt{q_i} + \sqrt{p_i})^2} = \sum_{i=1}^k (\sqrt{q_i} - \sqrt{p_i})^2 = \\ &= 2 - 2 \left\langle \sqrt{\frac{p_i}{q_i}} \right\rangle_q = 2 - 2 \left\langle \sqrt{\exp[-\beta(\Delta u_i - \Delta F)]} \right\rangle_q, \end{aligned} \quad (\text{s. 2.4})$$

which represents the squared Hellinger distance. Unlike the symmetrized $\mathcal{C}^2$ -distance, D_H can be expressed as the ensemble average (denoted by angle brackets) of odds ratio p_i/q_i , i.e., it belongs to the group of f-divergences. This is essential for the ability to determine the distance from simulations without the knowledge of the full partition functions of the two systems. While determining the absolute value of partition functions is notoriously difficult, here only their ratio is needed. The last equality in Eq. (s.2.4) is applicable in the special case of the Boltzmann distributed probabilities of equilibrium systems. Moreover the square root transformation equalizes the variances of histogram counts (i.e., their statistical weights) and places thus all histogram bins on equal footing.[134] Therefore, only the average values of counts need to be known for consistent comparison of different histograms. The D_H distance is suitable for comparing histograms in which counts in individual bins are independent. However, for histograms in which every count is

registered in exactly one bin at a time, such as in the present case of collecting statistics of local configurations, the normalization condition $\sum_{i=1}^k p_i = 1$ constrains the distance to the form of Eq. (2-3), as illustrated in Figure 2-17b. For a single pair of histograms s and D_H^2 lead to identical optimization results, but the distinction must be made when combining different target data collected in several different histograms. For additional arguments leading to the choice of the statistical distance as the natural basis of comparison between statistical mechanical systems see Ref.[132].

Supplementary Note 2-2.3. Alternative models.

To investigate whether more information about the microscopic forces could be extracted from the image, we have also tested a model optimized by matching 60 different local configurations that take into account not only the nearest neighbors, but also the next-nearest neighbor atoms, i.e., diagonal neighbors sharing a vertex on a square lattice. The histograms of local configurations in the target and model systems in Figure 2-28 show a close match. The resulting optimized parameters $w_0 = 0.522$, $w_{-1} = -0.015$, and $w_1 = 0.059$ are very close to those obtained by matching 12 nearest neighbor configurations, indicating that the additional information did not provide significant constraints.

In addition to the 3-dimensional model distinguishing three types of neighboring atoms, we also tested a simplified 2-dimensional model with elemental segregation modeled only through effective interactions within the same layer. Here we do not assume that only same-layer interactions are important, but rather that the effect of interlayer interactions can be effectively captured in this simplified manner. The ultimate simplicity of the model and its compatibility with the character of the surface data also prevents overfitting and provides a single parameter controlling segregation. The single-parameter optimization resulted in $w_0 = 0.526$, which is very close to the value found for the full 3-dimensional model, again suggesting that the same layer interactions are the main driving force of the elemental segregation.

Appendix 2-3. Supplementary information for Deep data mining in a real space: Separation of intertwined electronic responses in a lightly-doped BaFe₂As₂

Supplementary Note 2-3.1. Bulk magnetic susceptibility and resistivity measurements

Single crystals of lightly Au-doped BaFe₂As₂ (x= 0.009, ~1%) were grown out of self-flux using a high-temperature solution-growth technique [157]. Temperature-dependent magnetic susceptibility, χ , decreases with decreasing temperature and drops abruptly below $T_N = 110$ K (Fig. 2-29), which also overlaps with the structural transition, T_s [157]. Upon further cooling, χ increases in magnitude, which may be associated with additional magnetic contribution. Our neutron data (unpublished) confirms that the intensity of a wave vector relevant to SDW does not change in the corresponding temperature range, suggesting that the additional contribution to susceptibility comes from a different origin than simply enhanced SDW AF order. Resistance, R , diminishes with decreasing temperature from room temperature ($R_{300K} \sim \text{m}\Omega\cdot\text{cm}$), rising below ~ 140 K, and shows slight upward trend below $110 \text{ K} \approx T_N = T_s$, and a couple of smaller features below (inset of Fig. 2-29) [157].

Supplementary Note 2-3.2. STM at 77 K and at 4 K.

Our STM observations at 77 K on a cleaved surface shows a square-like lattice (Fig. 2-30) with a unit cell of $(0.56 \text{ nm} \times 0.56 \text{ nm})$. This agrees well with measurements on BaFe₂As₂ reported in [193]. We note that we were not able to observe the same square lattice at 4 K on the same surface. Instead, a well-defined 1D modulation of a charge density was typically observed at 4 K (Fig. 2-31).

During the STM measurements at 77 K, we were usually able to observe 2 types of bright protrusions (marked by red and orange ovals in Fig. 2-30), which we relate to specific point defects in the crystallographic lattice. We assume that these defects are preserved upon cooling down to 4 K, although we were not able to get a clear STM image of these defects due to large inhomogeneity and 1D modulations of charge density in the 4 K topographic images.

Supplementary Note 2-3.3. Principal Component Analysis (PCA)

Using the results of k -means clustering as an input, we acquire separate $G_m(N_m, V)$ subsets, where N_m corresponds to number of observations in the cluster m , and V corresponds to number of parameters and is the same among all subsets (768 points in the -100 to +100 mV energy range). We perform PCA on each of the subsets to convert them into a product of expansion coefficients a_{ik} and eigenmodes w_k . The eigenvectors are orthogonal and are arranged in descending order by variance [186]

$$G_i(\omega_j) = a_{ik}w_k(\omega_j) \quad (s. 2.5)$$

As the data is not mean-centered, the first eigenvector is the average STS curve within the selected cluster, and all other eigenvectors are deviations from that average curve sorted in order of explaining maximum amount of variance within remaining data.

Dashed lines in Figure 2-14a-c in the main text demonstrate STS curves reconstructed from PCA by adding corresponding deviations to the average curve. Figure 2-14d-f shows the corresponding scree plots and cumulative variance explained for the first 10 eigenvectors within each cluster.

Supplementary Note 2-3.4. Bayesian Linear Unmixing: choice of number of endmembers

To establish a number of endmembers in BLU analysis, we first used the results of k -means and PCA analysis (see also Fig. 2-13 in the main text). The mean curves associated with clusters 2 and 3 in k -means method show a relatively small variance in PCA analysis (see Fig. 2-14 (b, c, e, f) in the main text) and can be therefore described in terms of a physically defined electronic “phases”. Furthermore, we also were able to see persistently the well-defined spectral features associated with these “phases” in BLU analysis as we varied a number of endmembers between $R=4$ and $R=8$ (described below). The situation is quite different for cluster 1, which shows relatively large variance in PCA analysis, and can be

in principle decomposed into several spectra associated with distinct electronic “phases”. The number of these “hidden” spectral behaviors can be estimated from PCA scree plot which suggests that the most relevant information associated with cluster 1 can be expressed by 4 PCA components (Fig. 2-32). We therefore use the $R=6$ endmembers in the BLU analysis of the full dataset, which fits within the physical constraints on the number of distinct electronic responses, which was defined in the main text.

We next proceed to confirming our choice of the total number of endmembers by over- and under-sampling procedure. The proposed optimal number of the endmembers is $R=6$, as determined earlier. Here, we demonstrate scenarios, in which the full dataset is BLU-unmixed into $R=5$ (under-sampling) and $R=7, 8$ (over-sampling) components. The results for endmembers and associated abundance maps are shown in Fig.2-33 and Fig. 2-34, respectively. We will denote the spectral curve associated with a specific endmember for each unmixing result as m_r^R . The spectral features associated with spin-density wave induced gap (m_4^R and m_5^R curves) appear for both $R=5$ (under-sampling) and $R=7$ (over-sampling). These are also the spectral features found earlier in k-means clustering.

On the other hand, it is clear that the under-sampling, $R=5$, leads to incomplete separation of other relevant electronic responses from experimental dataset. Indeed, impurity induced spectral features are not accurately revealed in $R=5$ case, in which the m_1^5 endmember exhibits only one broad peak [Fig. 2-33 (a)]. In addition, the spectral features of a pseudogap-like state are not seen for $R=5$. Comparison of endmembers between $R=5$ and $R=6$ scenarios [Fig. 2-33 (a) and 2-33 (b)] suggests that the inability to accurately reveal the pseudogap-like state in $R=5$ case is due to the partial transfer of the spectral weight from the “impurity phase” into the “pseudogap phase”, which blurs over the spectral gap features in the positive energy range in m_2^5 curve. The presence of such transfer is confirmed by the inspection of relative intensities of m_1^5 and m_1^6 curves in the energy range from 0 meV to +25 meV, which shows the intensity of m_1^5 curve is roughly twice smaller than that of m_1^6 curve in this region (intensities are estimated as areas under the curves in the corresponding regions).

We next examine the over-sampling scenario. Both impurity double resonance and pseudogap spectral features are persistently seen for $R=7$ and $R=8$, in a good agreement

with the results for $R=6$, as is confirmed from the inspection of m_1^R and m_2^R endmembers in Fig. 2-33 (b)-(d) and the associated abundance maps in Fig. 2-34 (b)-(d). We also note that the oversampling splits the m_5^7 curve [Fig. 2-11(c)] into the two curves represented by m_5^8 and m_6^8 endmembers [Fig. 2-33(d)]. This can also be clearly seen from corresponding abundance maps in Fig. 2-34(c) and 2-34(d). We characterize this split as an emergence of a pseudo-component which does not have a direct physical meaning and is a result of oversampling. We finally note that the observed fluctuations in the position of the dip in a pseudogap-like state [Fig. 2-35] are likely related to a slightly different extraction of a “noise” component (endmember 3) from the full dataset in each case.

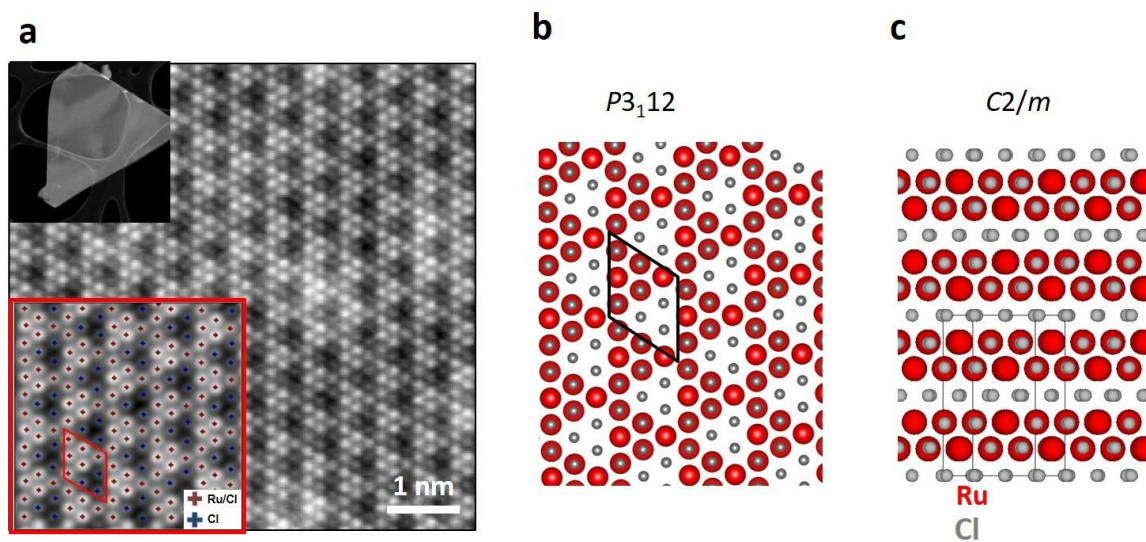


Figure 2-16. Comparison of room temperature STEM data with proposed structural models.

(a) Experimental STEM image; top left inset shows location of a flake on the grid. (b) $P3_12$ structural model. (c) $C2/m$ structural model. See also Supplementary Note 2-1.1.

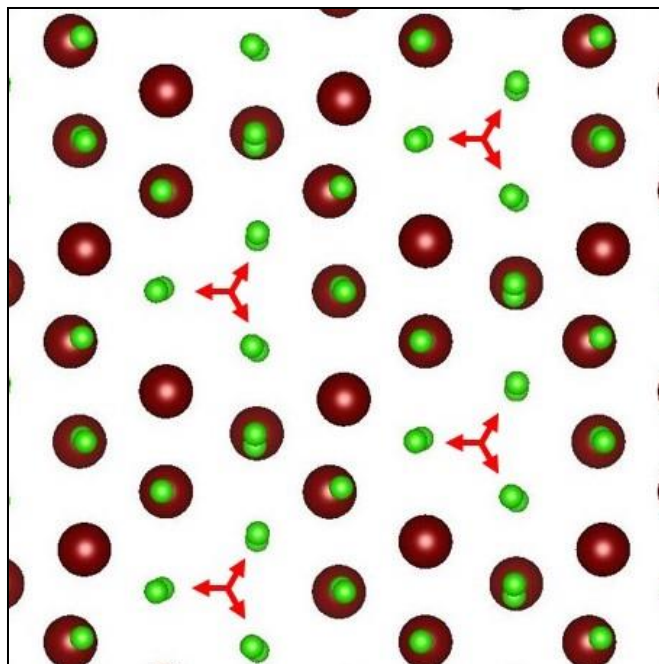


Figure 2-17. Schematics of octahedral distortion found in neutron diffraction data on single crystals (top view) that leads to enlargement of Cl-Cl intercolumn distances (schematically denoted by red arrows); this agrees with STEM observations on exfoliated films. See also Supplementary Note 2-1.2.

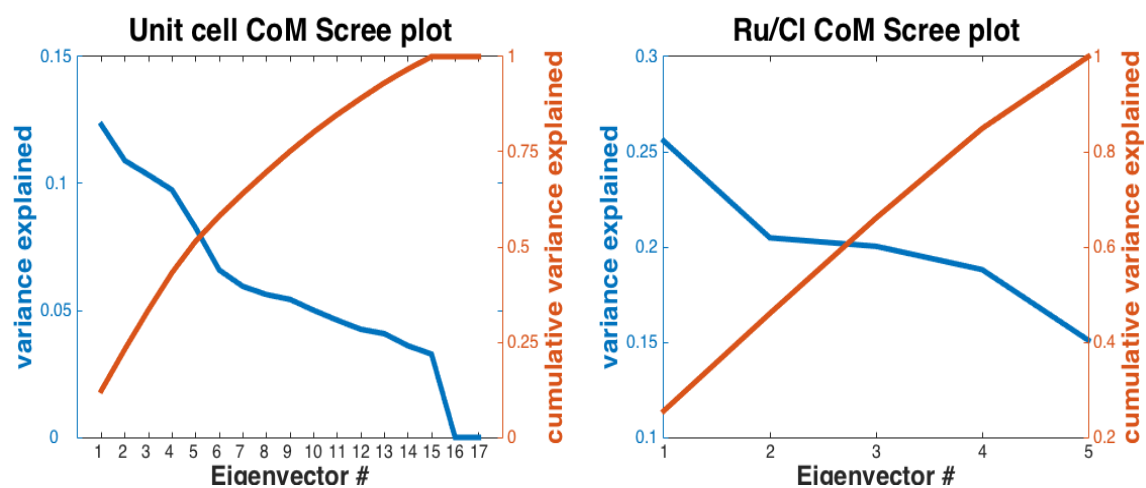


Figure 2-18. PCA scree plots.

(a) PCA scree plot for center of mass (CoM) of entire unit cell as origin and (b) center of mass of 6 Ru/Cl columns as origin.

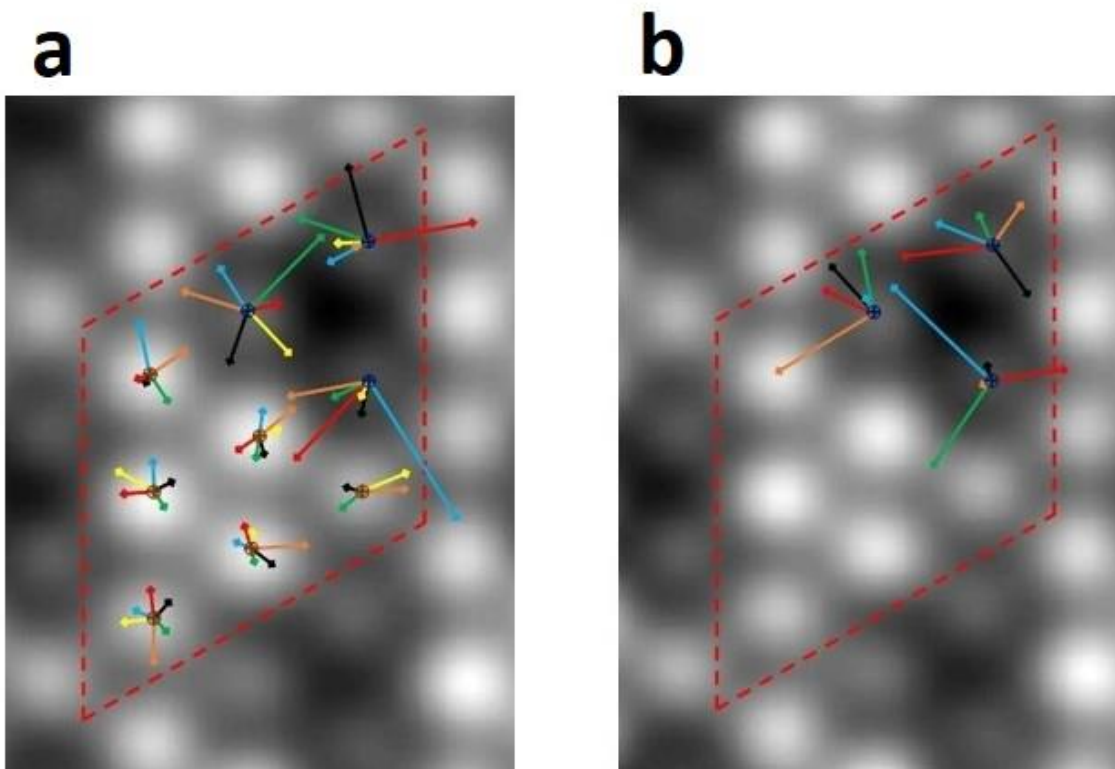


Figure 2-19. PCA-derived first 6 normal modes of displacement for (a) center of mass of entire unit cell as origin; (b) center of mass of 6 Ru/Cl columns as origin.

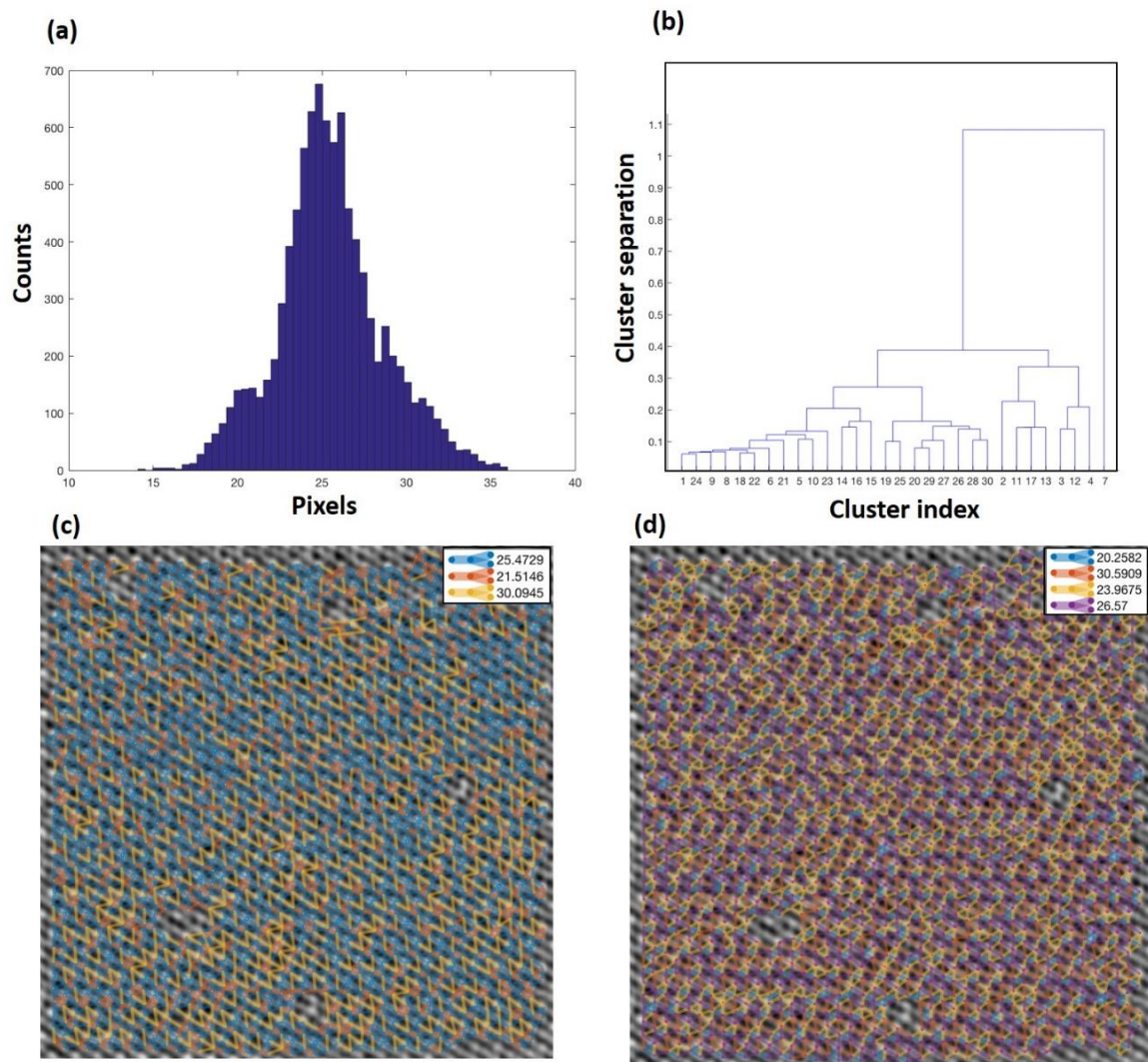


Figure 2-20. The k-means clustering analysis.

(a) Histogram of 6 nearest neighbor distances distribution from the STM image described in Fig. 4a of the main text (b) Dendrogram plot of hierarchical binary cluster tree. (c,d) Real-space data for distribution of the nearest-neighbor distances for 3 clusters (c) and 4 clusters (d).

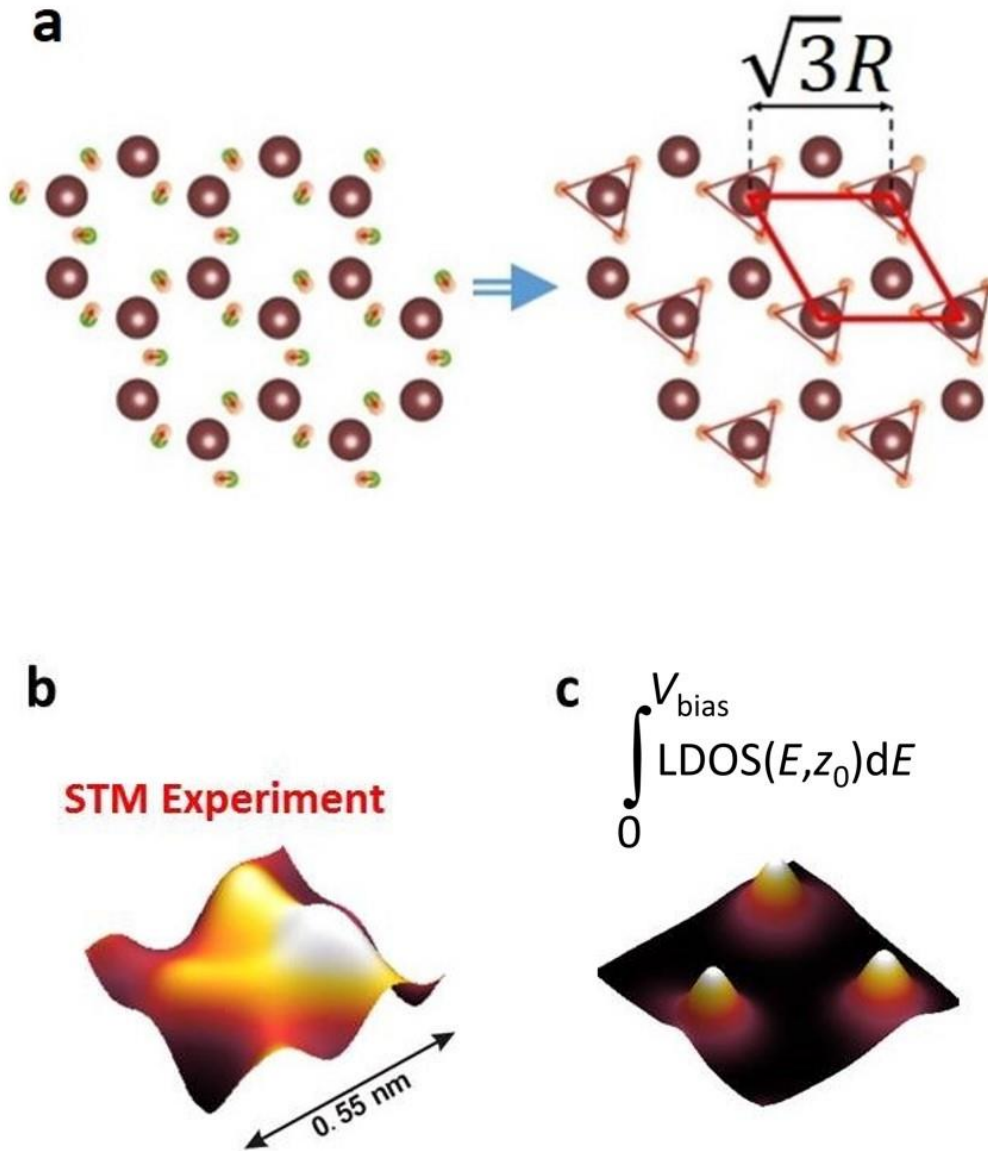


Figure 2-21. Lattice distortion as a possible origin of observed charge order.

(a) Left: Schematic view of the lateral component (projection onto xy plane) of Cl atoms displacement. The initial and final positions of atoms are denoted by green and orange balls, respectively. The actual displacements are exaggerated for a purpose of better visualization. Right: Projection of the resultant position of Cl atoms in the top layer that forms a structure with a $\sqrt{3}R \approx 600\text{pm}$ periodicity. (b, c) Experimental (b) and DFT-simulated (c) STM images showing a distribution of charge density at 3 “top” Cl atoms in the unit cell at bias voltage $V_{\text{bias}}=0.3\text{V}$. No charge modulation was found within the current DFT scheme.

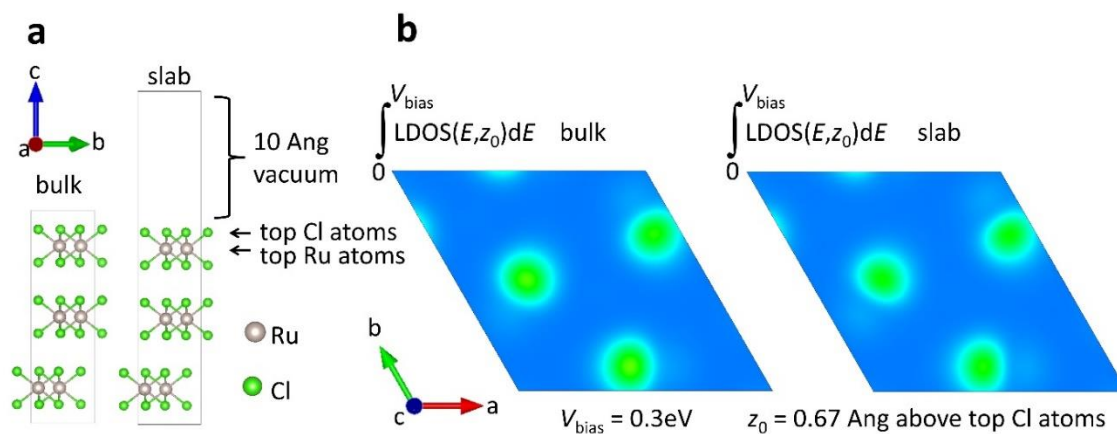
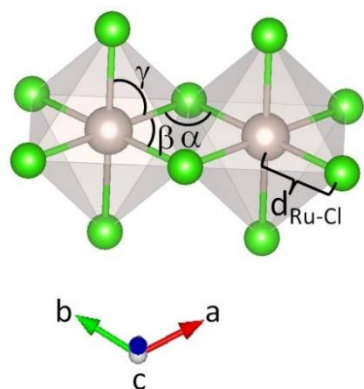


Figure 2-22. DFT STM simulations for bulk and surface models.

(a) Side view of bulk and slab unit cell. (b) Top view of energy integrated Local Density of States (LDOS) at 0.67 Ang above top Cl atoms from the bulk and the slab unit cell respectively.



	$\text{Ru}_6\text{Cl}_{18}$ bulk	top-layer $\text{Ru}_6\text{Cl}_{18}$ slab
$d_{\text{Ru-Cl}}$ (Ang)	2.362	2.363
α	92.48	92.44
β	87.52	87.56
γ	90.85	90.83

Figure 2-23. Comparison of average structural parameters obtained from the bulk and the top-layer of the slab.

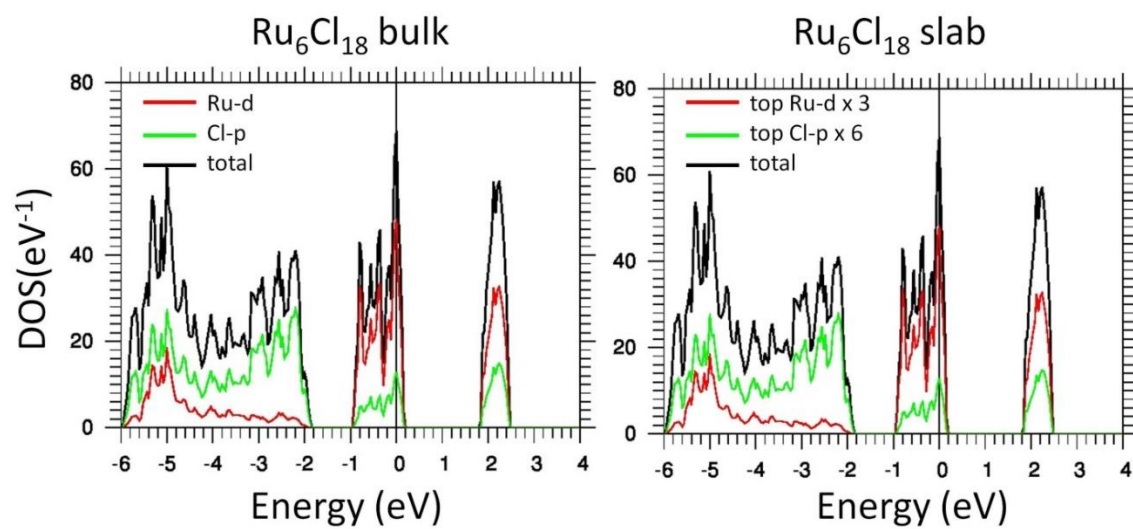


Figure 2-24. Orbital resolved Density of States (DOS) from the bulk and the slab unit cell respectively.

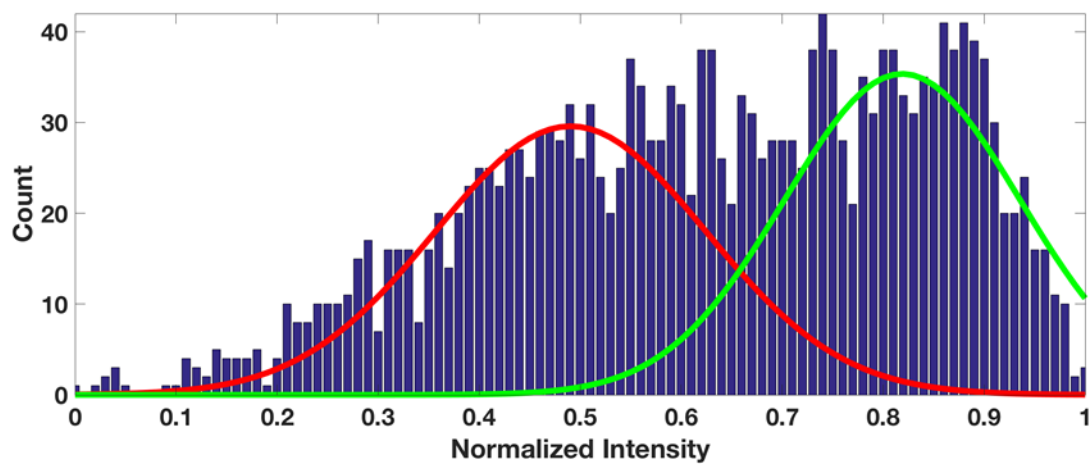


Figure 2-25. Histogram of collected atomic intensities and best fit of 2 Gaussian distributions to represent dark and bright atoms.

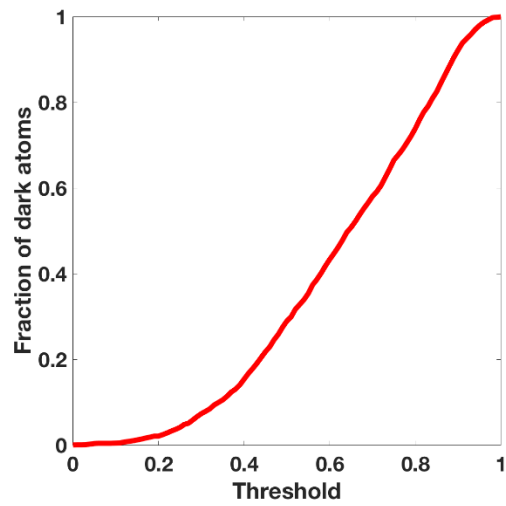


Figure 2-26. Fraction of dark atoms identified within the image as a function of selected threshold value. At threshold of 0.61 fraction of dark atoms is 45%, so it was chosen for the further analysis.

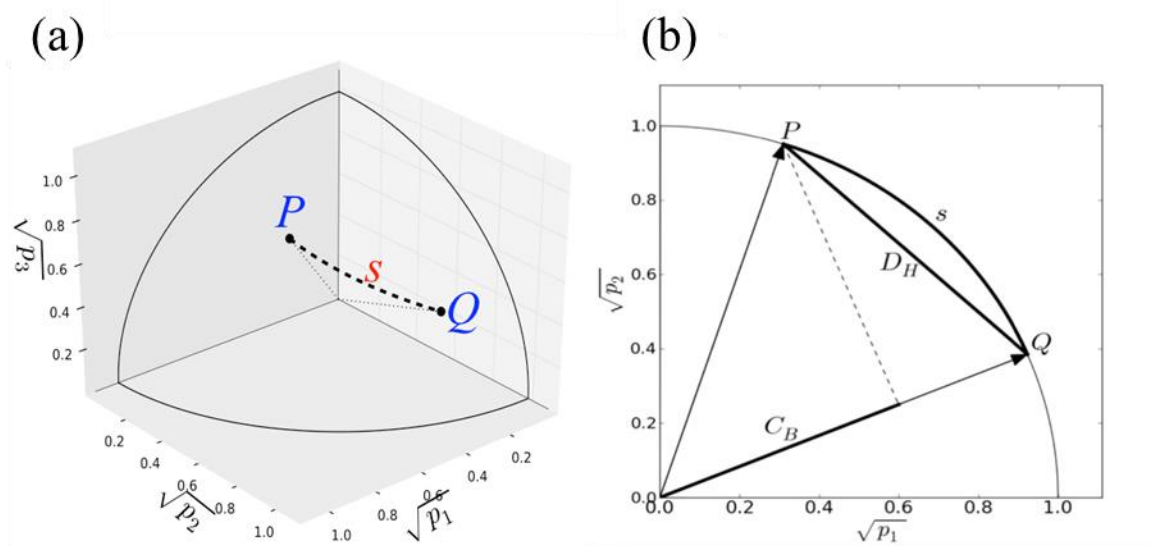


Figure 2-27. Illustration of statistical distance method.

(a) The probability space of trinomial distributions with a statistical distance s between distributions (i.e., statistical mechanical systems) P and Q . (b) The relations between the geodesic statistical distance s , the Euclidian Hellinger distance D_H , and the scalar product of P and Q vectors known as the Bhattacharyya coefficient C_B , on the example of binomial distributions

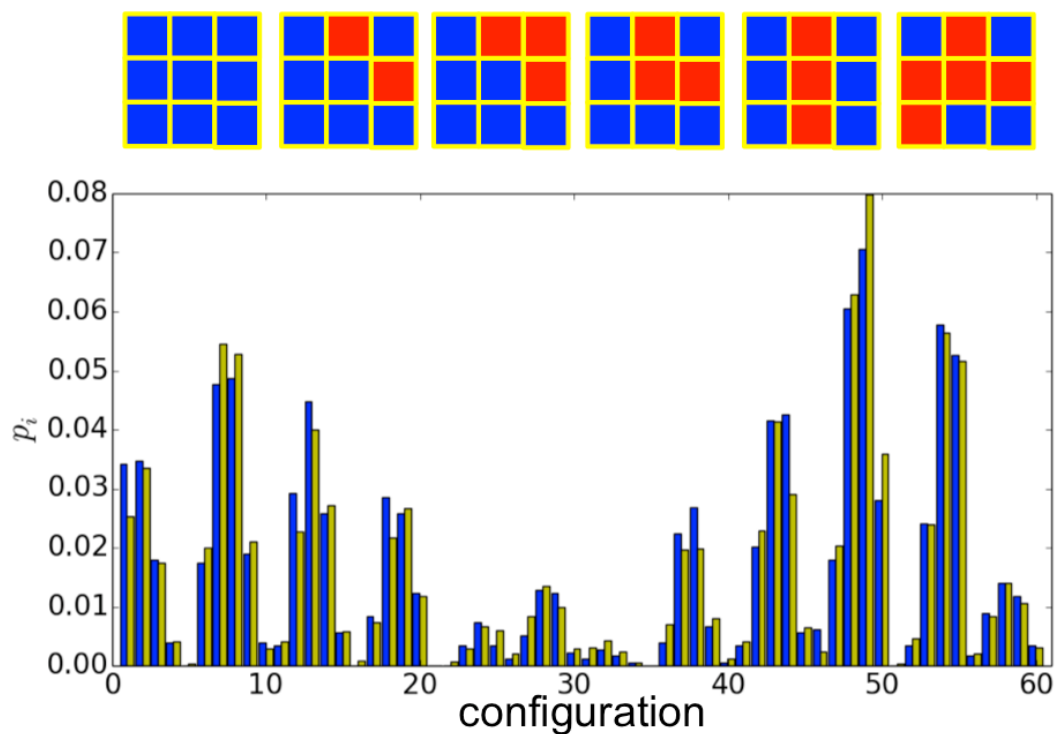


Figure 2-28. Top: Examples of six out of sixty local configurations, whose statistics was collected from the image and used as target information for model optimization. Bottom: The histograms of 60 local configurations in the target (blue) and model (yellow) systems.

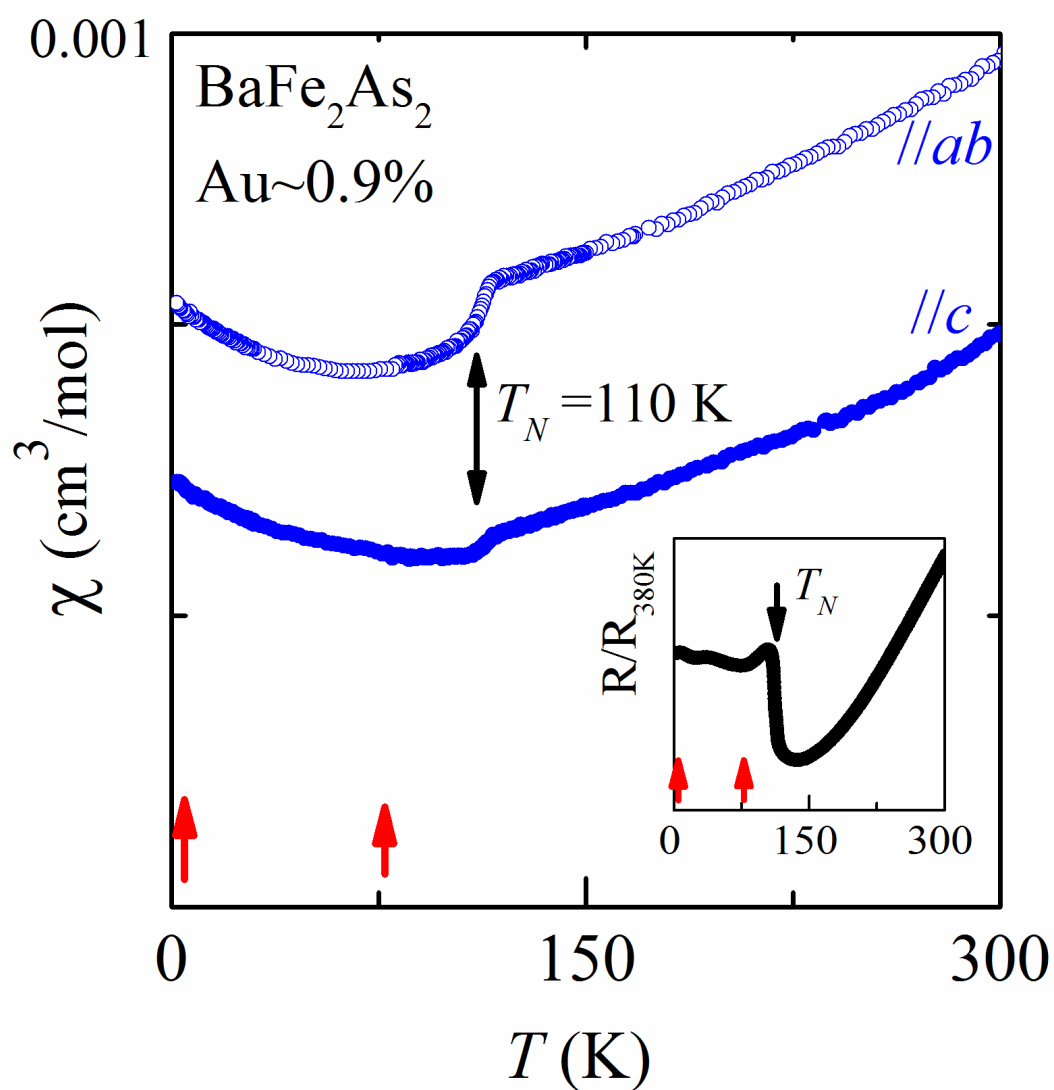


Figure 2-29. Temperature-dependent magnetic susceptibility and resistance (inset) results of BaFe_2As_2 , lightly doped with $\sim 1\%$ gold ($x=0.009$). χ is measured along the two crystallographic axes; T_N is inferred from data; red arrows indicate STM/S temperature points at 4 K and 77 K.

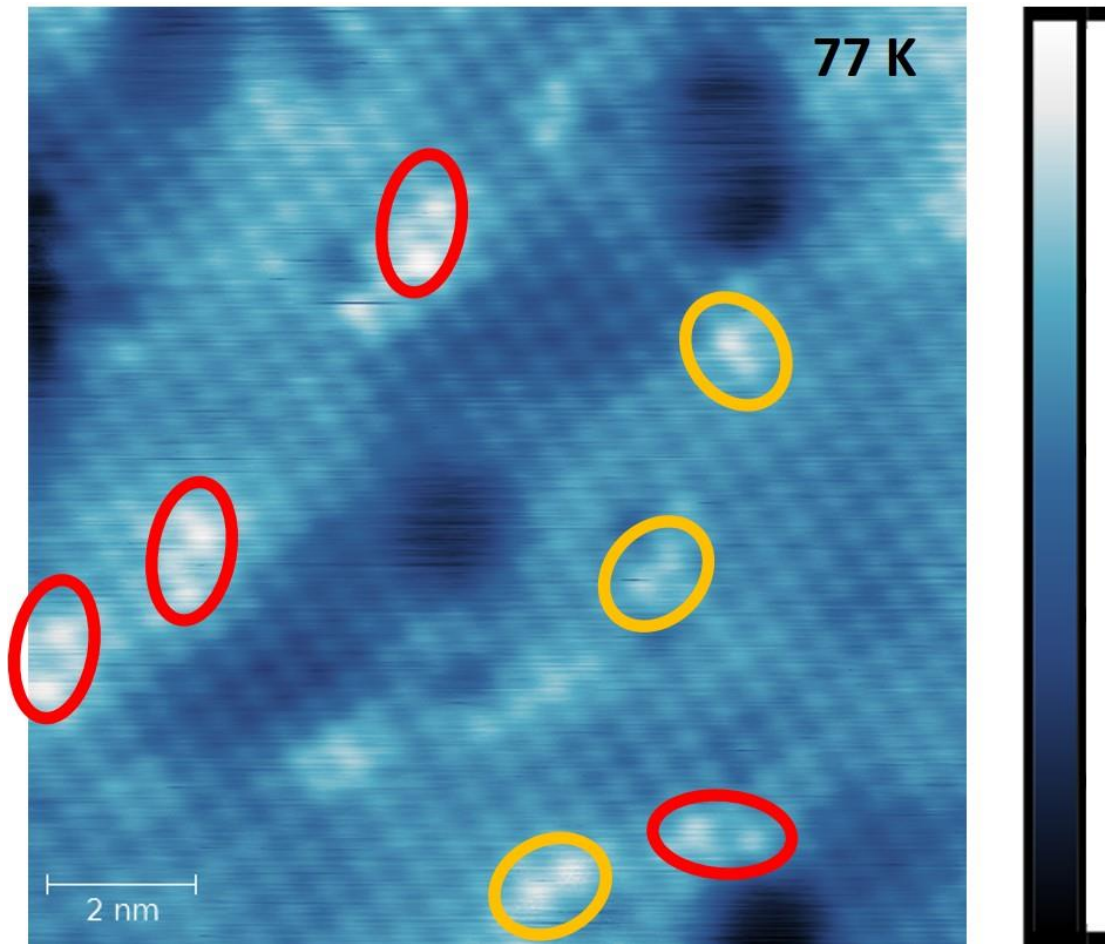


Figure 2-30. STM topographic image recorded at $T=77$ K for 1% Au-doped BaFe_2As_2 . Tunneling conditions $U_{\text{bias}}=100$ mV, $I_{\text{setpoint}}=1$ nA. Two common types of point defects are marked with red and orange ovals.

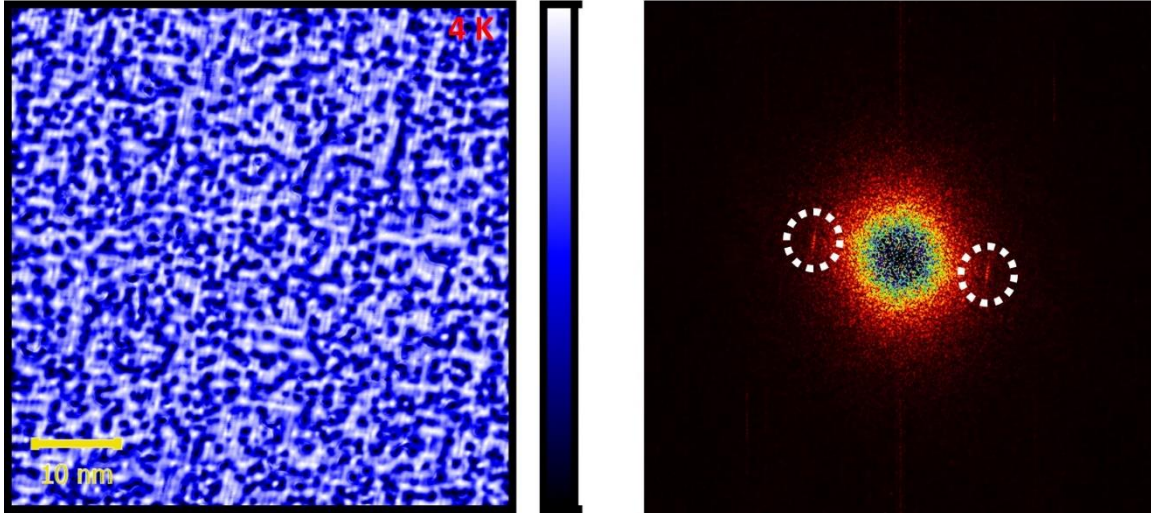


Figure 2-31. (Left) STM topographic image recorded at $T=4$ K for 1% Au-doped BaFe_2As_2 . Tunneling conditions $U_{\text{bias}}=-50$ mV, $I_{\text{setpoint}}=200$ pA. (Right) FFT on the topographic image on the left. Two spots corresponding to 1D modulation of a charge density are denoted by dotted circles.

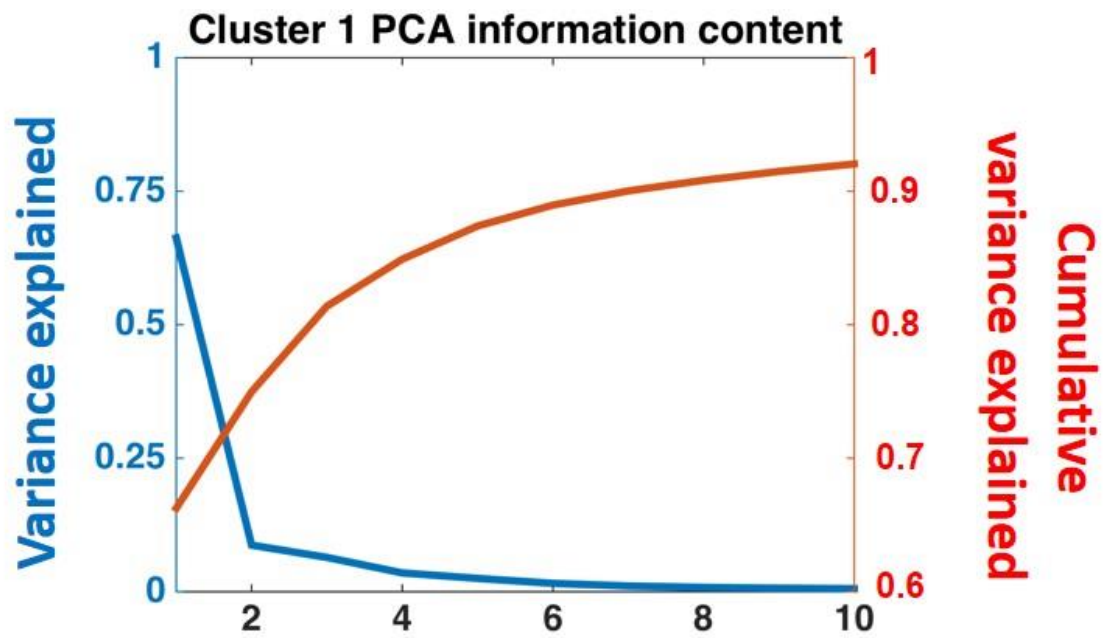


Figure 2-32. PCA scree plot describing variance in *k*-means derived cluster 1 (see main text)

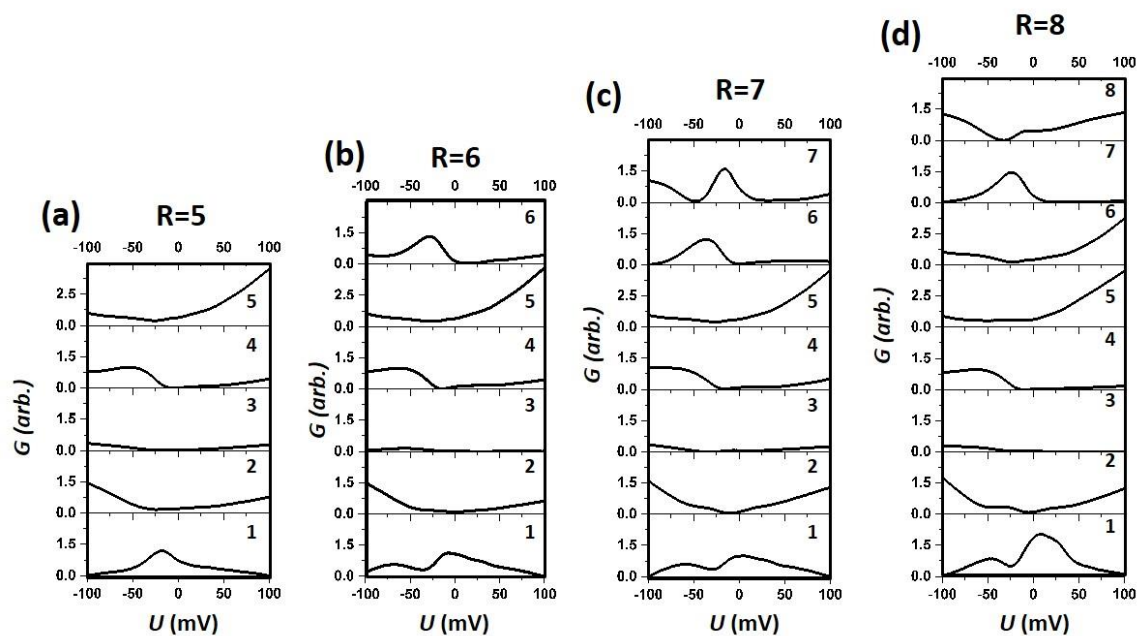


Figure 2-33. Bayesian endmembers for different number of total endmembers: $R=5$ (a), $R=6$ (b), $R=7$ (c), $R=8$ (d).

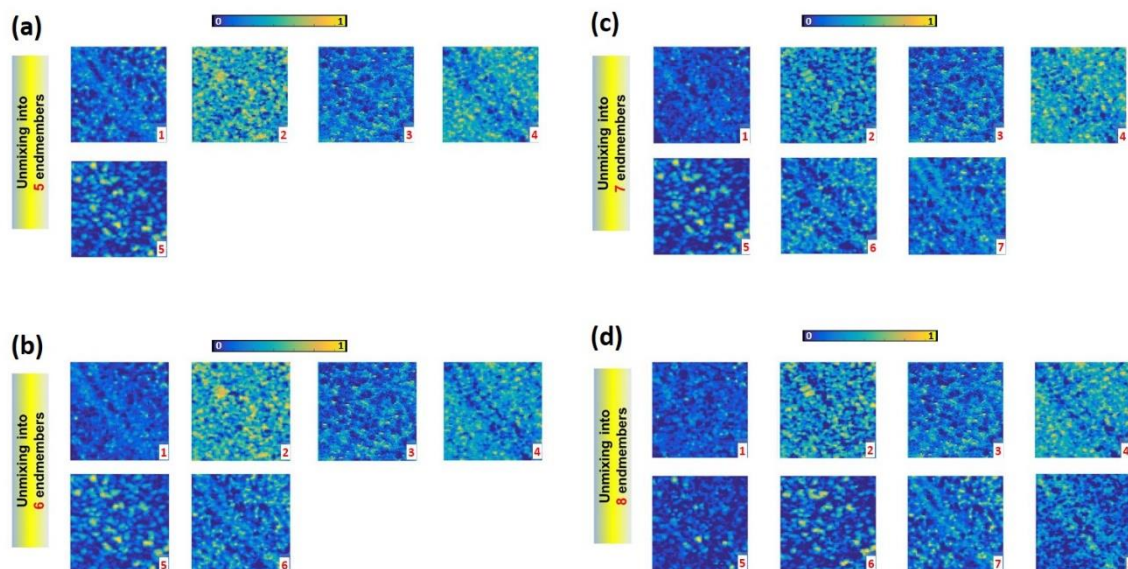


Figure 2-34. Abundance maps associated with different number of total Bayesian endmembers in Fig. B2: $R=5$ (a), $R=6$ (b), $R=7$ (c), $R=8$ (d).

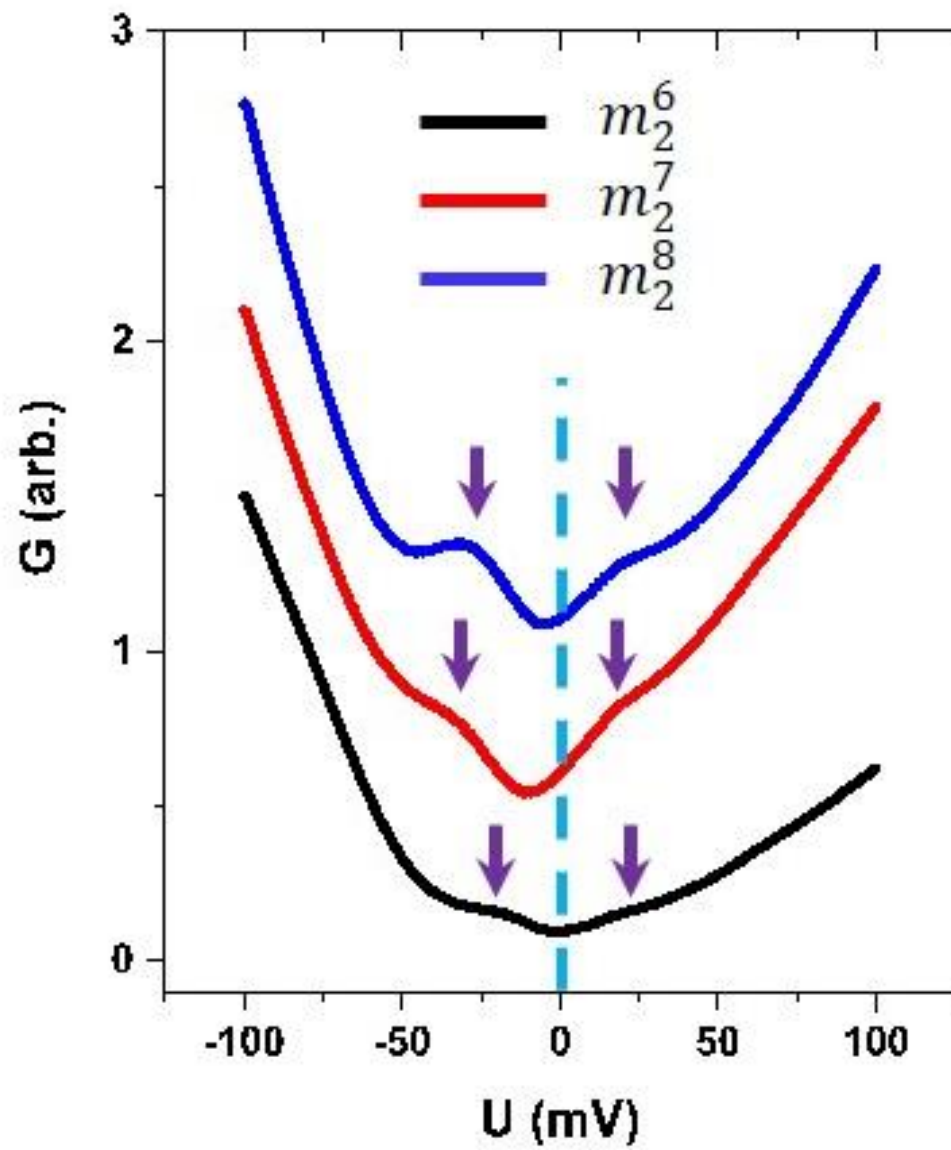


Figure 2-35. Endmember associated with a pseudogap-like state for BLU unmixing into $R=6$, $R=7$, $R=8$ components. Arrows and dashed line are guides for eye only.

CHAPTER 3.
TEACHING PHYSICS TO DEEP LEARNING MODELS

3.1 Introduction

Chapter 2 described some types of challenges we can encounter in materials science which require extraction of physically meaningful information from data. However, in all of those cases I had to create a customized way to achieve that. In this chapter I will address this issue by describing more universal approaches based on deep learning.

In the first two parts of this chapter, I will describe a deep learning framework based on fully-convolutional neural networks that allows identification of atomic positions in both STEM and STM imaging, in various types of lattices, without recreation of a training set, or hardcoding lattice and noise type assumptions, which is a significant breakthrough compared to methods described in Chapter 2.

The third part of this chapter addresses a slightly different challenge of classifying molecular states in STM images. While we have simulated images of molecules, and we know long-range periodicity in the material self-assembly, real experimental data is still very different, both visually and in terms of lattice disorder. However, we can introduce known physical constraints into the classification by using a combination of deep convolutional neural network and a Markov Random Field model.

Finally, once we know positions and classification of atoms or molecules in a material, we can use the data from large amounts of experiments or simulations to discover physics that cannot be recovered from a single data point. In particular, I cover utilization of variational autoencoder for uncovering phase transitions in complex materials in unsupervised manner.

3.2 Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations.

3.2.1 Publication information

A version of this article was originally published by Maxim Ziatdinov, Ondrej Dyck, Artem Maksov, Xufan Li, Xiahan Sang, Kai Xiao, Raymond R. Unocic, Rama Vasudevan, Stephen Jesse, and Sergei V. Kalinin in:

Ziatdinov, Maxim, Ondrej Dyck, Artem Maksov, Xufan Li, Xiahan Sang, Kai Xiao, Raymond R. Unocic, Rama Vasudevan, Stephen Jesse, and Sergei V. Kalinin. "Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations." *ACS nano* (2017). [194]

The article content is reprinted as published with figure numbers modified and supplementary material added at the end of the chapter as an appendix.

3.2.2 My contribution

Amount of imaging data collected by STEM has grown significantly. Fast, reliable, reproducible analysis of such data is more important, and applying deep learning for identification of atomic locations is a novel way to process such data.

The data for this project consisted of STEM images of graphene and MoSe₂ with defects. Both materials have hexagonal lattice. Traditional methods of atomic position extraction such as thresholding failed to provide physically meaningful results due to atoms forming continuous hexagons instead of local intensity maxima.

In order to achieve the goal of identification of features in the image and classification of them into carbon atoms, dopants, and vacancies, I proposed utilization of fully-convolutional neural networks. I proposed generation of training set from simulated atomic structures and augmenting them through image manipulation, as well as modification of training set after studying the results in conjunction with domain expertise. This allowed

extraction and classification of all features of interest in the images, allowing tracking of features in time and further refinement of the framework through weakly-supervised approach by performing graph analysis of the defect neighborhoods.

3.2.3 Abstract

Recent advances in scanning transmission electron and scanning probe microscopies have opened exciting opportunities in probing the materials structural parameters and various functional properties in real space with angstrom-level precision. This progress has been accompanied by an exponential increase in the size and quality of data sets produced by microscopic and spectroscopic experimental techniques. These developments necessitate adequate methods for extracting relevant physical and chemical information from the large data sets, for which *a priori* information on the structures of various atomic configurations and lattice defects is limited or absent. Here we demonstrate an application of deep neural networks to extract information from atomically resolved images including location of the atomic species and type of defects. We develop a “weakly supervised” approach that uses information on the coordinates of all atomic species in the image, extracted via a deep neural network, to identify a rich variety of defects that are not part of an initial training set. We further apply our approach to interpret complex atomic and defect transformation, including switching between different coordination of silicon dopants in graphene as a function of time, formation of peculiar silicon dimer with mixed 3-fold and 4-fold coordination, and the motion of molecular “rotor”. This deep learning-based approach resembles logic of a human operator, but can be scaled leading to significant shift in the way of extracting and analyzing information from raw experimental data.

3.2.4 Introduction

In the last decade, the proliferation of electron microscopy and scanning probe microscopy techniques have generated massive amounts of data on local chemical structure and atomic transformation.[195-197] Since the advent of aberration corrected Scanning Transmission Electron Microscopy (STEM), atomically resolved images of multiple materials classes

ranging from multiferroics, semiconductors, and superconductors have become common.[198-202] The further impetus to this field was given by development of atomically resolved dynamic studies, when the dynamic changes in matter on the atomic level are visualized. These traditionally include the thermal and chemical processes enabled by advanced thermal and environmental holders.[203, 204] More recently, progressively more attention is being attracted to the dynamic processes induced by the electron beam irradiation,[205-208] especially promising in the context of e-beam atomic fabrication.[209-211]

Similar advances are achieved in the field of atomically resolved scanning tunneling (STM) and atomic force microscopy (AFM). The recent famous examples include direct imaging of chemical bonds in molecules,[212] visualizing atomic collapse in artificial nuclei on graphene,[213] and inferring mechanisms behind fundamental physical phenomena, such as high- T_c superconductivity, from single atom defect induced scattering patterns.[214] In addition, the STM allows manipulating matter on atomic and molecular levels via mechanochemistry[215] or by electron injection from the STM tip.[216]

The characteristic aspect of both STEM and SPM fields is the generation of a large volume of high-veracity experimental data in the form of static or dynamic images. While providing immediate visualization of atomic structures, the interpretation of this data is traditionally limited to quantitative aspects, e.g. highlighting features such as presence of structural and topological defects, interfaces, etc. In the last decade, a number of approaches emerged based on the quantification of STEM and STM data. In these, atomic coordinates extracted from image analysis are projected on (postulated) mesoscopic order parameter field, providing information on local ferroelectricity, chemical strains, and octahedral tilts. The workflow for such analyses typically comprise the steps of the identification and refinement of atomic positions and subsequent transformation of measured atomic coordinates into the physical quantities. In STEM, this process is enabled by the fact that the ideal image represents a nucleus and the STEM transfer function is usually monotonic and well behaved; in STM and TEM analysis can be more complicated. However, in all cases to date physical models have been postulated by the observer based

on the qualitative observations, and quantitative information was obtained within the framework of the imposed model.

In the last several years, advances in machine learning brought significant developments to multiple areas of science and engineering, particularly (qualitative) image recognition.[217, 218] The main goal of the machine learning is to generalize from the available training examples to make accurate predictions/classifications on data samples that were not part of the training set. It is currently believed that one of the most promising areas of machine learning is the so-called deep learning (DL), which is the machine learning technique based on (deep) artificial neural networks.[219] The DL models can be used for prediction and classification of 2D and 3D, spatial and temporal data, such as images and videos. Despite that some of the key concepts of artificial neural networks have been researched as early as 1958,[220] it was only several years ago, partially due to overcoming computational limitations, that they gained widespread recognition. In particular, the convolutional neural networks were responsible for major breakthroughs in image and video recognition tasks,[46, 219] with seminal breakthroughs such as distinguishing with high accuracy between images of cats and dogs available on the Internet. Since then there was an explosion in applications of DL in various areas of everyday life (e.g., tagging photos on Facebook) as well as in scientific and engineering research (e.g. cancer detection, self-driving cars, and satellite imaging). Recently, the DL framework was applied to several problems in theoretical physics such as the search for new exotic particles at the Large Hadron Collider[39] and detection of phase transitions in lattice models.[221, 222]

Here, we present a “weakly supervised” approach to deep learning of atomically-resolved images, in which by starting with limited *a priori* information about types of defects in a sample, that is, a limited number of available labels in the DL scheme, we use the information on local atomic coordinates to identify a rich variety of new defect structures that were not explicitly included in the initial training set. Specifically, by applying Laplacian of Gaussian blob detection technique to the output of the fully convolutional neural network, we were able to make a transition from the classification of image pixels to chemistry-based classification of defects based on the position of chemical species, bond

coordination, bond length and bond angle. Using this approach, we identified multiple novel defect structures associated with silicon atoms implanted into graphene vacancies, as well as observed various chemical and structural transformations of defects on the surface for which we could identify both the location and chemical structure of the defect for each individual step of a “reaction”. This approach resembles the approach of a human operator with a background knowledge of the basic materials science concepts but without specific information on the structural and functional peculiarities in each particular material.

3.2.5 Results and Discussion

In order to obtain high quality, atomically resolved STEM images for the development and analysis described, we used a Nion UltraSTEM U100 STEM operated at 60 kV. Graphene samples were used as a model test case because of their robustness against the 60 kV beam and the ease with which interesting and highly-visible defects may be found or created. Moreover, since these defects are known to evolve/restructure under e-beam influence and each structure is usually only a single atom thick (providing clear delineation of each atom’s location), this system offers an ideal test bed for the methods described. In addition, we tested our methods on the STEM images of monolayer MoSe₂ doped with tungsten, Mo_{1-x}W_xSe₂.

For the analysis of the atomically-resolved experimental data on the model systems we employed the DL neural network model(s) that has an encoder-decoder type of architecture (Fig. 3-1a).[223] The encoder part consisted of convolutional layers for feature extraction followed by max-pooling layers for reducing the size of processed data. The number of filters in the convolutional layers was doubled after each max-pooling layer. The decoder (or “deconvolutional”) part of the network, whose role was to map the encoded low-resolution feature maps to full input-resolution feature maps, consisted of the same filters but in reverse order and *un*-pooling (or up-sampling) layers. The feature maps from the last convolutional layer of the network were passed to a softmax layer for the pixel-wise

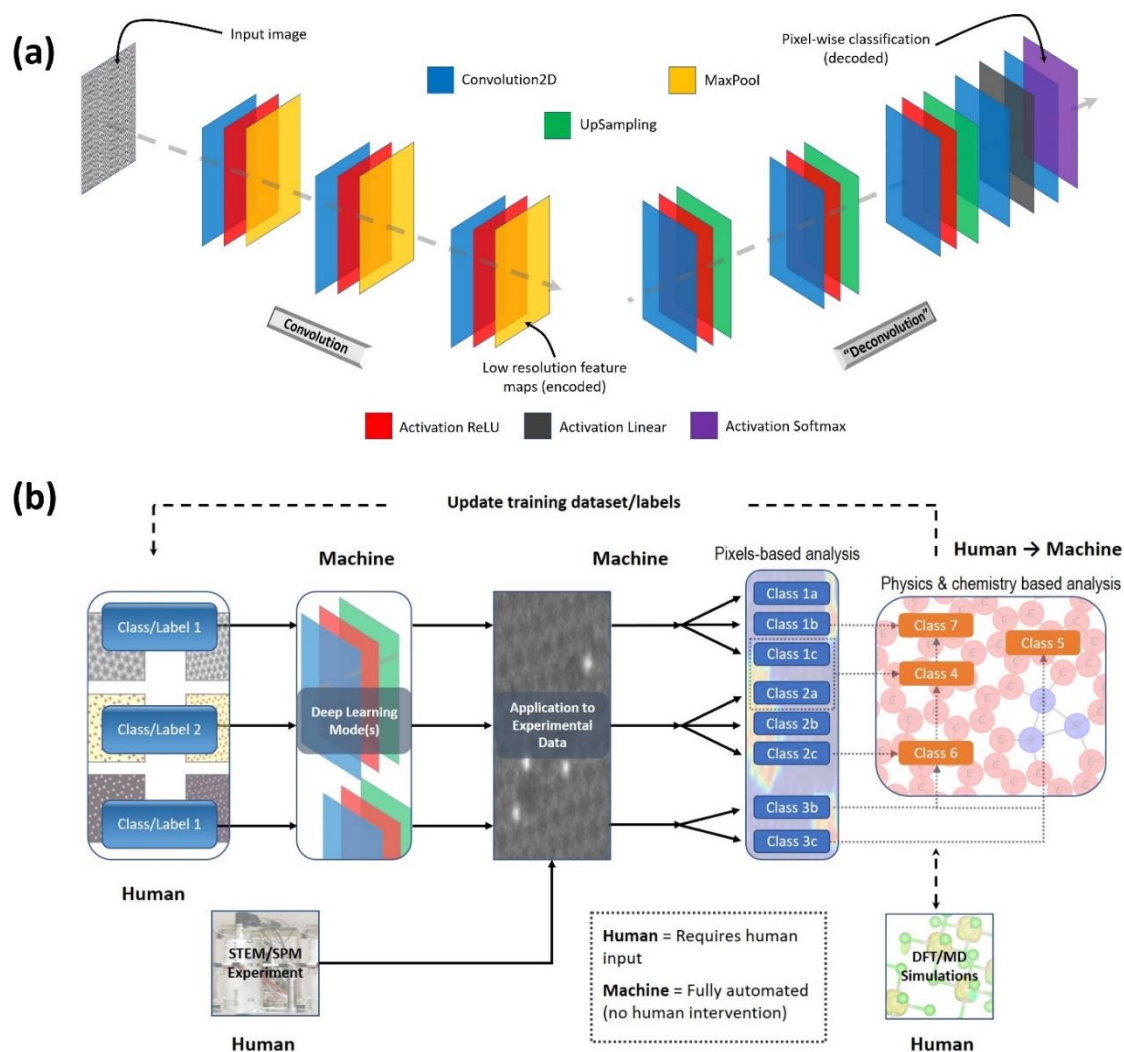


Figure 3-1. Application of deep learning to a problem of finding atomic species and defects in a crystal lattice.

(a) Schematic architecture of a fully convolutional neural network (FCN) that has an encoder-decoder type of structure (or convolution-“deconvolution” structure). The final softmax layer outputs a pixel-wise classification for atomic species and/or defects. (b) Schematics of a “weakly-supervised” approach towards identifying lattice configurations and defects in the experimental data (see text for details). The parts of the process that require (do not require) human input are denoted as ‘Human’ (‘Machine’).

classification providing us with information on the probability of each pixel being an atom or specific type of defect, depending on the classification scheme used. As we did not employ a fully connected, “dense” layer in our network architecture, these networks are referred to as fully convolutional networks (FCNs).[224]

We train an FCN model using the simulated, theoretical images of atomic lattice and defects and then apply the trained model to the experimental data. This is different from most of the approaches towards training of DL models, which usually utilize some existing database of experimental images with corresponding labels, and hence cannot be used to study novel materials for which no such databases exist. The theoretical images can be generated using atomic coordinates obtained from ab-initio or molecular dynamic simulations of the corresponding atomic structures. In the case of STEM images, the pixel intensities in the generated images are proportional to the square of atomic number Z for each atomic column. For the STM images, the intensity of the image is associated with electronic density of states at the Fermi level. In this paper we focus on the analysis of STEM experimental data; however, this approach is applicable to the STM datasets as well. In the following we are going to adopt a learning approach that could be generally referred to as a “weakly-supervised learning”. The idea behind this approach is that one starts with a few simple types (classes) of defect structures that enable a sample description in a very general manner (for example, classes associated with lattice and with 1-2 types of defects), trains a classifier, and then applies it to experimental data. In such a case, the output of the DL model will typically show, in addition to the classes that it was trained on, structures that can be an agglomeration of defects belonging to one class (e.g. clustering of point defects with various types of “bonding” to the lattice and to each other), those belonging to different classes (e.g. clustering of point defects along an extended 1D defect), as well as new subclasses for each class (e.g. presence of point defects associated with dopants of different chemical origin). These newly discovered structures can be then analyzed either qualitatively or quantitatively (e.g. through DFT or MD) and used to expand the training set and retrain the DL model. Note that this approach, which is schematically illustrated in Fig. 3-1b, is very similar to the way human experts explore and discover new types of physical and chemical behaviors in experimental data.

Following the ideas outlined above we start a construction of the training set. First, we note that most of the point (i.e. atomic-size) defects in typical STEM and STM datasets can be classified as belonging to two classes, commonly referred to as the depression and protrusion. The former includes missing atoms (STEM and STM), substitutional dopants with lower atomic number Z (STEM), and areas of depleted charge density at the Fermi level (STM), while the latter is commonly associated with adatoms (STEM and STM) and implanted dopants that have either a higher atomic number (STEM) or create additional states at the Fermi level (STM). We therefore prepare the training sets for 3 separate classes, namely, i) lattice without defects, ii) vacancy, and iii) dopant atom (with higher Z). The vacancy is created simply by removing a single carbon atom from the lattice. To create a structure with dopant, we remove a carbon atom(s) and place a different atomic element in its position whose intensity is $\sim 2 \cdot I_C$, where I_C is the STEM intensity of carbon atom in the graphene lattice. As our purpose is to “teach” a network only the basic concepts of what vacancies and dopants usually look like in the microscopic images, no relaxation of the lattice structure was performed after removing atom or implanting the dopant atom. We will show later in the paper how such a relatively simple training procedure allows finding and classifying more complex defects in the experimental data even from different materials that were not explicitly included in the training set. To account for a variation in instrumental parameters during image acquisition (e.g. different sample orientation, noise and blurring levels) as well as model uncertainties (e.g. sample thickness) we perform an augmentation of our training data by applying a set of image transformation operations to our theoretical image(s) (see also Methods and Figure 3-26).

Pixel-based classification

The application of the trained FCN model(s) to experimental data allows us to identify positions of individual atoms in the lattice as well as location and type of different types of defects. Figure 3-2 illustrates a procedure for obtaining atomic positions from raw experimental STEM data. First, the raw experimental image is fed into the trained FCN (accuracy on the test set $\sim 97\%$, see Figure 3-27) that outputs a map of probabilities (Pr) for each pixel being an atom (Fig. 3-2b, e). We then threshold the map

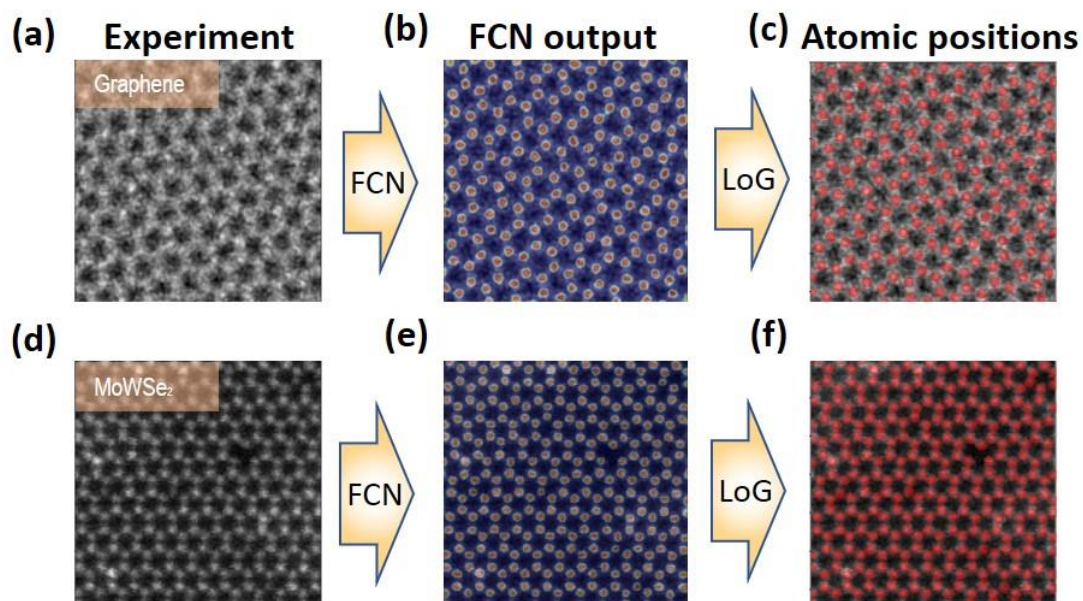


Figure 3-2. Finding atomic positions from raw experimental images.

(a-f) Illustration of how the atomic positions are obtained from raw experimental data on two different materials, graphene (a-c) and $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ (d-f), using fully convolutional neural network (FCN) and Laplacian of Gaussian (LoG) blob detection method. The FCN maps and atomic positions are overlaid on the experimental images.

at some specific value (usually at $Pr \sim 0.8-0.9$) to produce a binary image with small blobs of a circular shape corresponding to atoms in the lattice. Finally, the coordinates of the atomic centers are calculated via a Laplacian of Gaussian (LoG) blob detector technique[225] (Fig. 3-2c, f).

Remarkably, the FCN allows to find atoms in the vertices of hexagons even when atoms do not produce any characteristic local maxima in the image (Fig. 3-2a-c). This becomes possible because when tasked with finding atomic positions the FCN considers what is known as deep, high-level features in the image such as the shape of an individual hexagon (sides length, angles), as well as shapes of its neighbors. While the FCN was trained on theoretical images of graphene, we demonstrate that it is possible to use it for finding atoms in a different material that has similar but distinct atomic lattice structure. Specifically, we apply our FCN model to locating atoms in the STEM of $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ that has two inequivalent atomic sites (in terms of their intensities/atomic number) in the unit cell associated with a single Mo atom and with two stacked Se atoms.[202] As we show in Fig. 3-2d-f the network trained on simple honeycomb structure of graphene can accurately predict locations of atomic centers in the $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ system.

We then proceed to the deep learning based analysis of atomic structures in the presence of defects and lattice reconstructions. Figure 3-3a shows a STEM image of graphene with two types of defects: vacancies (lattice regions with missing carbon atoms) and dopants (Si atoms). The FCN outputs for Si dopants and vacancies are shown in Fig. 3-3c and 3-3d, respectively. It is worth noting that while the FCN was trained only on a honeycomb lattice with and without *single* vacancies it was able to i) find vacancies characterized by multiple missing lattice atoms, and ii) identify non-hexagonal, 5-7 reconstructions (Stone-Wales defects[226]) in the graphene lattice near the defect regions (Fig. 3-3e). Interestingly, in order to find vacancies, a neural network is essentially looking for regions of a surface where the hollow sites are larger than normal (see Fig. 3-3d). In the Supplemental Material we examine which surface areas a network is usually paying attention to when searching for vacancies and dopants. In general, our analysis of the DL network behavior suggests that the way these networks search for a vacancy, as well as dopants, is similar to the way

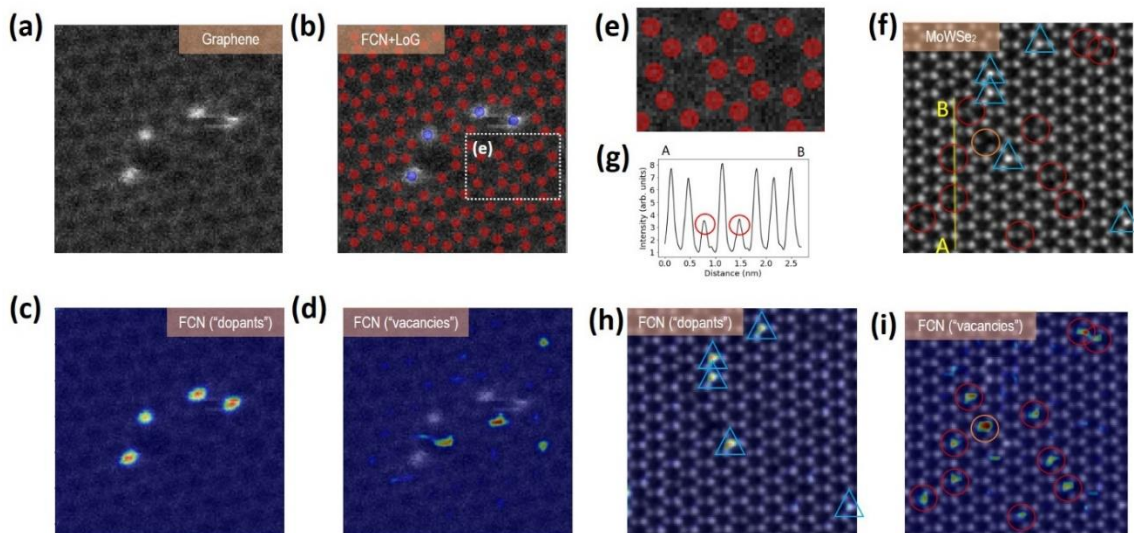


Figure 3-3. Pixel-wise classification of atomic defects and lattice reconstructions.

(a) Experimental STEM image of graphene with defects. (b) Application of FCN and LoG to extracting locations of atoms and silicon dopants from data in (a). (c, d) FCN output for class “dopants” (c) and class ‘vacancy’ (d) overlaid on the experimental image. (e) Zoomed area from the box in (b) showing that the DL model trained to find atoms in hexagonal graphene lattice was able to map the 5-7 reconstructions (formation of Stone-Wales defect) in graphene. (f) Experimental STEM image of $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ ($x \approx 0.18$), where the triangular and circle shapes correspond to the results from FCN outputs for class “dopants” (h) and class “vacancies” (i), respectively, overlaid on the experimental image. The orange circle in (f, i) denote a “full” vacancy with two missing Se atoms, whereas the red circle denotes a “half” vacancy with one missing Se atom. The intensity profile along the AB line in (f) is shown in (g).

the human eye would distinguish between regular lattice structure (with top, bridge and hollow sites) and a defect that breaks lattice periodicity.

Similar to our earlier analysis of the clean lattices, we apply the FCN model trained on graphene with defects to find dopants and vacancies in a single layer of a different material, $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ (Fig. 3-3f-i). The FCN output for a class “dopants” in $\text{Mo}_{1-x}\text{W}_x\text{Se}_2$ is shown in Fig. 3-3h. The model is clearly able to detect tungsten dopants which are characterized by a larger intensity in STEM.

It is worth noting that the FCN was able to distinguish accurately between the variations in the intensity associated with two different sublattices and that associated with the presence of dopants, even though such specific information was not included in our training dataset. We further analyzed the output of the FCN model for a class “vacancy”. Here our model found two types of blobs (denoted by orange and red circles in Fig. 3-3i) characterized by different intensities (softmax probabilities) and widths and associated with “full” vacancies (two Se atoms missing from a column) and “half” vacancies (one Se atom missing from a column[202]) (Fig. 3-3g). Importantly, it was not our initial goal to search for the “half” vacancies, nor was the network specifically trained to discover such structures.

Chemical structure-based classification

In the preceding section, we showed a deep learning based analysis of atomic lattice and defect structures based on classification of image pixels. We now demonstrate that it is possible to use DL networks, in combination with LoG and simple graph representations, to extract relevant structural/chemical parameters such as coordination number, bond lengths and angles, and to classify defects based on their chemical structures.

As a model example, we analyzed a 3-fold and 4-fold coordinated silicon defect implanted in a graphene lattice.[227, 228] By applying LoG to thresholded outputs of FCN for class “lattice” and class “defect” we obtained information about positions of atomic centers and locations of defects (Fig. 3-4a-e). We then extracted coordinates of the lattice atoms and, if applicable, other defect atoms within a specified radius ($\sim 3a-4a$, where a is a lattice vector) around the dopant atom(s). The extracted structure is represented by a graph

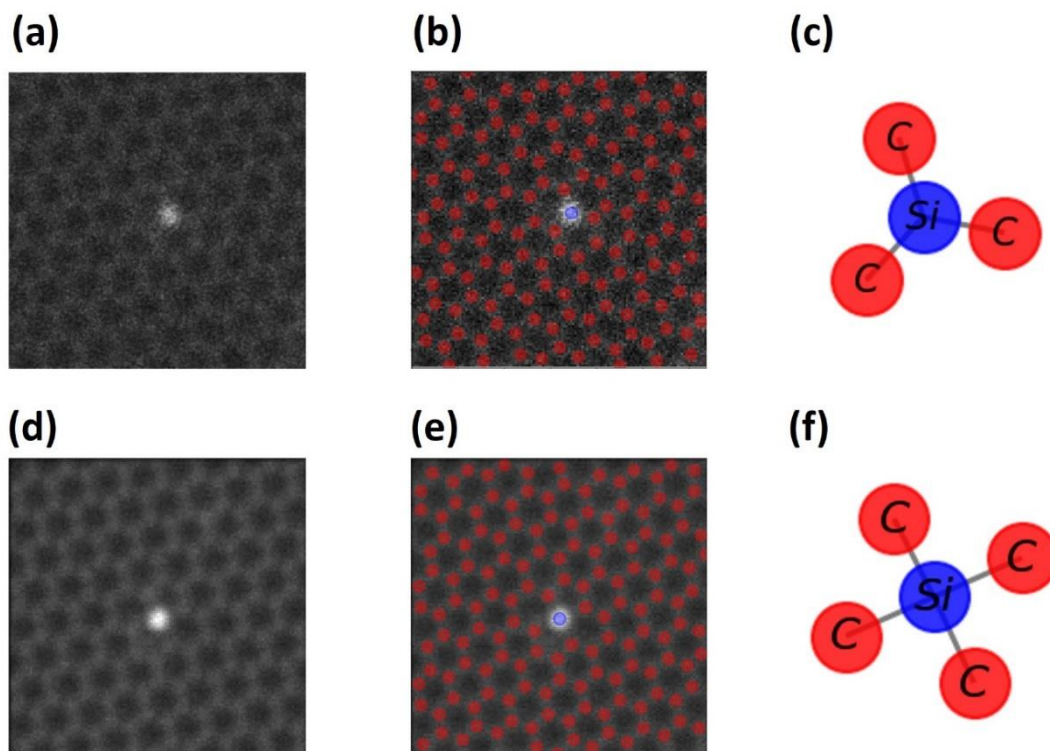


Figure 3-4. From analysis of pixels in image to analysis of single defect chemistry.

(a) Experimental image of 3-fold coordinated Si defect. (b) Result of applying FCN and LoG to data in (a). (c) Graph representation of single defect structure for 3-fold Si defect (see text for details). (d-f) Same for 4-fold coordinated Si defect.

$G = (N, E)$. We then applied a number of constraints based on simple chemistry rules to nodes N and edges E to adjust/modify our graph structure. Specifically, we introduced constraints for a maximum possible number of chemical bonds (edges) for an atom (node) of each type and for a maximum possible length for each type of chemical bond. In addition, we distinguished “edge” lattice atoms bonded to the “foreign” chemical species at the defect sites from the atoms in the bulk. Finally, for composite defects, we ensured that there were no unphysical connections between different parts of the defect, e.g. that a node representing a “vacancy” (region where atom(s) is missing) is not connected to nodes representing atomic species. The coordinates of the defect center were calculated as a “center of the mass” of the obtained graph nodes.

This approach allowed obtaining a graph representing a dopant atom(s) and a neighbor lattice (or other dopant) atom(s) to which it is connected directly via a chemical bond (Fig. 3-4c, f) while removing all the “irrelevant” atoms which do not form a direct chemical bond with a dopant atom. Such an approach represents a transition from a pixel-based classification of defects to a classification of defects based on their chemical structure. Indeed, in the case of the 3-fold and 4-fold coordinated Si atoms, the pixel-wise classification based on our initial training set categorized them as the same type of defect. In contrast, our chemistry-based classification scheme allowed identification of an important difference between the two dopants based on the details of their bonding to the carbon lattice. We can also use this approach to extract values of bond lengths and bond angles. For example, we found that a single 3-fold coordinated silicon dopant is characterized by a relatively small distortion in *apparent* Si-C bond lengths, $d_{Si-C} \cong 150$ pm. This can be explained by the presence of out-of-plane distortions of the Si dopant in the 3-fold coordination as predicted by DFT,[228] which is not possible to resolve in the STEM experiment. The angles between projections of Si-C bonds onto a 2D plane are approximately $\alpha_{C-Si-C} \cong 120^\circ$. In the case of 4-fold coordinated Si, the bond lengths and bond angles are $d_{Si-C} \cong 172$ pm and $\alpha_{C-Si-C} \cong 90^\circ$, respectively, which is in general agreement with DFT calculated parameters for planar Si-C structures (e.g. 177-179 pm in isoatomic 2D form $Si_{0.5}C_{0.5}$ [229])

We further extended the deep learning approach to track the transformation of a Si-lattice defect as a function of time. Specifically, we explored switching from 3-fold to 4-fold coordination and vice versa (Fig. 3-5) in a frame by frame fashion. The defect class was assigned based on its chemical structure, that is, the number of Si-C bonds, instead of a collection of pixels. This allowed us to obtain both the location of the defect and its type for each frame (Fig. 3-5b, c). We note that this can be extended to the analysis of more complex surface transformations involving multiple defects in larger image stacks and for larger surface areas.

We then applied this method to the analysis of complex defect transformations in graphene that involved more than one dopant atom. Figure 3-6a-d show four consecutive frames of the STEM experiment that starts with an image of a region with a single dopant and ends with an image of the same region but with two dopants that form a dimer-like structure. We first applied FCN and LoG to map the positions of all dopants and lattice atoms. This information was used to construct chemical structures of single defects shown in Fig. 3-6e-h. The initial configuration was a 4-fold coordinated Si (Fig. 3-6 a, e). It then underwent a transition to a distorted 3-fold coordinated silicon located next to a carbon vacancy (Fig. 3-6b, f). The vacancy then moved from right to left inducing a change in the defect orientation that resembled an application of reflection and translation operations to the defect (Fig. 3-6c, g). The presence of a carbon vacancy resulted in a significant distortion of the C_3 symmetry of the defect in frames 2 and 3 (Fig. 3-6f and 3-6g) compared to the nearly ideal structure of an isolated defect in Fig. 3-4a-c. The C-Si-C bond angles were (150°, 115°, 95°) and (165°, 102°, 93°) for defects shown in Fig. 3-6f and 3-6g, respectively. We therefore treated the defects in frames 2 and 3 as vacancy-Si-lattice complexes. This represents a situation in which the analysis based on the details of chemical structure of the defects and nearby lattice atoms allowed us to identify a new, composite type of defect that represents a mixture of two defect classes. Finally, in the fourth frame (Fig. 3-6d, h) the graphene lattice “captured” another Si atom which resulted in the formation of a peculiar $\frac{3}{4}$ -dimer structure. Here two Si dopants formed a covalent bond with each other and were bonded to two “edge” C atoms (“left” dopant) and three “edge” C atoms (“right” dopant), respectively (Fig. 3-6h).

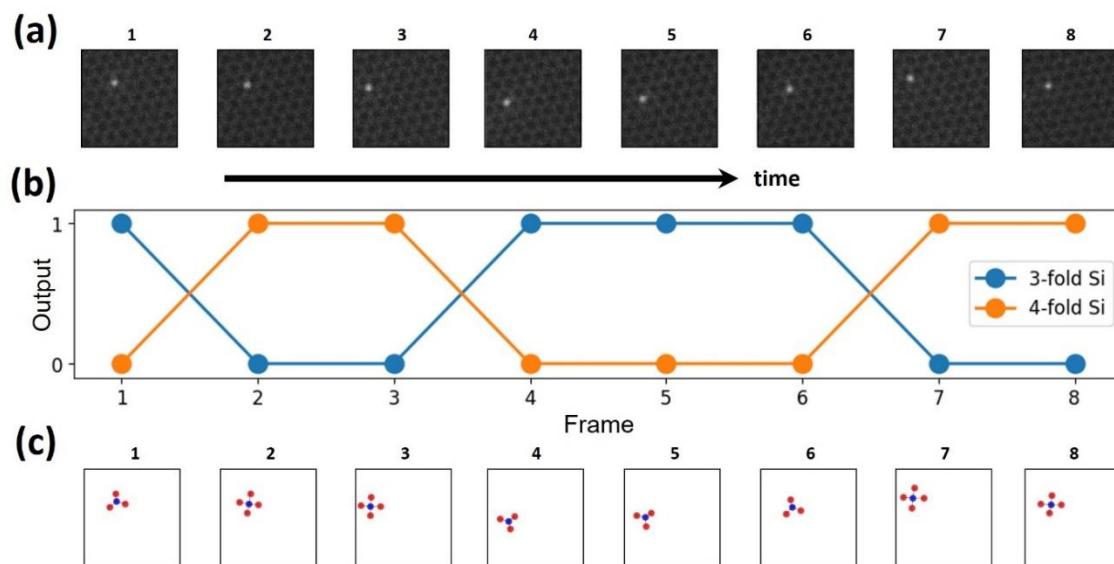


Figure 3-5. Tracking a reversible switching between different coordinations of Si dopant.

(a) STEM imaging of the same Si dopant over time. (b) Classification of the defect type (3-fold Si vs 4-fold Si) for each frame in (a). (c) Localization of the defect within each frame (the size of the boxes in (c) is identical to that of STEM images in (a)).

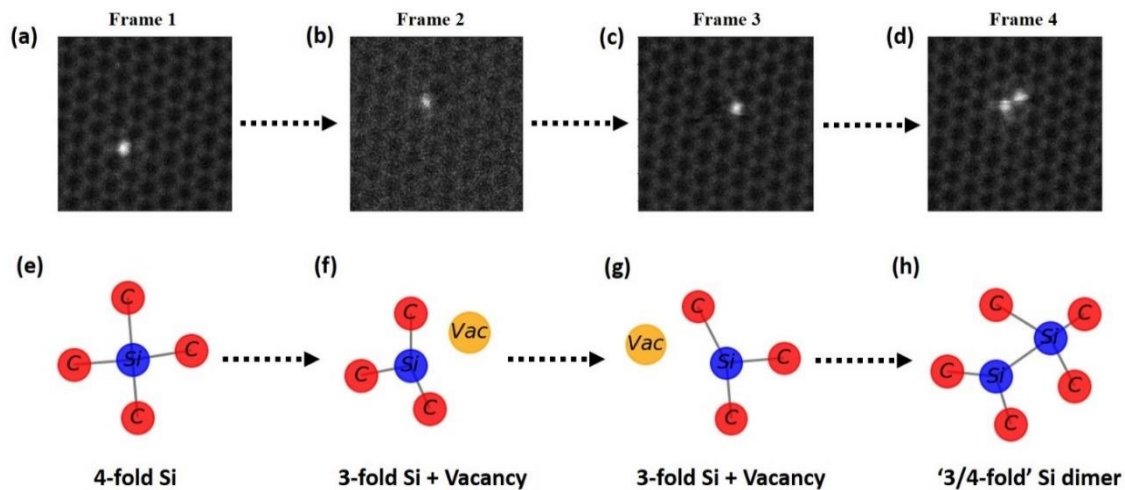


Figure 3-6. Tracking complex defect transformations on the surface.

(a-d) STEM based tracking of defect transformations on graphene surface. (e-h) Corresponding single defect structures extracted by applying FCN, LoG and graph representations to data in (a-d)

Finally, we extended this framework to the analysis of the defect structures that represent a clustering of Si atoms in the form of a propeller (hereafter, molecular “rotors”). We selected three different frames (Fig. 3-7a-c) obtained from STEM atomic manipulation of the molecular “rotor”. The decoded results representing a single defect chemical structure for each individual case are shown in Fig. 3-7d-f. The difference between the first frame (Fig. 3-7a) and the second frame (Fig. 3-7b) is the rotation of the “rotor” structure by 60°. In the third frame (Fig. 3-7c), an additional Si atom was placed in the center of the “rotor” structure resulting in a new defect structure as shown in Fig. 3-7f.

We found in our analysis that in the 4-Si-rotor structure, the distance between Si atoms in the corners of a triangle increases by ~30% compared to 3-Si-rotor structure, effectively breaking the bonds between corresponding Si atoms. On the other hand, the average length of bonds between the central Si atom and the Si atoms in the corners was comparable to that in carbon lattice, ~ 150 pm. This suggests a scenario in which the Si atom in the center is covalently bonded to three other Si atoms in the corners but is displaced along the z-direction which leads to the effective shrinking of the observed bond lengths. We note that as the number of dopant atoms located next to each other within a single defect structure increases, the model trained on a single isolated dopant may return less precise results for the values of bond lengths and bond angles. This, however, can be easily improved by using the newly identified defects as training samples/labels in the updated model (see Fig.3-1b).

Having extracted coordinates of all the building blocks of the individual defects we can obtain relevant information on their behavior and evolution over time (e.g. spatial trajectories, interaction and ‘switching’ processes) for larger datasets including STEM ‘videos’ in a matter of seconds. As an example, we show in Fig. 3-8 the analysis of the spinning behavior of the molecular “rotor” over the span of 72 STEM frames. Here, we are particularly interested in i) detecting in the automated fashion rotational switching of the “rotor” and ii) analyzing its internal structure such as bond lengths and bond angles for each frame. For the former case, we calculated an angle parameter that describes a relative orientation of three Si atoms with respect to the defect’s center of the mass. By taking into account C_3 symmetry of the defect, the rotational switching such as the one shown in

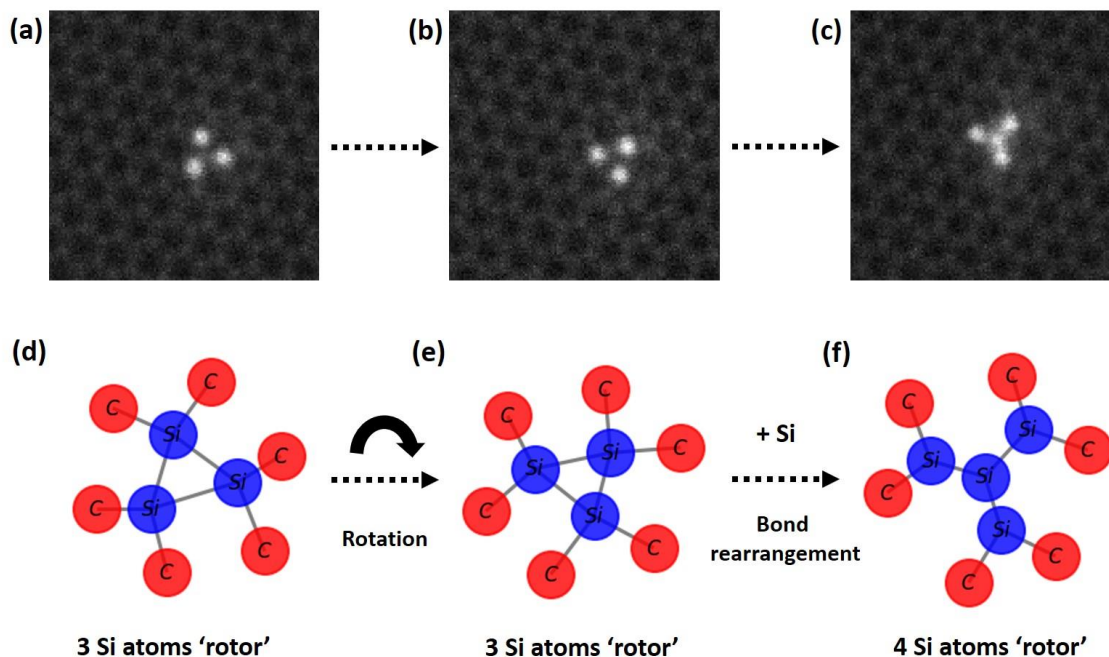


Figure 3-7. Tracking complex defect transformations on the surface during STEM atomic manipulations.

(a-c). STEM based manipulation of “rotor” structure. (d-f) Corresponding single defect structures extracted, in an automated fashion, after applying FCN, LoG and graph representations to data in (a-c)

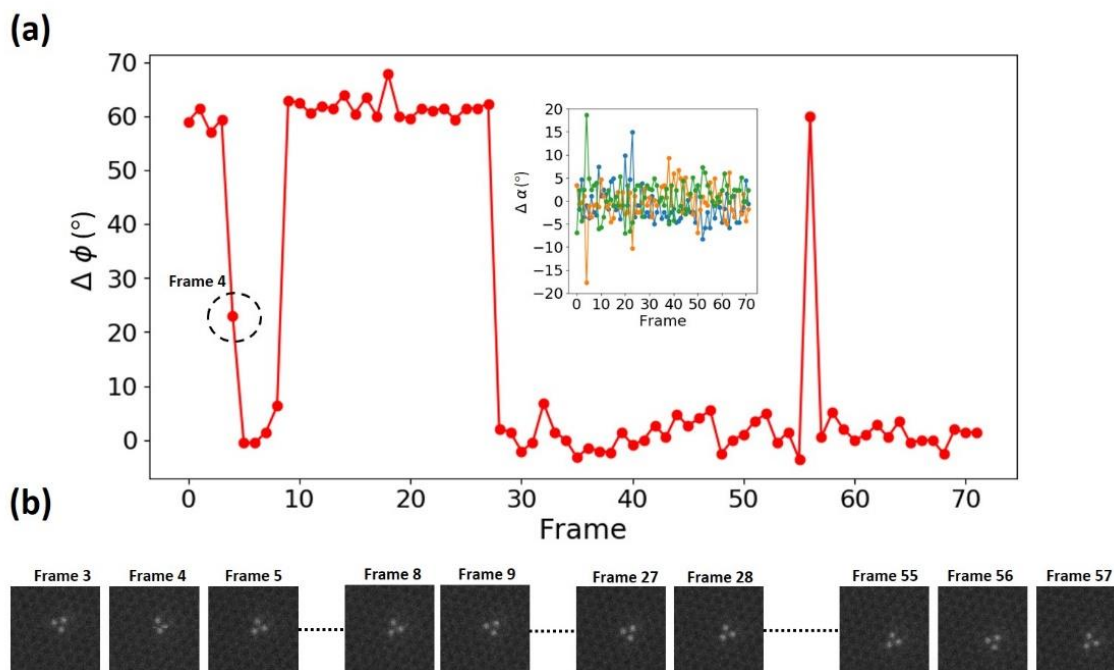


Figure 3-8. Analysis of “rotor spinning”.

(a) Behavior of an angle parameter ($\Delta\phi$) that describes (relative) orientation of the “rotor” in the image; the angle parameter was constructed using the information on coordinates and type of atomic species in the defect, which was extracted by applying FCN and LoG to STEM movie consisting of 72 individual frames (note that count starts from 0). The inset shows deviation from 60° for three angles in the defect’s Si-Si-Si triangle. (b) Individual frames of the STEM movie corresponding to the detected switching events.

Fig. 3-7a,b and 3-7d,e is described by the change in angle $\Delta\phi = 60^\circ$. Such switching events can be easily seen in Fig. 3-8a and occur at frames 3-(4)-(5), 8-9, 27-28, 55-56, and 56-57 (Fig. 3-8b). To analyze the internal structure of the defect we have plotted a deviation of angle in the Si-Si-Si triangle (inset in Fig. 3-8a) from the case of an equilateral triangle (see also Fig. 3-7 d,e)). This analysis suggests that while there are some fluctuations in the values of the Si-Si-Si bond angles, typically within $\pm 5^\circ$, the overall Si cluster structure remains largely intact. Interestingly, both plots show an anomalous behavior for frame 4. By visualizing this frame for the original experimental dataset, we found that it corresponds to an intermediate step of the rotational switching (i.e. the rotation occurred halfway through the image acquisition). We would like to emphasize that this type of analysis is not limited to dopants with higher Z and can be applied to other types of defects, and their motion and transformation, such as movement and rotation of vacancies and vacancy-dopant clusters, once their chemical structures and coordinates are extracted via the outlined deep learning based framework.

3.2.6 Conclusions

To summarize, we have developed a deep neural network based workflow for identifying positions of atoms in a lattice, type and positions of defect structures, and tracking complex defect transformations from un-processed STEM experimental data. To accommodate limited a priori knowledge about possible defect types (similar to background knowledge of general materials scientist without specific information), we have developed a “weakly supervised” approach, in which we start with just a few simple types (classes) of defects and then use information on local chemical structure to identify a rich variety of defects that were not in the initial training set. In addition, we extracted information relevant to complex defect transformations such as a reversible change between 3-fold and 4-fold coordinated Si dopants in a graphene lattice, formation of a Si dimer consisting of both 3-fold and 4-fold coordinated dopant atoms, and changes in the atomic structure and chemical bonding in the molecular “rotor”.

Based on the advances presented in this paper, we envision a development of a “self-driving” microscope for atomically resolved imaging in the not too distant future. In such a microscope, a neural network will be looking through the “eyes” of a microscope on a sample surface and returning information about the position and types of atomic species and defects in real time. Depending on the network’s output, the microscope will make necessary actions/adjustments (e.g. move the scan frame, or “switch” to the manipulation regime) similar to the way a self-driving car adjusts the steering wheel upon encountering an obstacle on the road.

3.2.7 Methods

Sample Preparation

Graphene was grown via atmospheric pressure chemical vapor deposition (APCVD) on a Cu foil. [230] The Cu foil was then spin coated with poly(methyl-methacrylate) (PMMA) as a mechanical stabilizer. The Cu foil was dissolved in a bath of ammonium persulfate-deionized (DI) water (0.05 g/ml). The graphene/PMMA film was transferred to hydrogen chloride (HCl) diluted in DI water bath to remove the ammonium persulfate, followed by a DI water rinse. The graphene/PMMA film was placed on a TEM grid and annealed on a hot plate at 150 °C for ~20 minutes to adhere the grid to the graphene. The PMMA was subsequently dissolved away in an acetone bath, followed by an isopropyl alcohol rinse. Finally the grid was annealed in an Ar-O₂ (450 sccm/45 sccm) environment at 500 °C for 1.5 hours to prevent e-beam induced hydrocarbon deposition in the STEM.[231, 232]

2D Mo_{1-x}W_xSe₂ monolayers were grown on SiO₂/Si substrates at 780 °C by a low pressure CVD method. [233] Poly(methyl methacrylate) (PMMA) was first spun onto the SiO₂/Si substrate with monolayer crystals at 3500 rpm for 60 s. The PMMA-coated substrate was then floated on 1M KOH solution, which etched the silica epi-layer, leaving the PMMA film with the monolayer crystals floating on the solution surface. The film was then transferred to deionized water for several times to remove residual KOH. The washed film was scooped onto a Cu TEM grid covered with lacey carbon. The PMMA was then

removed with acetone and the samples were then soaked in methanol for 12 h to achieve a clean surface with flakes.

STEM experiment

STEM imaging was performed using a Nion UltraSTEM U100 STEM operated at 60 kV. The images were acquired in high angle annular dark field (HAADF) imaging mode and were introduced to the DL network without any post processing.

Deep Learning

All the DL networks employed in our paper have an encoder-decoder type of structure and were implemented using Keras 2.0 (<https://keras.io>) python deep learning library, with the TensorFlow backend. We found that generally having separate FCN models tailored to find atoms in the lattice and specific defects allows more flexibility when it comes to the analysis of composite defects compared to a single FCN model that is trained to find probabilities for all the classes (lattice + defects) simultaneously. The encoder part of the network for finding lattice atoms/vacancies/dopants typically consisted of 3 convolutional layers with 8/8/2, 16/16/4, and 32/—/— kernels, each of the size 3×3 and stride 1, activated by a rectified linear unit (ReLU) function. Each convolutional layer was followed by a max-pooling layer with a 2×2 window and stride 2. The decoder (or “deconvolutional”) part of the network consisted of the kernels with the same size and stride plus the un-pooling layers. The last two convolutional layers have the number of kernels equal to the number of classes to be determined, with kernel size 3×3 (pre-last) and 1×1 (last) and stride 1 (note that there is only a linear activation $f(x) = x$ between these two convolutional layers). The feature maps from the final convolutional layer of the network were fed into a softmax classifier for pixel-wise classification, providing us with information on the probability of each pixel being an atom or specific type of defect. The Adam optimizer [234] was used with categorical cross-entropy as the loss function. To generate a large enough training set from theoretical image(s), we applied data augmentation procedure to the original synthetic image(s). Specifically, we used random horizontal/vertical shifts, rotations, zooming-in/out and shear transformations to generate a set containing ~2000

training images for each of the models. Each image was then corrupted with random noise and blurring in the ranges comparable to those typically observed in experiments. Plots of accuracy on the training and validation datasets over training epochs can be found in the Supplemental Material. The input images were resized to 128×128 using pixel area relation. The Laplacian of Gaussian based blob detection applied to the softmax output of the FCN was implemented with scikit-image image processing toolbox (<http://scikit-image.org/>), and the final graph representations were constructed using NetworkX (<http://networkx.github.io>).

3.3 Deep analytics of atomically-resolved images: manifest and latent features.

3.3.1 Publication Information

Original article was submitted for review to Nature Communications.

It was authored by Maxim Ziatdinov, Ondrej Dyck, Artem Maksov, Bethany M. Hudak, Andrew R. Lupini, Rama K. Vasudevan, Stephen Jesse, and Sergei V. Kalinin, and is available as preprint at:

Ziatdinov, Maxim, Ondrej Dyck, Artem Maksov, Bethany M. Hudak, Andrew R. Lupini, Jiaming Song, Paul C. Snijders, Rama K. Vasudevan, Stephen Jesse, and Sergei V. Kalinin. "Deep analytics of atomically-resolved images: manifest and latent features." *arXiv preprint arXiv:1801.05133* (2018). [235]

3.3.2 My contribution

While framework described in the previous section was able to identify and classify features in STEM images of hexagonal materials, it had failed to provide physically

meaningful results for other types of lattices and images. In this section, the work addresses generalization of deep learning methods to STEM and STM images of different materials. I proposed and implemented removing periodic constraints from training datasets by including randomly generated distributions of atomic positions with only constraint being distance from other atoms, as positions cannot overlap. I trained a deeper model on such a dataset which allowed to identification of atomic positions within STEM images of Al/Si interface and WS₂, as well as atomic positions in STM images of La_{5/8}Ca_{3/8}MnO₃, and molecular positions and rotations in STM images of sumanene molecules.

3.3.3 Abstract

Recent advances in scanning transmission electron and scanning tunneling microscopies have allowed researchers to measure materials structural and electronic properties, such as atomic displacements and charge density modulations, at an angstrom scale in real space. At the same time, the ability to quickly acquire large, high-resolution datasets have created a challenge for rapid physics-based analysis of images that typically contain from several hundreds to several thousand atomic units. Here we demonstrate a universal deep-learning based framework for locating and characterizing atomic species in the lattice, which can be applied to different types of atomically resolved measurements on different materials. Specifically, by inspecting and categorizing features in the output layer of a convolutional neural network, we were able to detect structural and electronic “anomalies” associated with a presence of point defects in tungsten disulfide monolayer, non-uniformity of the charge density distribution around specific lattice sites on the surface of strongly correlated oxides, and transition between different structural states of buckyball molecules. We further extended our method towards tracking, from one image frame to another, minute distortions in the geometric shape of individual Si dumbbells in 3-dimensional Si sample, which are associated with a motion of lattice defects and impurities. Due the applicability of our framework to both scanning tunneling microscopy and scanning transmission electron microscopy measurements, it can provide a fast and straightforward way towards

creating a unified database of defect-property relationships from experimental data for each material.

3.3.4 Introduction

Nano- and atomic scale imaging and spectroscopy have become the hallmark of modern times, allowing unprecedented access to materials structure and functionalities. [195-197] Electron beam techniques such as scanning transmission electron microscopy (STEM) now allow performing single atom imaging [236] and probing the local states of quasiparticles and their interactions in the solid-state systems. [237, 238] Meanwhile, scanning probe techniques, such as atomic force and scanning tunneling microscopy (STM), enabled a broad range of studies of materials surfaces, including imaging of magnetic [239] and ferroelectric [240] domains, single molecule recognition for biological applications [241], as well as atomically resolved imaging of physical and chemical processes on metallic and semiconducting surfaces. [52, 213, 242]

The continuous development in these techniques and emergence of *in-operando* imaging gives rise to progressively increasing volumes of data, as exemplified by the development of dynamic STEM, [203, 204] where one can observe, in real time, changes in the structure of a sample under the external perturbations, as well as e-beam atomic manipulation techniques. [205, 206, 211] These developments necessitate rapid image analysis, ideally immediately during acquisition.

This task becomes particularly relevant in the context of recently discovered controlled (manually) e-beam modification of materials, such as movement of single [243-245] and multiple [246] Si atoms over several lattice sites in graphene. Particularly, in order to achieve fast and automated, feedback-based atomic-level manipulation of materials structure, it is necessary to “teach” a microscope to localize and recognize various atomic objects (atoms, atomic columns, defects) from one STEM (or STM) frame to another in real time, as well as to interpret the newly created (during a manipulation stage) and previously unseen atomic features ‘on-the-fly’ in a way similar to how human operator

works. One of the most promising method for creating such an AI microscope is based on deep artificial neural networks that are capable of learning from the image in a way somewhat similar to humans. [46, 219]

Deep artificial neural networks have witnessed unprecedented success, progressing from recognizing pictures of cats and dogs on the Internet to defeating a human champion in one of the most sophisticated human games in the world in just about three years. [219] The deep learning (DL) is now being actively used in various engineering applications, such as object detection and recognition [247] and unsupervised text translation, [248] to name just a few. More recently, deep learning models have been utilized in the domain of theoretical physics for the analysis of data collected at the Large Hadron Collider [39] and detection of phase transitions in lattice models. [221, 222]

There are three critical aspects to consider before applying DL analytics to scientific imaging. The first aspect is a network's topology. For atomic imaging, for example, it is crucial to locate accurately all the atomic species in the image, hence, the model's output layer must provide a pixel-wise classification map of the same size as the input image. The second critical aspect is the creation of an appropriate training set. The currently available DL models pre-trained on objects from the macroscopic world cannot be used for analysis of atomic species and defects. When creating a new training set specifically for atomically-resolved imaging, one of the key issues is a trade-off between a model's ability to generalize to different types of atomic structures (e.g. lattices with different periodicities and symmetries) and its accuracy (e.g. how precise is a localization and identification of atomic species and defects). The question is: shall one use a different model for each lattice type or instead create a universal model that is able to analyze all types of lattices? Because the former approach will not work for complex materials for which several phases (and interface regions) are usually present in a single image, it is preferable to have one model capable of recognizing multiple types of lattices (and associated lattice disorder), provided that a trade-off between generalization and accuracy is acceptable for physics-based interpretation of the obtained results. Finally, the third aspect is whether a further processing of the DL model output can potentially yield more information. Here, one must be aware of several important differences with the applications of DL in a "real"

macroscopic world, particularly, with those DL models used for segmentation tasks, such as fully convolutional networks. [249] Indeed, while a fully convolutional neural network can be an optimal tool for many (static) image segmentation tasks, [249, 250] the interpretation of the results of such segmentation for the macroscopic world applications is rather straightforward and requires little to no further analysis (that is, simple visualization of the output is sufficient in the most cases). The situation, however, is entirely different, even for static images, in the world of atoms and molecules, where small changes in shapes and relative positions might imply emergence of certain physical phenomena (e.g. formation of charge and/or magnetic ordering). In this case, the outputs of the DL model can be refined by using certain theory constraints, for example, by applying a theory-informed Markov model to the output of the last convolutional layer. In addition, exploratory data analysis and anomaly detection methods can be applied to the DL model output to uncover crucial information about atomic defects in the system and their impact on lattice properties.

Here we have developed a deep learning based approach that allows finding atomic species, as well as lattice defects, in the atomically-resolved images from different materials without specifying the type of materials lattice. It works with both STEM and STM images and is not sensitive to noise and other common types of experimental image distortion. We further extended our framework in such a way that it allows to perform exploratory data analysis, as well as anomaly detection and unsupervised and semi-supervised clustering, on the output of the DL model to uncover information linked to materials atomic structure and electronic properties.

3.3.5 Results and Discussion

The schematics of our approach for analysis of experimental data using a single DL model for different materials and types of atomic lattice is shown in Fig. 3-9. It starts with collecting experimental data (e.g. from STM or STEM observations) and feeding it into the deep convolutional neural network. The network has an encoder-decoder type of

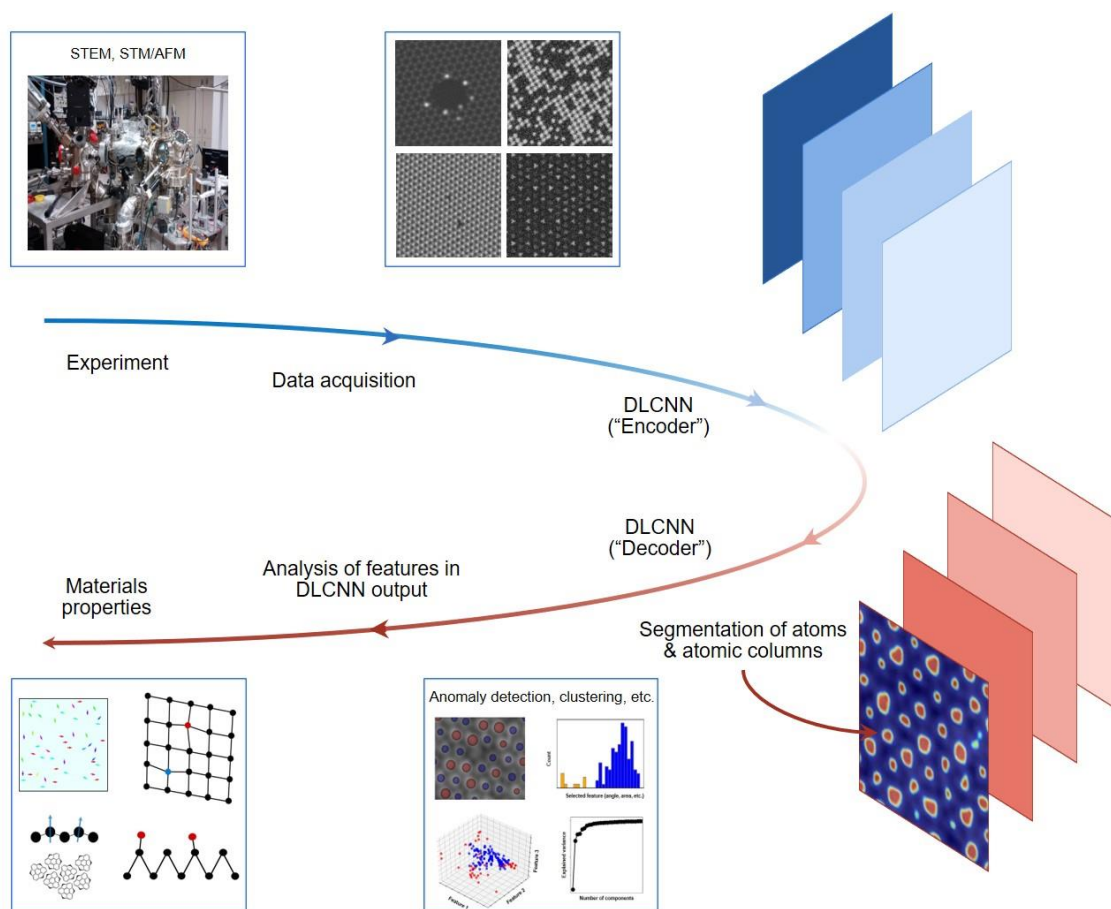


Figure 3-9. Deep learning based identification of atoms, atomic columns and defects.

The experimental data is passed to a deep fully convolutional neural network, which has an encoder-decoder type of structure (see Methods for details on network's architecture). The final, *softmax* layer of the network typically shows a set of well-defined features which descriptors (intensity, area, circularity, etc.) are linked to physical processes defining shapes of atomic columns and/or certain experimental anomalies. By performing exploratory data analysis as well as applying appropriate clustering and/or anomaly detection algorithms to the *softmax* layer (model's output) one can learn important information about the atomic structure under consideration. See text for more details.

architecture [250] such that an input image and the model output have the same dimensions (a crucial requirement for accurate localization of atomic species), and is trained using simulated images of atoms. As the model does not contain fully-connected (“dense”) layers, it is referred to as a fully convolutional neural network (FCN). [249] We removed hard periodic constraints during simulation of training images, which allowed us to train a network that can find atoms in different types of lattices (see Supplemental Material). The typical output of FCN for atomically-resolved images represents a set of well-defined circular features on uniform background, where each feature corresponds to individual atom and/or atomic column.

The key here is that variations in geometrical shapes of these features (e.g., area and circularity) are not random but are linked to physical properties of the system (shape of atomic columns) as well as possible experimental anomalies such as abrupt changes in the tip apex shape in STM. Hence, a simple exploratory data analysis of the FCN output layer as well as more advanced data analytics techniques such as unsupervised classification/clustering and anomaly detection can yield a wealth of information on material's properties. Furthermore, in addition to performing a pixel based analysis of the FCN output, it is possible to construct physical descriptors of the system using parameters of the extracted atomic contours (area, circularity, orientation/angle, etc.) as well as other relevant parameters such as the nearest and next nearest neighbor distances between the centers of extracted contours and column intensity (summed over small box area around the detected center), and to use these descriptors as an input into the above mentioned data analytical models.

We start with illustrating how a deep learning model trained on the *simulated* data without hard periodic constraints can find atoms in different periodic lattices of real experimental images (Fig. 3-10). We first use our model to identify atomic positions in a STEM image obtained at an Al-Si interface region (Fig. 3-10a). Here the model was able to locate positions of atomic columns in both the Al lattice (bottom of the image) and Si lattice (top of the image), as we show in Fig. 3-10b. By contrast, if we use a model trained on a honeycomb, graphene-like lattice used in our earlier study, [194] it will find atomic positions in the Si lattice, but will fail to identify accurately atoms in Al lattice as it will try

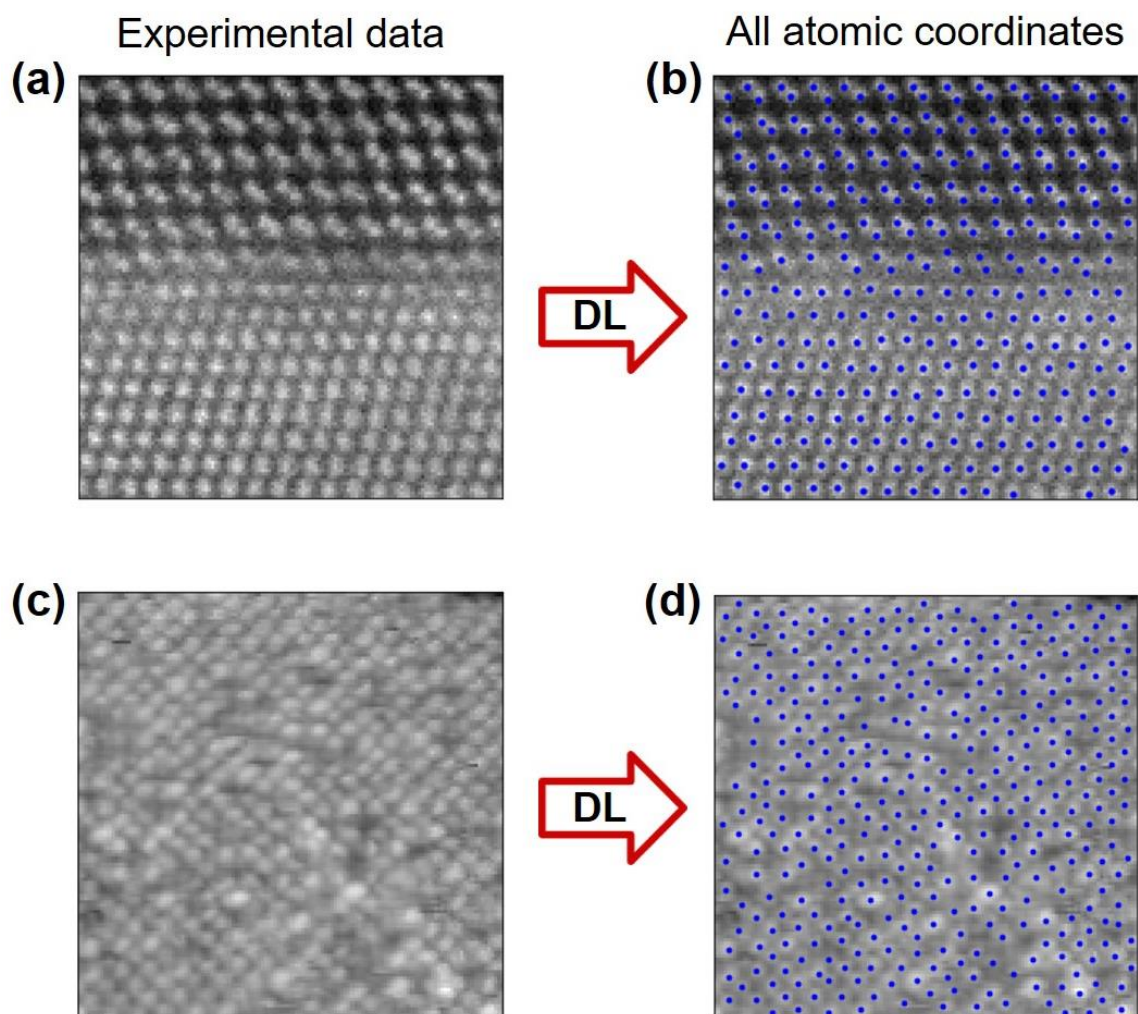


Figure 3-10. Using a single DL model to identify positions of atoms/atomic columns in the images of different materials obtained from two different experimental techniques.

(a-b) STEM image of Al/Si interface (a) and identified positions of all atomic columns (b).
 (c-d) STM image of $\text{La}_{5/8}\text{Ca}_{3/8}\text{MnO}_3$ (LCMO) film (c) and identified positions of all surface atoms (d).

to “place” additional atoms in-between Al atoms to turn it into a honeycomb-like structure (see Supplementary Material).

To ensure a universality of our model for more different types of atomic lattices, as well as a different experimental technique, we applied it to an STM image of $\text{La}_{5/8}\text{Ca}_{3/8}\text{MnO}_3$ (LCMO) film (other STEM and STM examples will be shown later in the text as well), which has a relatively high degree of lattice disorder (Fig. 3-10c).

Unlike the STEM image, the STM image is associated with electronic density of states around the Fermi level. While this can provide invaluable information about electronic behavior at the atomic scale, it also makes it more difficult (compared to STEM) for most image analysis techniques to locate atom-like or molecular features in the image and to categorize different structural and/or electronic states. In fact, automatic analysis of STM images is notoriously difficult compared to STEM, reflecting the delocalized nature of STM signal as compared to delta-function like original image for STEM.

In addition, on a somewhat deeper technical level, the STM images are usually characterized by large amounts of scars (strokes) due to interaction of an STM tip with a surface, which further complicates the task of finding atoms/molecules. Despite all these obstacles, our model was able to accurately identify positions of surface atoms in the unprocessed image. We would like to emphasize that neither of two atomic lattice structures was explicitly included in our training set. This means that our model trained on simulated STEM data was able to learn the concept of what an atom/atomic column is and use it to find positions of atoms in various lattice structures in real experimental images, without being biased to one lattice type (and its periodicity). Furthermore, it was able to apply the learned concept of “atoms” to the analysis of an STM image, even though we did not train it specifically on STM data. This is important because to date the analytical methods developed for STEM have rarely been applied to STM, partially because the STM data is much “noisier” (due to tip-surface interaction), harder to match with simulated data (the image is a convolution of tip and sample density of states and is also affected strongly by out-of-plane surface deformations) and typically has lower pixel resolution. However, our results clearly show that the deep convolutional neural networks, which are not sensitive

to pixel-level details and instead are able to capture “deeper” and more abstract concepts of the atomic lattices, can easily overcome these constraints.

We next illustrate how a simple exploratory analysis of the DL model output can provide important information regarding types of atomic species in the image, using a STEM image of Mo-doped WS₂ (Fig. 3-11a) as an example. Note that the S atoms are almost non-differentiable from background noise in the STEM experiments on WS₂ and hence the lattice appears as having a trigonal structure instead of a honeycomb-like structure.

While finding the likely locations of S atomic columns using FCN is possible in principle, we will not perform such an analysis in this study. The output of FCN (*softmax* layer) for this image is shown in Fig. 3-11b. It essentially represents a probability map of each pixel belonging to an atom/atomic column. The key observation here is that variation in shapes of circular-like features in the *softmax* output is not random, but is linked to actual changes in the shape/intensity of atomic projections in the STEM image. As such, features representing dopants/contaminations with lower (higher) Z are characterized by a smaller (larger) area (Fig. 3-11b). Note that the model was *not* specifically trained to return features of different sizes for atoms with different intensities. The analysis of other relevant geometric parameters such as circularity may provide additional information about lattice distortions, “switching” during the scan (particularly in STM images), or even about some more subtle effects such as a character of rapid atom oscillations in the shallow potential well of a 2D system.

We now illustrate how the search for “anomalies” in the output of DL model can be in principle automated. Once we establish a probability of each pixel belonging to an atom or to a background (that is, being relevant or irrelevant), we can extract a contour of each identified feature associated with atoms and/or atomic columns and characterize it based on its geometric descriptors such as area, circularity, orientation/angle, etc. For STEM image of the doped WS₂, a plot of area and circularity descriptors for all the extracted contours as well as associated histograms are shown in Fig. 3-11c.

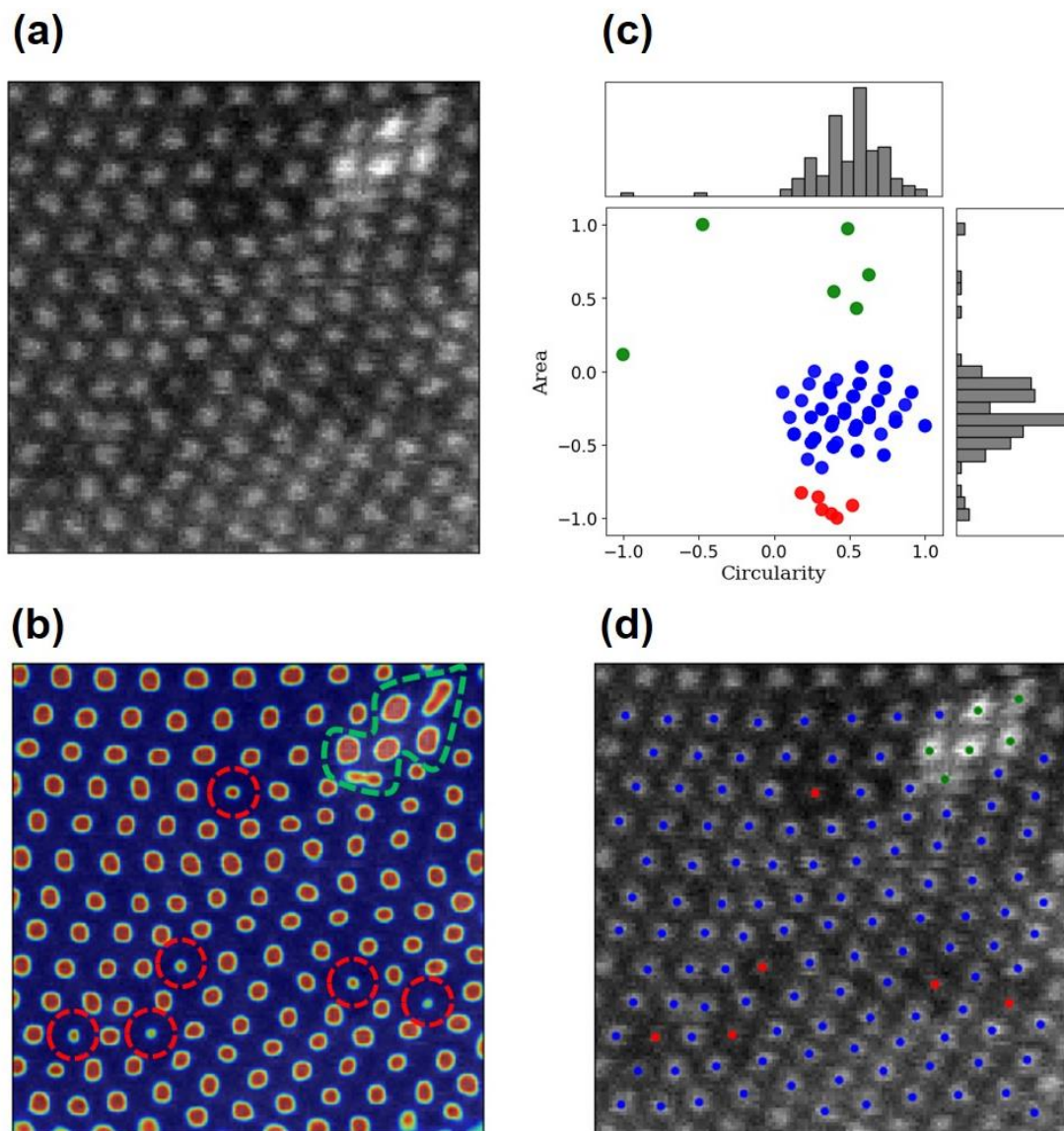


Figure 3-11. Identification of anomalies in softmax layer of DL model.

(a) Experimental STEM image of Mo-doped WS_2 . (b) *Softmax* layer of DL model for image in (a). “Contaminations”/adatoms and Mo dopant are shown by green and red dotted lines, respectively. (c) Plot of descriptors (normalized) associated with area and circularity of individual features in the *softmax* layer with corresponding histograms for each descriptor. The anomaly detection method (one class support vector machine) followed by semi-manual refinement (1 data point removed from anomaly “cluster”) allowed to separate the extracted descriptors into clusters associated with “contaminations”/adatoms (green) and Mo dopants (red). (d) Atomic positions overlaid on the experimental image (coloring scheme the same as in (c)).

We can then classify features from DL output based on the constructed descriptors using clustering techniques and/or anomaly detection methods. For data in Fig. 3-11a, the anomalies extracted via one class support vector machine method are plotted along with the regular lattice in Fig. 3-11d. The red circles identify what is likely to be Mo dopants (lower Z atomic number compared to W), whereas the green circles correspond to “contaminations”/adatoms that frequently form on WS_2 surface under e-beam irradiation. In Fig. 3-12 we illustrated several more representative examples of successful application of this approach to STEM experimental data.

Specifically, in Fig. 3-12a-b we show another STEM images of Mo-doped WS_2 for which, in addition to finding all dopants and atomic-size contaminations, our model was able to describe accurately a non-trivial shape of edges (*e.g.*, presence of both regular lattice atoms and dopants at the edge in Fig. 3-12a,d or considerable irregularities in the edge shape in Fig. 3-12b,e). In Fig. 3-12c,f we were able to identify accurately positions of Bi dopants in a Si lattice, as well as locations of Si atomic species themselves.

In STEM experiments on 3-dimensional materials, the variations in the shapes of atomic columns may originate from certain atomic (re)arrangements along the columns that are too small to be resolved as separated features but are large enough to distort the shape and orientation (if the initial shape is not circular) of the atomic columns. We will now illustrate how our DL-based model can be used to extract all the relevant information about geometrical parameters of such distortions. As an example, we use STEM images of a Si lattice with Bi dopants (Fig. 3-13). Here, in addition to non-uniformity introduced to the Si lattice by Bi dopants, we expect additional forms of the lattice “disorder” caused by atom motion and, potentially, lattice reconstructions under the e-beam impact.

By tuning the threshold intensity value used for the extraction of atomic contours from the DL model output we can make a transition from analysis of individual Si columns to analysis of individual Si dumbbells (see Supplementary Material). Specifically, we approximate each extracted Si dumbbell contour with an ellipse and analyze its orientation (angle) with respect to the chosen frame of reference (Fig. 3-13a). As our data is now represented by the collection of well-defined contours on an “empty” background, we can easily track changes in the geometrical descriptors of a specified contour (or all contours

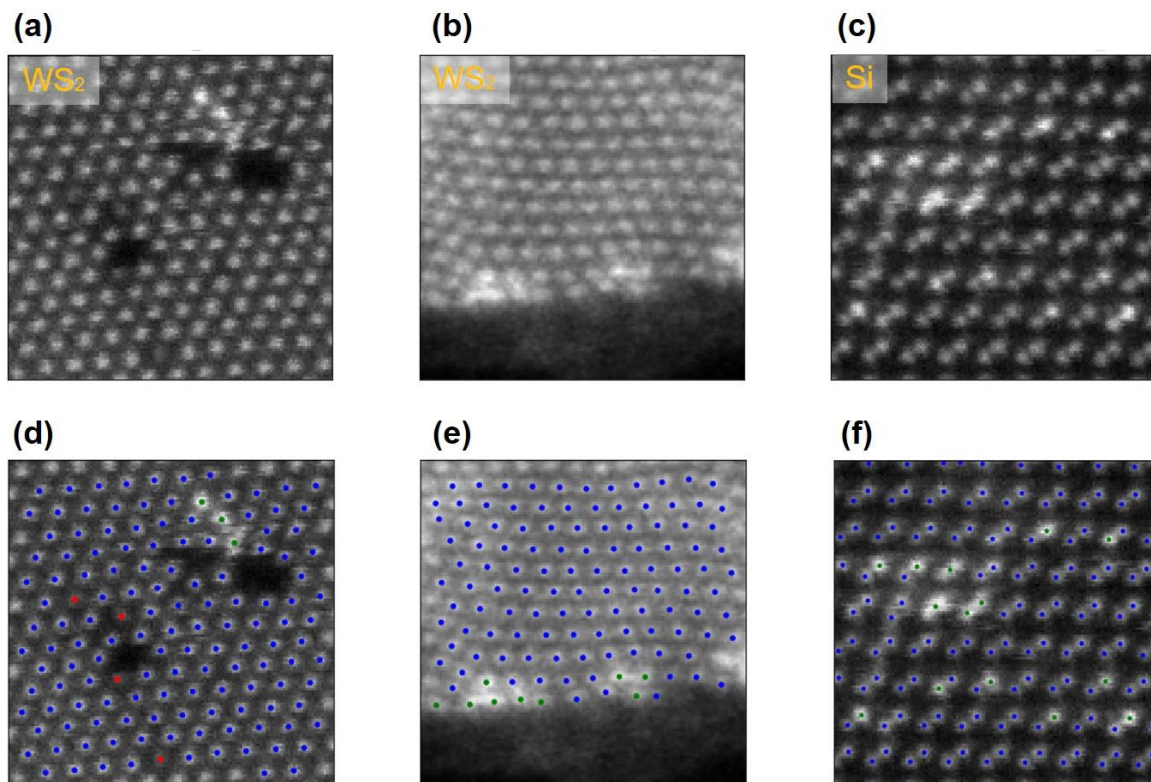


Figure 3-12. Identification of lattice atoms and lattice disorder from STEM images using a single DL model.

(a-c) Experimental STEM images of WS₂ with small holes in the lattice (a) and step edge (b), and Si with Bi dopants (c). (d-f) Corresponding atomic positions (colormap meaning the same as in Fig. 3).

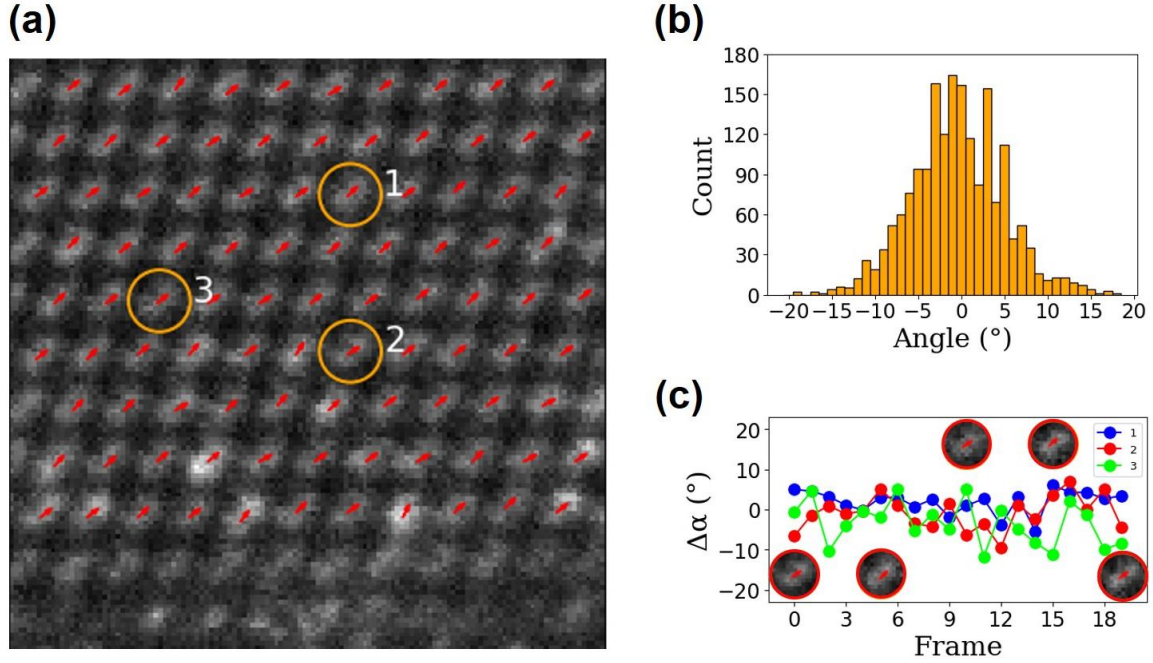


Figure 3-13. Analysis of atomic columns distortion in 3-dimensional Si sample by tracking orientation of individual Si dumbbells.

(a) First (out of twenty) frame of STEM movie on Bi-doped Si sample. Orientation of Si dumbbells extracted from the output of DL model is shown by red arrows. (b) Distribution of Si dumbbell angles (with respect to the [001] direction) for twenty movie frames. (c) Tracking changes in orientation of individual Si dumbbells (orange circles in (a)) across twenty movie frames (count starts from zero). For several selected frames, the orientation of dumbbell #2 is shown by the red arrow overlaid on top of experimental data.

at once) through multiple STEM frames (STEM “movie”) as we demonstrated in Fig. 3-13c To track each “atomic contour” during the movie analysis (or acquisition) we first store the coordinates of the centers of all contours from the first frame and then locate the same contours in the subsequent frames by finding the closest contour for each pair of coordinates. This approach works if an overall shift between the first and the last frame due to sample drift is smaller than a half-distance between two neighboring dumbbells. For larger sample drifts one must update the stored coordinates after successfully locating the same contour(s) in each frame (note that the same idea can be used “in reverse” for drift correction between consecutive frames). This ability to obtain detailed information on the changes in orientation of individual structural units in an automated fashion can provide important clues about types of disorder present in the material's volume. It may be particularly useful for analysis of long-range correlations in complex 3D systems with various types of charge and/or magnetic (dis-)order, such as complex cobaltates. [251] Furthermore, as we show below, it can be easily extended towards analysis of non-uniform charge density distribution in STM images of strongly correlated systems.

We proceed to applying the same DL model to analysis of STM images. Specifically, we apply our DL model to the STM data on an LCMO film. This material is characterized by a complex phase diagram with competing structural, chemical and electronic/magnetic orders which to date remains inadequately understood. [252] In some cases, such a competition can lead to formation of full or partial charge order characterized by a non-trivial shape and distribution of charge density "blobs". [253] Even without formation of charge order, analysis of charge density distribution around different lattice positions can provide important clues on local chemical environment and electronic environment. For example, since the STM images are imaging O atoms (these have the highest density of states near the Fermi level [254]), one can expect that depending on the local cation environment, the imaged atom may display elongated features caused by altered local electronic density, which would be expected when comparing La^{3+} to Ca^{2+} in this system. The raw experimental STM image is shown in Fig. 3-14a. In Fig. 3-14b we plotted the anomalies in the distribution of charge density at atomic lattice sites, detected by a one class support vector machine and associated with variation in circularity and area of

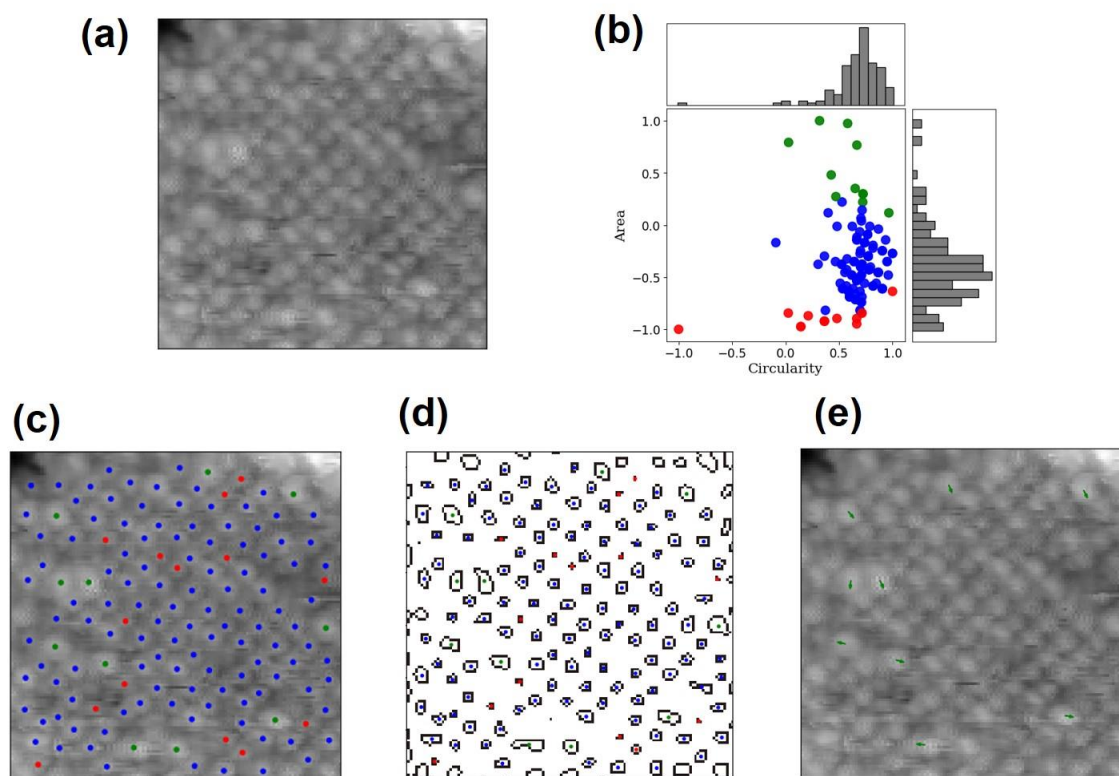


Figure 3-14. Identification of anomalies in *softmax* layer of DL model for STM data.

(a) Experimental STM image of LCMO surface. (b) Plot of geometrical descriptors (normalized) associated with area and circularity of individual features in the *softmax* layer with a corresponding histogram for each descriptor. (c) Atomic positions overlaid on the experimental image. The color labels are based on data analysis shown in (b). (d) Visualization of “atomic contours”, which in this case provides information about non-uniform distribution of charge on the surface. (e) Green arrows show schematically the distortion of electronic charge density for atoms labeled green in (c) and (d)

individual features. The detected anomalies are plotted together with the experimental image and “atomic contours” in Fig 3-14c and 3-14d, respectively. Interestingly, by more closely analyzing the larger features (marked green) we found that the charge density appears to be “leaking” in one direction (that is, charge distribution is not identical at the ends of blob’s long axis) which can be associated with non-homogeneous local chemical environment.

We finally show that it is possible to use our model for analyzing molecular-resolved STM data. Specifically, in Fig. 3-15 we demonstrate a DL based analysis of STM image of sumanene molecules (also called buckybowls) on a gold substrate. The molecules can have different orientation with respect to the substrate (azimuthal rotations) and, under certain experimental conditions, can switch between their bowl-up and bowl-down structural conformations. [255] Interestingly, our model, which was trained to find atoms, was able to locate local maxima (hereafter, “blobs”) in charge density distribution inside molecules (Fig. 3-15b). The latter is a product of hybridization with a substrate and can be also affected by electronic and structural state of an STM tip apex. The ability of our model to see these characteristic “blobs” can be viewed as if model treats each molecule as a cluster of “atoms” (not to be confused with actual, much larger number atoms inside a molecule). For the unperturbed molecular structure, positions of such “blobs” within each molecule allows us to determine an orientation (azimuthal rotation) of the molecule with respect to the substrate (blue markers in Fig. 3-15b-e). On the other hand, the arrangement of “blobs” within each molecule is altered for a molecule that undergoes a transition between different structural (and/or rotational) states. We found that such a transition can be detected by our model (red markers in Fig. 3-15b and 3-15c). This ability to detect a transition between different molecular states via a deep learning framework can prove to be an invaluable tool for single molecule manipulations in future fully autonomous microscopes.

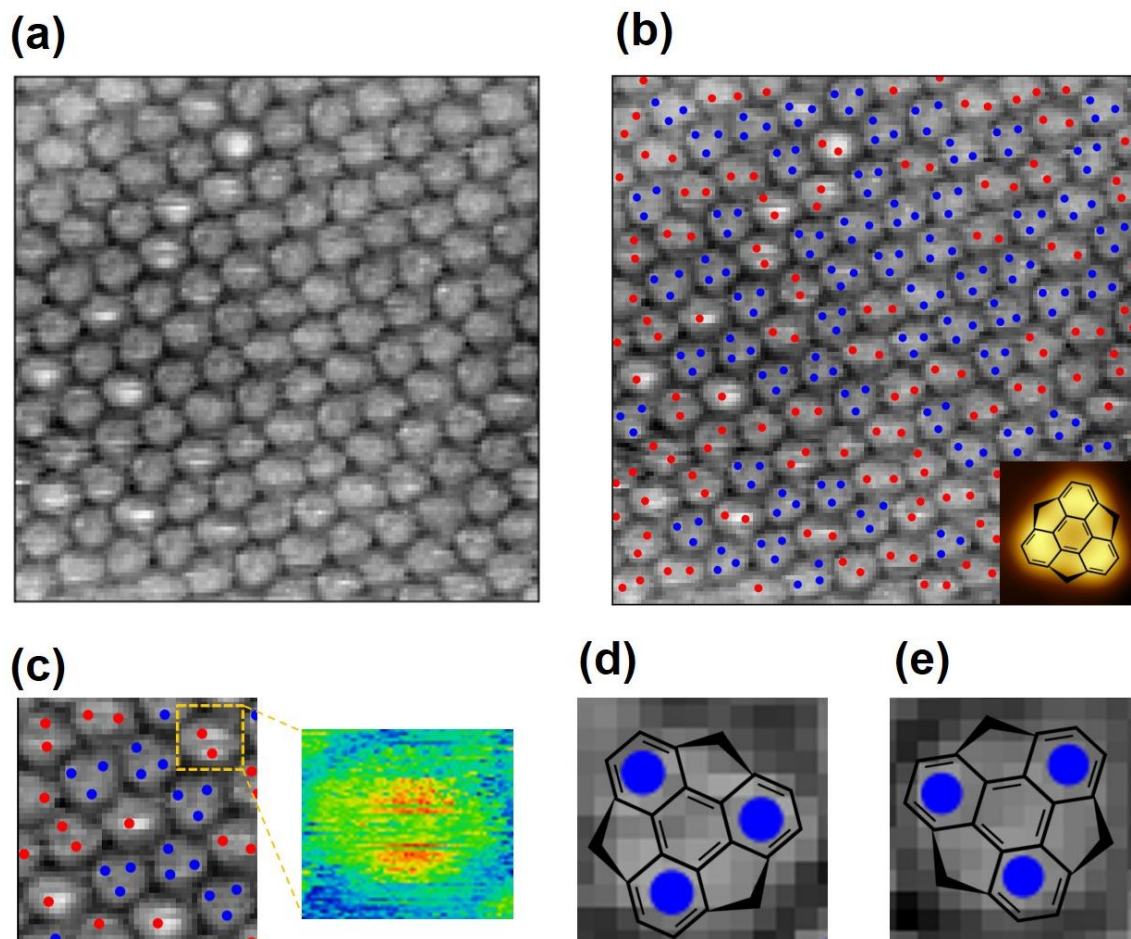


Figure 3-15. Identifying molecular rotations and detecting a switching between different molecular conformations via DL.

(a) Experimental STM images of sumanene monolayer on gold surface. (b) Decoded molecular configurations. Molecules that do (do not) undergo a switching between different conformations are marked red (blue). The positions of blue markers correspond to a molecular orientation on the substrate as can be clearly seen from a zoomed area in (c). Zoomed-in illustrative example of a molecule undergoing a switching event during the scan is shown for one of molecules in (c). (d-e) Molecular chemical structure overlaid on top of the decoded structure for two different orientations. The overlay was done according to the distribution of charge density maxima determined from DFT STM simulations of molecules on the gold substrate (see inset in (b)). Note that a match is not perfect in part due to the model's uncertainty, but also because of interaction between molecules and STM tip as well as potential admixture of different rotational states and the associated dynamic averaging in STM experiment.

3.3.6 Conclusions

In summary, we have developed a universal deep learning based analytical framework for fast and accurate analysis of atomic lattice structures across a wide range of materials. Within this framework, we first utilize a deep convolutional neural network for locating and extracting contours of individual atoms and atomic columns from the experimental images. We then construct and analyze geometrical descriptors of the extracted contours such as area, circularity, and orientation (angle) via exploratory data analysis and unsupervised anomaly detection/clustering. The variations in the contour's shape are linked to materials structural and electronic properties, such as presence and motion of dopants and contaminations on the surface and in the bulk, as well as a non-uniform spatial distribution of electronic charge density in the vicinity of certain surface lattice atoms. We further extended our approach towards analyzing rotations and detecting conformational changes of molecules and tracking specific geometrical descriptors associated with the selected atomic columns in STEM “movies” of 3D material. We foresee that due to its ability to quickly analyze raw data obtained from different types of atomically-resolved experimental measurements (structural data from STEM, electronic properties from STM), the created framework opens a pathway towards creation of comprehensive libraries of defect-property relationships for each technologically relevant nanoscale material.

3.3.7 Methods

STEM experiment

STEM imaging was performed using a Nion UltraSTEM U100 STEM operated at 60 kV for all images except that shown in Figure 3-10(a), which was acquired at 100kV, and Bi-doped Si. A Nion UltraSTEM U200 aberration-corrected scanning transmission electron microscope, operating at 200 kV was used for images of Bi-doped Si. The images were acquired in high angle annular dark field (HAADF) imaging mode and were introduced to the DL network without any post processing.

STM experiment

STM imaging of LCMO films was performed *in situ* at room temperature with an Omicron VT system at an operating pressure $< 3 \times 10^{-10}$ Torr using mechanically cut Pt/Ir tips in constant current mode. The images were taken with the tip negatively biased in respect to a grounded sample. STM imaging of sumanene molecules was performed at room temperature in a constant-current mode using the Japan Electron Optics Laboratory (JEOL) JSPM-4500S system at an operating pressure $\sim 10^{-11}$ Torr. The STM tips were prepared by electrochemical etching of tungsten wire and were further cleaned by Ar^+ ion sputtering (ion energy = 1.0–3.5 keV).

Deep Learning

The DL networks were implemented using Keras 2.0 (<https://keras.io>) Python deep learning library, with the TensorFlow backend. The convolutional neural network had an encoder-decoder type of structure. The first block of encoder consisted of two back to back convolutional layers (20 kernels of size 3×3 and stride of 1 in each layer) followed by max-pooling layer (size 2×2 , stride 2), while the second block had three back to back convolutional layers (number of kernels in each layer 40, size 3×3 , stride 1) followed by max-pooling layer (size 2×2 , stride 2). The decoder part had the same structure but in the reversed order and was followed by a dropout layer (rate set to 0.25). The last two convolutional layers had the number of kernels equal to the number of classes to be determined (size 3×3 (pre-last layer) and 1×1 (last layer) and stride 1). The feature maps from the final convolutional layer of the network were fed into a *softmax* classifier for pixel-wise classification, providing us with information on the probability of each pixel belonging to an atom. The Adam optimizer [256] was used with categorical cross-entropy as the loss function. The initial simulated STEM images produced by a Multislice algorithm [257] were augmented to generate a large enough training set of about 3000 images. The augmentation was performed by applying random horizontal/vertical shifts, rotations, zooming-in/out and shear transformations. Each image was also corrupted with Poisson noise and blurring in the ranges comparable to those typically observed in

experiments. The accuracy of our model on the test set was about 97 %. The extraction of contours and associated geometrical descriptors was performed in part using the Open Source Computer Vision Library (<https://opencv.org>).

3.4. Learning surface molecular structures via machine vision.

3.4.1. Publication Information

A version of this article was originally published by Maxim Ziatdinov, Artem Maksov, and Sergei V Kalinin in:

Ziatdinov, Maxim, Artem Maksov, and Sergei V. Kalinin. "Learning surface molecular structures via machine vision." *npj Computational Materials* 3, no. 1 (2017): 31.

The article content is reprinted as published with figure numbers modified and supplementary material added at the end of the chapter as an appendix.

3.4.2. My contribution

I proposed and implemented a machine vision framework for identification of both structural and rotational states of sumanene molecules in a self-assembled monolayer imaged by STM. Sumanene molecules can occupy 2 structural states (bowl-up or bowl-down), as well as multiple rotational states. Simulations of the self-assembly have previously identified that optimal configuration is one bowl-down molecule surrounded by 6 bowl-up molecules, however, visual inspection of experimental data have indicated that this configuration does not always form. Additionally, distinguishing between bowl-up and bowl-down structural states was not accurate due to significant overlap in histogram of intensities of 2 states. In addition to experimental images, 2 simulated STM images constructed from DFT simulations were available, representing each structural state.

I started by constructing additional simulated images for rotational states based on initial configurations. From previous simulations and visual inspection, it was established that 4 distinct rotational states with 30-degree increments were present within the data, with additional rotations being degenerate due to the individual molecules shape which can be described roughly as a equilateral triangle. Additionally, theory has indicated that neighboring molecules of the same rotation are highly unlikely due to repulsion.

I proposed modeling the self-assembly structure as a Markov Random Field (MRF), where each molecule was a node in the graph and was connected to its 6 nearest neighbors. Intensity of the molecule would be proportional to the probability of belonging to bowl-down class (and inversely proportional to bowl-up class probability), which was assumed to be a sigmoid function. The pair-wise probabilities were assumed from the simulation with a disorder parameter p introduced: bowl-down/bowl down pair was assigned probability $(0 + p)\%$, while bowl-down/bowl-up a probability of $(100 - p)\%$. The possible neighbors of bowl-up molecules were assigned 50% probability each. To identify individual probabilities of belonging to a rotational class, I applied a deep neural network trained on a training set generated from initial simulations through data augmentation. For pairwise probabilities, I used equal probabilities for each rotational class except the same one, which was set at 1%.

Having defined the model, I used Markov Chain Monte Carlo (MCMC) sampling to create synthetic images. I used rotational averaging, Gaussian smoothing, and Poisson noise to simulate image altering effects observed in the experiment.

I created an averaged image from all available rotations and used correlation surface with thresholding and connected components to identify positions of all the molecules in the image and extract 25 by 25 pixels images centered around them. Using those images, I calculated sum of the pixel intensity values as input to structural state probability calculations and applied neural network to get rotational class probabilities. Structural state determination was improved from >10% error for mean intensity threshold to <1% for MRF model. Rotational state error for neural network only was found to be >3% and was improved to ~2% with application of MRF.

I applied the framework to experimental data and achieved full information extraction. I computed pair density functions for the molecular states, and further calculated Moran's I correlation between structural and rotational states, which hinted at the reaction path of the bowl-inversion process.

3.4.3. Abstract

Recent advances in high resolution scanning transmission electron and scanning probe microscopies have allowed researchers to perform measurements of materials structural parameters and functional properties in real space with a picometre precision. In many technologically relevant atomic and/or molecular systems, however, the information of interest is distributed spatially in a non-uniform manner and may have a complex multi-dimensional nature. One of the critical issues, therefore, lies in being able to accurately identify ('read out') all the individual building blocks in different atomic/molecular architectures, as well as more complex patterns that these blocks may form, on a scale of hundreds and thousands of individual atomic/molecular units. Here we employ machine vision to read and recognize complex molecular assemblies on surfaces. Specifically, we combine Markov random field model and convolutional neural networks to classify structural and rotational states of all individual building blocks in molecular assembly on the metallic surface visualized in high-resolution scanning tunneling microscopy measurements. We show how the obtained full decoding of the system allows us to directly construct a pair density function—a centerpiece in analysis of disorder-property relationship paradigm—as well as to analyze spatial correlations between multiple order parameters at the nanoscale, and elucidate reaction pathway involving molecular conformation changes. The method represents a significant shift in our way of analyzing atomic and/or molecular resolved microscopic images and can be applied to variety of other microscopic measurements of structural, electronic, and magnetic orders in different condensed matter systems.

3.4.4. Introduction

The symmetry properties of crystalline and molecular systems associated with a long-range periodicity of their assumingly ideal ‘lattices’ serve as a cornerstone for deriving electronic, magnetic, and optical functionalities of technologically relevant materials. Experimentally, these properties are usually accessed via scattering techniques that provide information about spatially averaged (over the probing volume) site occupancies. The knowledge of average structural parameters underpins classical physical descriptions based on concepts of order parameters, average compositions, and symmetry averaged thermal, phonon, and disorder properties. At the same time, there is growing realization that the effects of local structure exemplified by disorder can often lead to novel functionalities absent in averaged models.[258-262] For example, one interesting scenario of disorder occurs when there is a distinction between local symmetry associated with individual building blocks and global symmetry imposed by underlying lattice.[258, 259] A resultant interplay between local and average symmetry opens new pathways to understand and optimize optical, magnetic, electronic, and thermal properties of certain disordered systems.[259, 262]

Exploring and controlling different types of disorder both in periodic and non-periodic structures is therefore crucial for the applications and basic science alike. In the last decade, the progress in scanning probe and transmission electron microscopies (SPM and STEM) have allowed a real-space unit cell scale mapping of electronic and structural orders in materials, making them the perfect tools for analyzing such distorted systems.[263] For example, the subfields of SPM such as non-contact atomic force microscopy and scanning tunneling microscopy are known to provide an unprecedented, angstrom-resolved visual insight into a nature of chemical bonds[212] and spatial behavior of electronic density of states on a surface[213], respectively. Such capabilities result in ever-growing stream of the vast amounts of high quality (resolution) experimental data that requires adequate analytical methods for extracting from it a relevant physical and chemical information.[80] Concurrent with advances in real-space experimental microscopic measurements, contemporary theoretical, ab-initio modelling allows detailed study of atomic/molecular structures, their electronic, magnetic, and optical properties.[264-266] However, many

interesting functionalities in disordered molecular and/or atomic systems are defined on length scales at which the number of possible molecular or atomic configurations (and hence, computational cost) grow exponentially. Similarly, effects of local symmetry breaking and disorder are often manifested in minute ($\sim$ pm level) distortions of the bonding geometry or effective molecular shapes.[114, 267-269] This suggests a necessity for pathway to integrate an experiment and certain elements of theory that would allow an automated and highly efficient inspection and interpretation of experimental image consisting of a large number of individual atomic and/or molecular units ($\sim 10^2$ - 10^3) in a fashion of full information extraction, linking both minute deviations in local structure and large-scale assembly properties in statistically significant manner.

Here we use an approach based on a synergy of Markov random field, convolutional neural network, and ab-initio simulations for performing a full decoding of various orders associated with symmetries of individual building blocks (molecules) on the underlying lattice (substrate). We apply this method to explore molecular interactions in 2D molecular film of bowl-shaped sumanene molecule on gold substrate, where an individual molecule in each lattice point can reside in multiple (structural and rotational) configurations. The obtained full decoding at the nanoscale level allows us to directly construct both relevant pair density functions – a centerpiece in analysis of disorder-property relationship paradigm, [258] as well as more complex structural descriptors. This in turn allows us to explore how individual blocks may form certain short-range orders, as well as to analyze potential (spatial) correlations between multiple order parameters, and to use the obtained information for constructing a reaction path for molecule conformational changes in the self-assembly.

3.4.5 Results

Model experimental system.

As a model system, we chose a self-assembly of bowl-shaped π -conjugated sumanene molecules [270] (hereafter, buckybowls) on a gold (111) surface. Unlike most of the planar

molecules, the buckybowls are characterized by an additional structural degree of freedom associated with their bowl-up (**U**) and bowl-down (**D**) conformations (Fig. 3-16a,b). [255, 271] The raw experimental STM image in Fig. 3-16c shows a nanoscale area of gold (111) surface covered with self-assembled ad-layer of buckybowls (see Supplementary Note 3-2.1). The global two-dimensional fast Fourier transform performed on the data from Fig. 3-16c reveals a presence of two hexagonal patterns (inset in Fig. 3-16c), that are rotated by 30° with respect to each other, with their lattice constants different by a factor of $\approx\sqrt{3}$. Such reciprocal space structure indicates an alternation of STM tunneling current at every third molecule in the self-assembly. This can be explained by the formation of the so-called **2U1D** structure [271] in which every third molecule appears in the bowl-down state, and is associated with an increase in STM tunneling current (see Figure 3-29), whereas the rest of the molecules reside in the bowl-up state. The relatively weak intensity of FFT peaks associated with **2U1D** structure (inner hexagon in the inset of Fig. 3-16c) suggests only a partial formation of **2U1D** structure across the field of view. In addition to the degrees of freedom associated with **U** and **D** states, each buckybowl can preside in several azimuthal rotational states. A simple visual inspection of several randomly selected areas of the STM image (such as the one illustrated in Fig. 3-16d) as well as an application of more advanced statistical tools such as principal component analysis [272] (Fig. 3-16e) suggests that a likely number of rotational classes needed to be considered for this dataset is four (see Supplementary Note 3-2.2 and 3-2.4, as well as Figure 3-30 and Figure 3-31). While a presence of three rotational states on (111) surface was expected from earlier studies [271], we assign an occurrence of an additional rotational class to a presence of imperfections of the molecular film and/or underlying substrate.

Due to the relative proximity of ground state energies in **U** and **D** conformations [255, 271], it is likely that certain perturbations will induce a transition between the two structural states yielding a deviation from ideal **2U1D** periodic structure in the molecular film. Indeed, visual inspection depicted in e.g. Fig. 3-16d shows a presence of disorder (i.e., distortion of periodicity associated with **2U1D** structure) in both **U** and **D** structural states, as well as in distribution of the molecules rotational classes. Recent experimental and

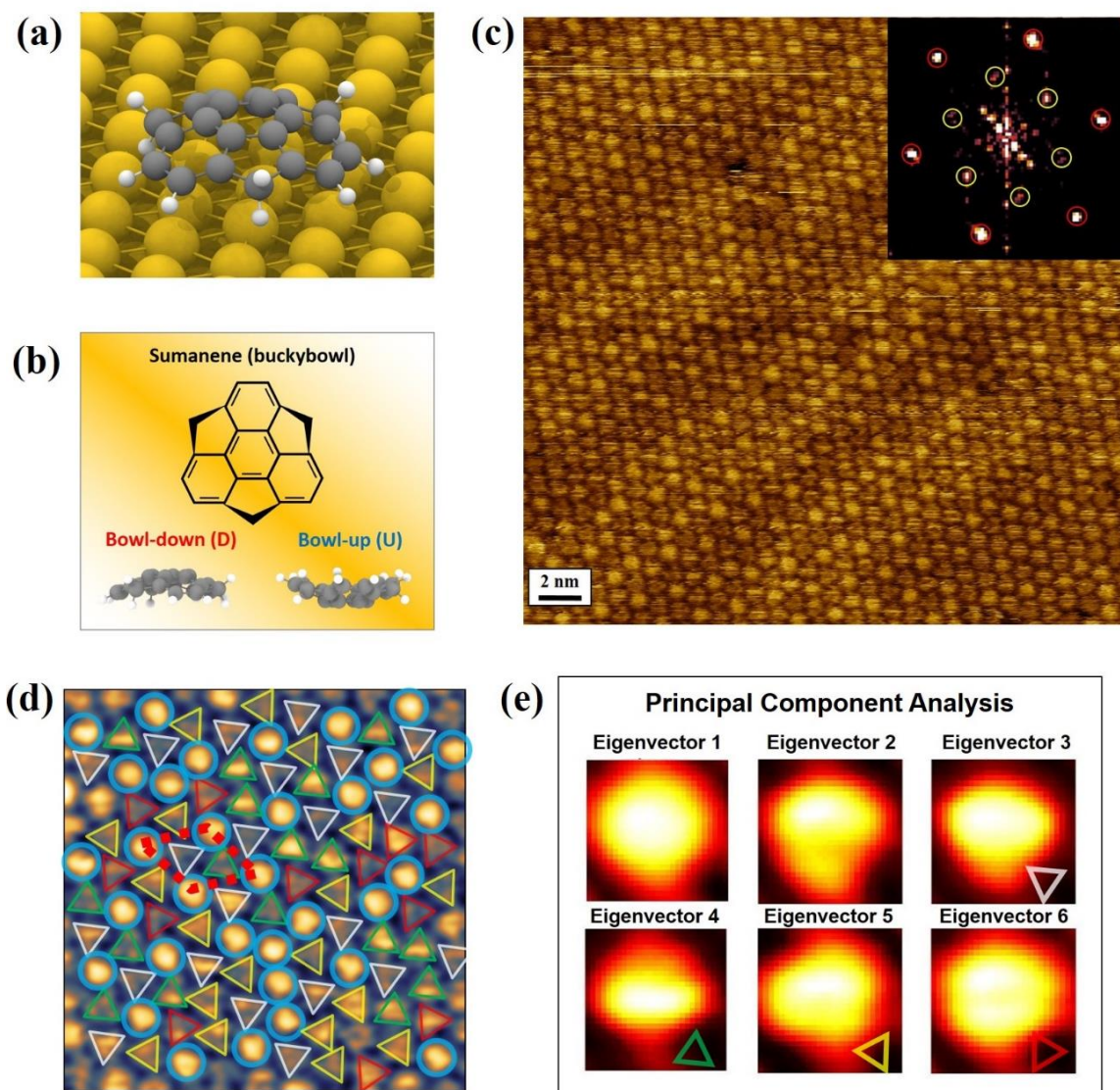


Figure 3-16. Description of the system and physical priors.

Schematics of (a) sumanene molecule on a gold substrate and (b) bowl-up (**U**) and bowl-down (**D**) conformational states. (c) Experimental STM image (raw data) of sumanene ad-layer structure on Au(111) surface. Inset shows a global Fast Fourier Transform of the STM image. The inner hexagon marked with yellow circles is due to formation of **2U1D** structure in certain parts of the image. (d) Visual inspection of different conformational states (**U** and **D**) and rotational states from selected area of the STM image. The **D** states are marked by blue circles. The **U** states are marked by triangular whose orientation reflects a presence of different rotational states. (e) Principal component analysis on the dataset containing all the 938 molecules extracted from (c). The first eigenvector is the mean value whereas the rest five eigenvectors correspond to the largest variances in the dataset (see also Supplementary Note 3-2.4) and can be linked to presence of different rotational classes (schematically indicated in the bottom right corner) observed e.g. in (d).

theoretical studies suggest that a controllable formation of various architectures in 2D buckybowls ad-layer is feasible through a manipulation via the SPM tip [255, 273] or physical adsorption of certain chemical species [271], which could potentially lead to a realization of information storage molecular device or systems for molecular level mechanical transduction. Furthermore, due to a presence of multiple rotational states in addition to bowl up/down structural conformations, such system can be viewed as an ideal playground for probing an interplay between multiple order parameters, the molecular analog of multiferroic systems. [16] However, one of the critical issues in these efforts lies in being able to identify ('read out') all the individual building blocks in different molecular formations on a scale of hundreds and thousands of molecules. This requires tools that would enable a classification of non-strictly periodic structures in the STM images as well as extracting information on the 'internal' structure of individual units (molecules) in an automated and reliable fashion. Unfortunately, while average image analysis methods such as Fourier transform (FFT) and principal component analysis (PCA) of the STM data described above are useful in establishing physical priors, these methods alone are not sufficient for obtaining an accurate information on spatial distribution of the structural and rotational molecular states in such systems and on the possible spatial correlation between the corresponding order parameters. Below we demonstrate how adoption of Markov random field model and convolutional neural networks aided by density functional theory (DFT) simulations of STM images allows to classify bowl-up/-down structural states and different rotational classes in an automated and accurate fashion (Fig. 3-17).

Molecular self-assembly as Markov Network. Markov Random Field (MRF) is a mathematical model that allows representing a long-range order of a system through defining only local interactions. [274, 275]. We use as an example the molecular system of buckybowls that can reside either in different structural conformations (**U** or **D** states); later we also apply this scheme to the analysis of multiple azimuthal rotational states of buckybowls. The posterior probability distribution of the model can be described using Bayes's formula as:

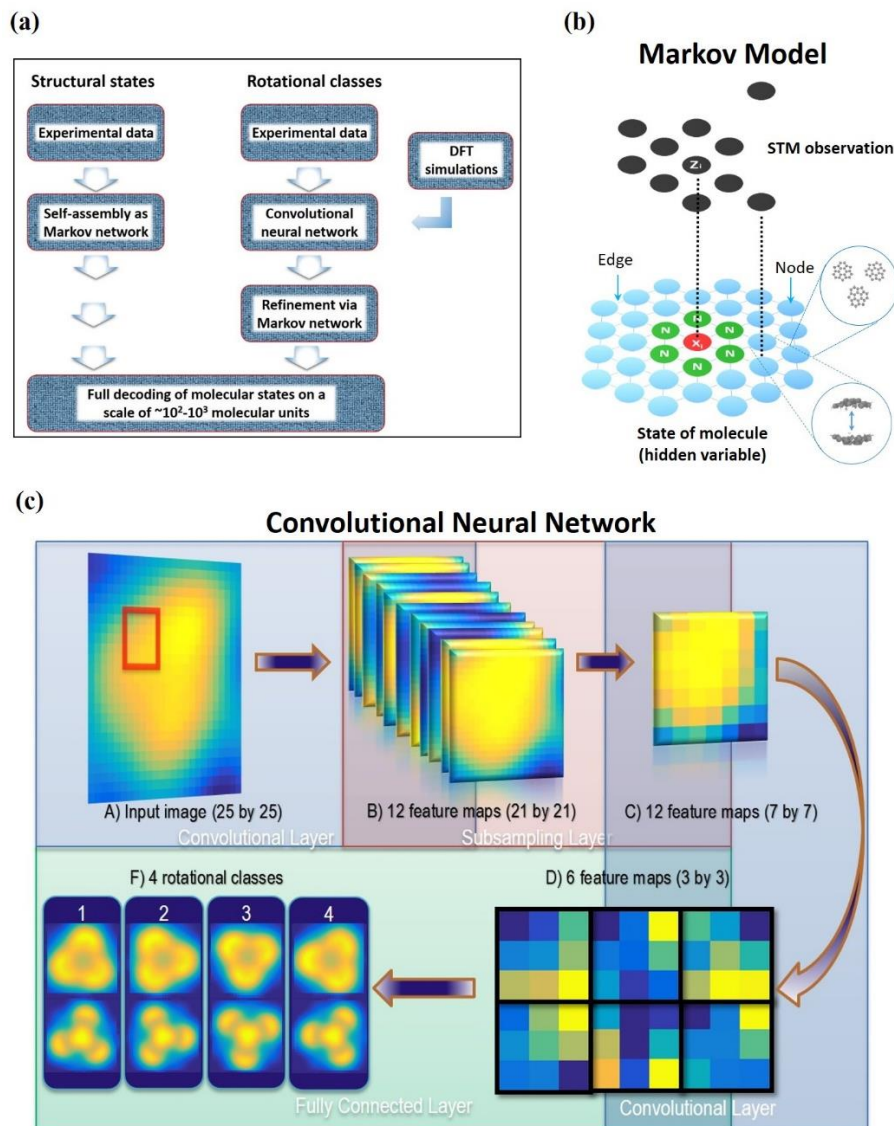


Figure 3-17. Workflow overview.

(a) Block chart diagram of the workflow for classification of **U** and **D** states, and different rotational classes. (b) Graphical Markov model structure used for analysis of a molecular self-assembly of buckybowl. Here Markov network on a regular lattice acts as a prior over hidden variables (associated with the state of the molecule) in a model which is coupled to an array **z** of experimental observations (STM signal). (c) Schematics of convolutional neural network (CNN). The 12 convolutions of a size (21px $\times$ 21px) are generated by applying 12 kernels of a size (5px $\times$ 5px) to the input image. The filters are shifted across the image with a step size of 1px. These convolutions are subsampled into 12 maps of a size (7px $\times$ 7px) using average pooling technique. The second convolution layer is formed by applying 6 kernels to an input from the previous layer. At the end of the network, a fully connected layers contains four neurons corresponding to different rotational classes.

$$P(X = x|Z = z) = \frac{P(Z = z|X = x)P(X = x)}{P(Z = z)} \quad (3.1)$$

where $P(X = x|Z = z)$ describes the probability of molecule belonging to a state x given the observation z (information from the experimental image) and is proportional to likelihood $P(Z = z|X = x)$ of the particular configuration leading to observed outcome multiplied by the prior distribution probability $P(X = x)$ of such configuration in the absence of observation and based purely on our assumptions about the model. The $P(Z = z)$ plays a role of a normalization constant. The priors $P(x)$ can be represented *via* MRF, which makes use of an undirected graph $G = (V, E)$, where $V = \{1, \dots, n\}$ is the vertex set associated with random variable X , and E is a set of edges joining pairs of vertices. The underlying assumption of Markov property is that state of an element is explicitly dependent only on the states of its neighbors,

$$P(X(i)|all\ lattice\ points\ except\ i) = P(X(i)|nearest\ neighbors\ of\ i) \quad (3.2)$$

Importantly, the explicit Markov structure *implicitly* carries longer-range dependencies; hence, it directly translates into the physics of our problem. Note that these priors are directly linked to the fundamental physics of the system, namely the presence of short-range interactions in molecular assembly which are now explicitly taken into account during image analysis.

We represent our data in a form of graph in which each molecule is represented as a node (vertex), and edges are connections to each molecule's nearest neighbors (Fig. 3-17b). The k-d tree method with Euclidean distance metric is used to identify up to 6 nearest neighbors for each molecule. The posterior distribution $P(X = x|Z = z)$ of an MRF can be factorized over individual molecules [276] such that

$$P(x|z) = \frac{1}{Z} * \prod_{\langle i,j \rangle} \Psi_{ij}(x_i, x_j) \prod_i \Psi_i(x_i, z_i) \quad (3.3)$$

where $\Psi_i(x_i, z_i)$ represents unary potential given observation z , $\Psi_{ij}(x_i, x_j)$ are pairwise potentials over the connected neighbors, and Z is the partition function over the posterior

MRF. The potentials are defined based on our knowledge about physicochemical properties of the molecular system. For analysis of molecules conformational changes, each node in our model can reside either in **U** state or **D** state. We then assign the unary potentials $\Psi_i(x_i, z_i)$ over molecular states based on the proximity of a particular molecule's intensity in the STM image to the threshold value between the states T . The simplest threshold value is the mean value of all intensities after normalization and outlier removal (Supplementary Note 3-2.3). Therefore, node probabilities are calculated as

$$\Psi_i(x_i = 1, z_i = I_i) = \frac{1}{1 + \text{Exp}[S * (T - I_i)]}, \quad (3.4a)$$

$$\Psi_i(x_i = 2, z_i = I_i) = 1 - \Psi_i(x_i = 1, z_i = I_i) \quad (3.4b)$$

where $I_i \in [0,1]$ is the intensity of a given molecule i , and S is a parameter that controls the growth rate of the logistic function. This results in two logistic functions, which classify the molecular intensities far away from the threshold as belonging to their corresponding class with probability of 1, but provide more flexibility in the region around the threshold value itself. We proceed to assigning pairwise potentials $\Psi_{ij}(x_i, x_j)$ for our molecular system. The optimal **2U1D** configuration proposed above is characterized by six **U** molecules surrounding one **D** molecule, such that **D** molecule is never allowed to have the nearest neighbor in the same bowl conformation. As we are interested in the distortion of an ideal structure, this condition is relaxed by introducing a disorder parameter p . The new probabilities used in our MRF model are summarized in Table 3-1. Finding an exact solution to MRF model is intractable in our case as it would require examining all 2^n combinations of state assignments, where n is the number of molecules, that is, about 1000 for examined images. However, one can obtain close approximate solution by using a max-product loopy belief propagation method, which is a message-passing algorithm for performing inference on MRF graphs, with unary and pairwise potentials as an input [277, 278] (See also Supplementary Note 3-2.5). We note that by tuning a graph structure and/or form of the potentials one can easily apply this approach to

other molecular order parameters (such as lateral rotations, as we will show later in the paper) or even different molecular architectures.

Table 3-1. Assignment of pairwise potentials based on our knowledge about the molecular system

Node/Neighbor	Bowl-down (D)	Bowl-up (U)
Bowl-down (D)	(0+p)	(1-p)
Bowl-up (U)	0.5	0.5

Classification of azimuthal rotations via convolutional neural network. To determine an azimuthal rotational state of each molecule in the image in an automated fashion, we employ an approach based on convolution neural network (cNN).[279] The cNN based image analysis has been successfully used in recent years in various areas of science and engineering ranging from cancer detection to satellite imaging,[280, 281] but has yet to be applied to atomic-resolved and molecular-resolved imaging.[282] The schematics of cNN adopted for the present study is shown in Fig. 3-17c. It consists of two convolutional layers interspersed with a subsampling/pooling layer, and a fully connected layer. The convolution layer is formed by running *learnable* kernels (‘filters’) of the selected size over the input image (or image in the previous layer), whereas the sub-sampling layer uses average pooling technique to reduce the size of the data. Fully connected layer at the end of the network contains as many neurons as the number of classes/states to be predicted. The learning of kernels is performed through a convolutional implementation of the backpropagation algorithm.[283]

The cNN is trained on a set of synthetic STM images (25000 samples) obtained from DFT simulations of different rotational classes (see more details in the next paragraph). Note that an information on the bowl conformational states (**U** and **D**) is inferred from the MRF analysis and is not therefore treated by cNN (when treated, the adopted cNN scheme produces much poorer overall accuracy in **U/D** states classification compared to MRF analysis).

Generation of synthetic STM data. To ascertain the applicability and robustness of our machine learning and pattern recognition methods for general STM data, we start with constructing a synthetic dataset on a model system. We generate synthetic STM images by Markov chain Monte Carlo sampler using inputs from DFT calculations of electronic charge density distribution in the molecules. We work under the commonly adopted assumption that ‘realistic’, experimental STM image can be viewed as a ‘distorted’ DFT simulation of a charge density distribution in the system (Fig. 3-18a).[284, 285] For the molecular system under consideration, one possible type of the distortion is an admixture of another azimuthal rotational state to a given structural configuration of an individual molecules. Indeed, if two (or more) states are separated by a relatively low activation barrier, such as in the case of buckybowls’ distinct rotational states,[255] the system may switch between these states during the acquisition of STM tunneling current over the molecule. As a result, the STM image will be a dynamical average of two (or more) states.[286, 287] This effect may be especially pronounced during the room-temperature measurements, small tip-sample separation distances, or high setpoint current density. In addition, the blurring effect associated with a presence of the STM tip and a signal-dependent Poisson noise were incorporated in our model (Fig. 3-18a). Here, blurring defines the convolution with the STM probe function, whereas Poisson noise is associated with the tunneling statistics.

Testing our methods on synthetic data. The MRF approach results in a remarkably accurate identification of molecular **D** and **U** states in scenarios where the distribution of the STM intensities in the synthetic data closely resembles the experimental one (note that different rotation angle with respect to the substrate results in a variation of STM signal intensity[255]). This is illustrated on synthetic dataset described in Fig. 3-18b,c for which only 4 out of total 1225 molecules are misidentified by our classification scheme (Fig. 3-18d). The overall total error rate (ratio of misidentified molecules) as a function of p -value and intensity distributions is shown in Fig. 3-19. Generally, this approach vastly outperforms simple mean-value and/or average value thresholding (Fig. 3-19a) and allows accurate classification of **U** and **D** states for a wide range of intensity distributions (Fig. 3-19b) where no estimations regarding the p -value is available *a priori*. In the case of data

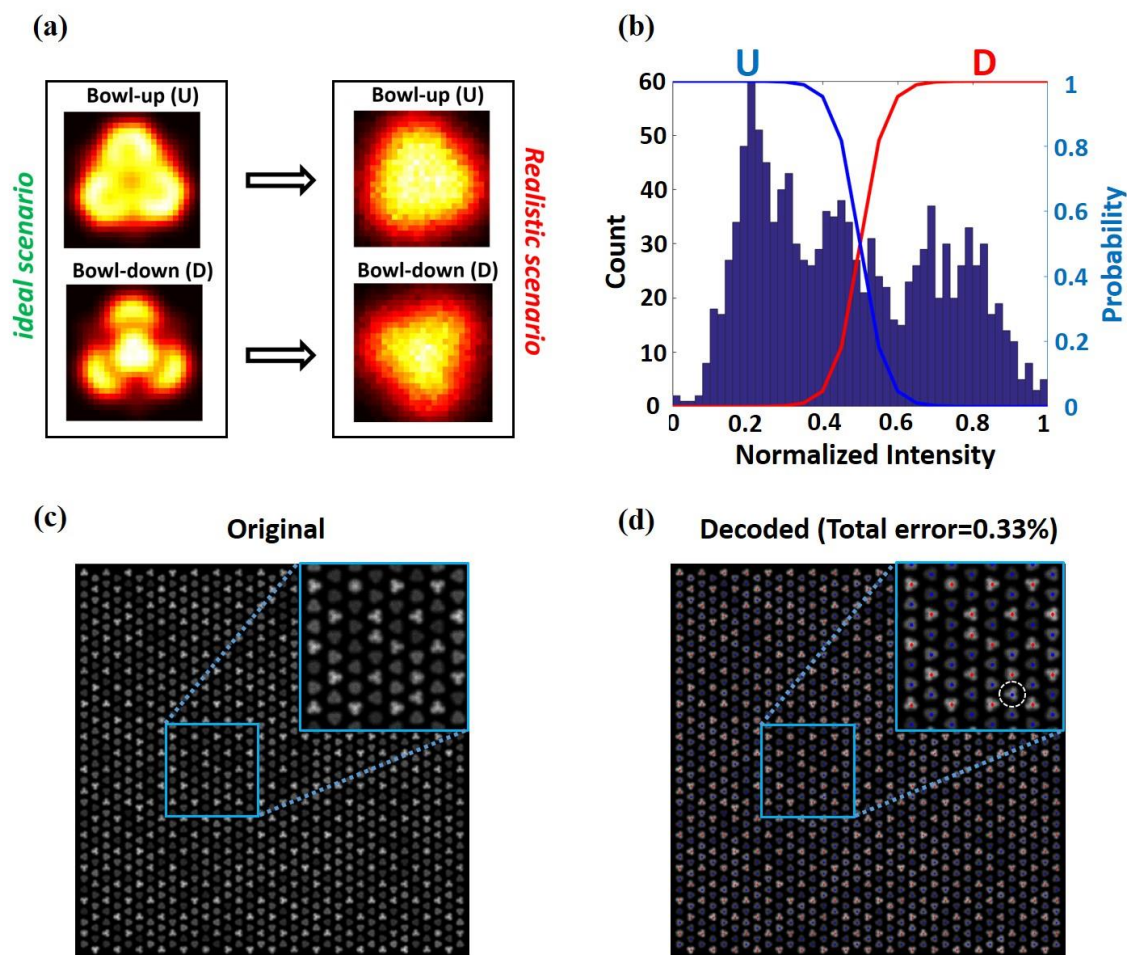


Figure 3-18. Generating and analyzing synthetic data.

(a) Schematics for generating synthetic data for each molecule. Left: DFT-simulated STM images of bowl-up and bowl-down configurations. Right: Same images corrupted by admixing a proximate in energy azimuthal rotational state, blurring (effect of STM tip) and Poisson noise. (b-d) Application of graphical Markov model to synthetic data generated by Markov chain Monte Carlo sampler using inputs from density functional theory calculations of electronic charge density distribution in molecules. The number of molecules in synthetic dataset is 1225. (b) Intensity distribution for synthetic data. Two logistic functions overlaid (see text for details). (c) Real space distribution of molecules in synthetic dataset. (d) MRF based decoding of **D** and **U** states from image in (c). The **D** and **U** states are denoted by red and blue, respectively. The total error (ratio of misidentified states) was 0.33 % or about 4 (out of 1225) molecules. One of the misidentified state is denoted by circle in the inset in (d).

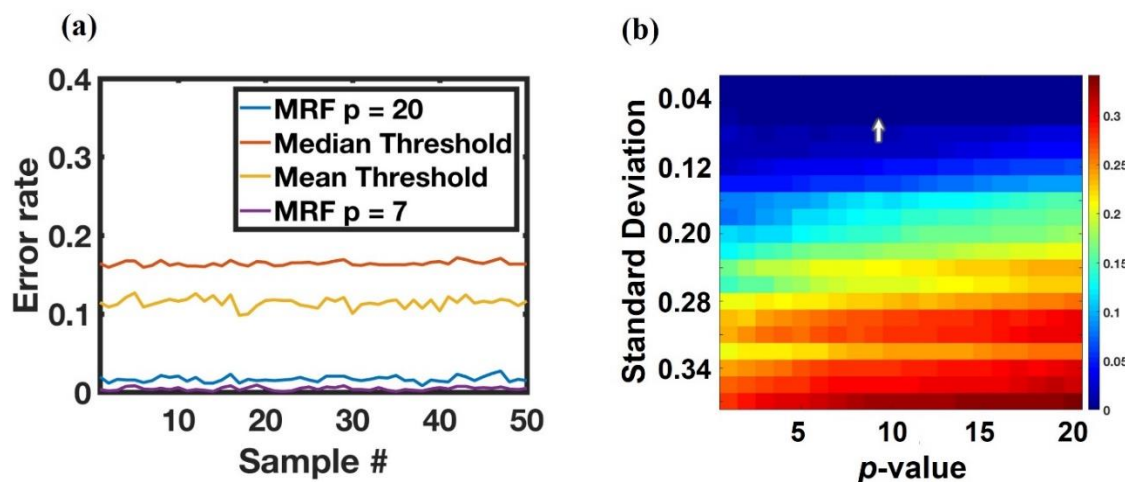


Figure 3-19. Analysis of MRF error rate.

(a) Comparison of error rates, determined as a proportion of misidentified bowl-up and bowl-down conformations, for MRF analysis with two different p values (7 and 20), median thresholding, and mean thresholding. (b) Error rate as a function of standard deviation of normalized STM intensity distributions and an optimization parameter (p -value). The arrow shows the value of these parameters in the synthetic data used in Fig. 3-18b,c.

described in Fig. 3-18b,c, for example, increase of the p -value by a factor of ~ 3 would result in total error increase by less than $\sim 1\%$ (see Fig. 3-19a).

We proceed to extracting information about the azimuthal rotational state of the individual **U** and **D** units in the synthetic STM image (Fig. 3-20a). The dependence of classification accuracy on the hypothetical admixture ratios of a proximate rotational state (which are potentially located ‘in-between’ the rotational classes used in our classification scheme) and the cNN error are shown in Fig. 3-20b and 3-20d, respectively. It is easy to see from Fig. 3-20b that one can obtain a reliable classification of molecules rotational states even for relatively large ratio of the selected admixed state.

The unique aspect of the proposed approach is that it is possible to incorporate certain physical constraints, into the cNN-based analysis for obtaining more accurate decoding results. Here, we incorporate the effect of steric repulsion between molecules. In this case, we can use the cNN-calculated probabilities of azimuthal rotational states as prior probability distributions for another MRF model. Consider, for example, that two nearest-neighbor molecules are highly unlikely to have the same azimuthal rotational states if they reside in the same conformation state (either **U** or **D**) in the self-assembly.[271] We may therefore assign 1% probability of each class to have a neighbor of its own class and equal 33% probabilities to have a neighbor of other 3 rotational classes (see Table 3-2). Total probabilities are then normalized to sum to 1. Then, similar to earlier description, we perform decoding using loopy belief propagation in order to acquire more accurate solution (Fig. 3-20c).

Application to experimental data. Having verified that our algorithm is capable of working on synthetic data that mimics the ‘laboratory conditions’, we move to applying it to real experimental STM images of buckybowls on Au(111) substrate described in Fig. 3-16. The FFT mask with Hamming window is first applied to the STM image for removing a large-scale periodic contribution from the substrate. The MRF decoding of **U** and **D** states and the cNN-based decoding of azimuthal rotational states are summed in Fig. 3-21a. The physical priors and classes for MRF and cNN were taken from FFT and PCA analysis, respectively (see Fig. 3-16). A simple visual inspection of results in Fig. 3-21c can confirm a high accuracy of our method for experimental data.

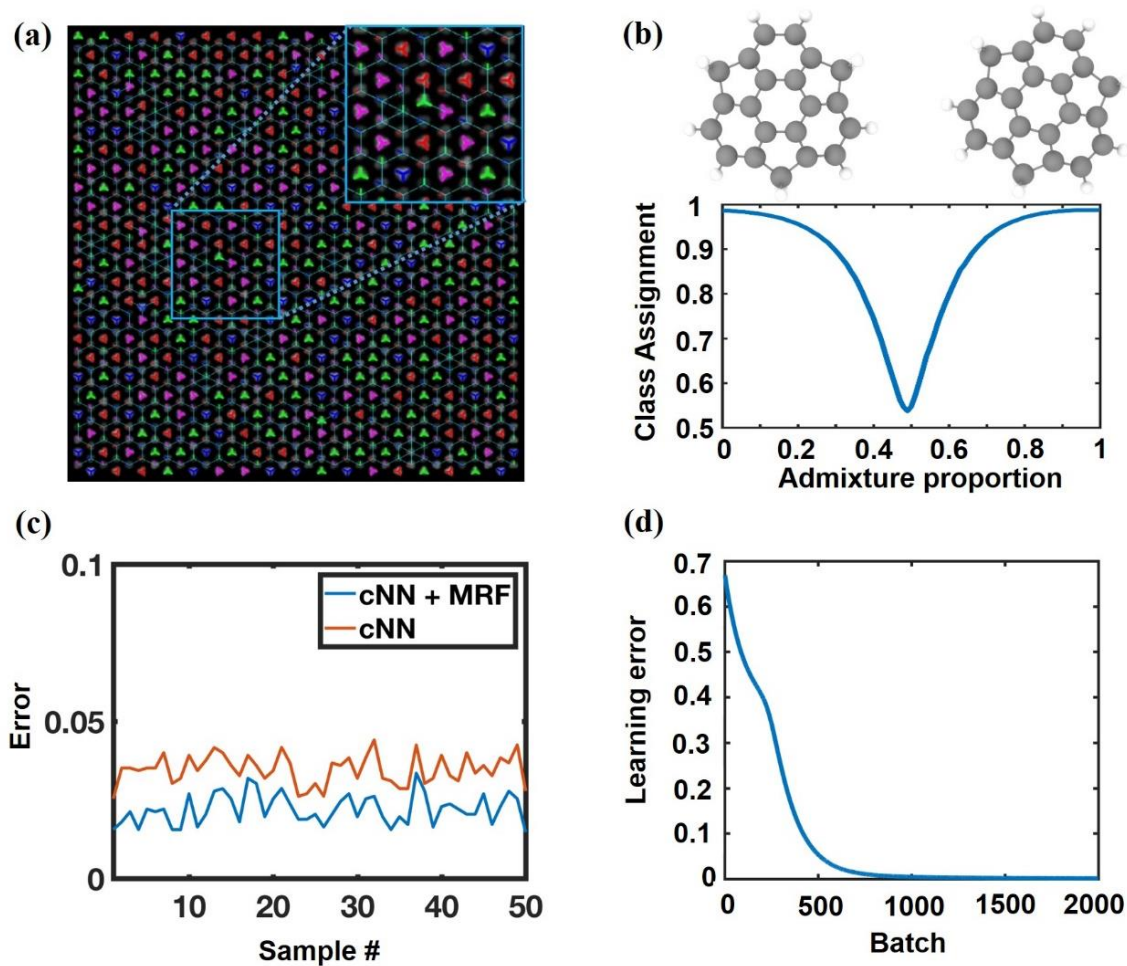


Figure 3-20. Convolutional neural network (cNN) decoding.

(a) Convolutional neural network (cNN) produced decoding of rotational states. Thick-line tri-pointed stars describe the rotational classes of **D** states, whereas thinner lines show rotational classes of **U** states (both **U** and **D** states were classified via MRF prior to applying cNN). (b) The dependency of cNN accuracy (probability of a correct state assignment) on the admixed proportion of a different rotational state. The two rotational cases considered are shown at the top of the image (c) Comparison of error rate for cNN only and cNN+MRF approaches. (d) cNN learning error.

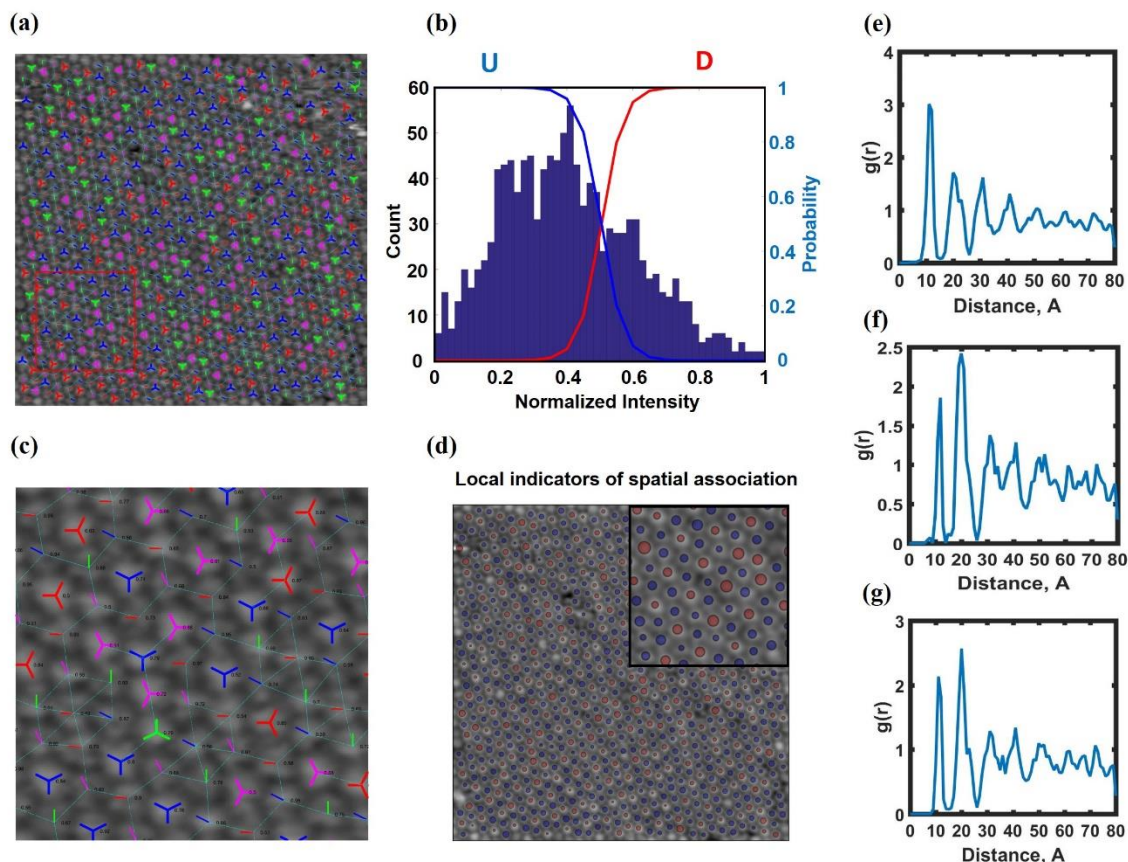


Figure 3-21. Application of MRF and cNN to experimental data of buckybowls on Au(111).

(a) Decoding of bowl-up/down states ($p=7$) and rotational states for the experimental image described in Fig. 3-16c. For cNN analysis, an additional MRF refinement of rotational classes was performed using the same probabilities as for the synthetic data. (b) Histogram of STM intensities for all identified molecules. (c) Zoomed-in area from red rectangle in (a) where numbers denote an accuracy of state determination. (d) Spatial correlation analysis based on local Moran's I statistics. The molecules are color-coded such that molecules in **D** and **U** states are presented as red and blue circles respectively; the size of the marker is scaled such that it is proportional to a value of local Moran's I . (e-g) Pair correlation function constructed from decoded experimental data for all molecular states (e), only bowl-down states (f), and one of the rotational classes (g). See also Figure 3-32.

Once a full decoding is performed, it becomes possible to construct a pair distribution function (PDF) for molecular states of interest. In turn, these provide further insight into the nature of (dis-)order in molecular film, such as whether a disorder is correlated or random[258]. The PDFs for all the molecular states, bowl-down molecular confirmations, and one of the rotational classes are shown in Fig 3-21e-g. The molecules clearly display a well-defined long range *positional* order, as evident from Fig. 3-21e. On the other hand, an analysis of PDF for different molecular bowl conformations suggests that neither long range **2U1D** nor perfect long range **3U** orientational orders suggested previously[255, 271] are realized for a given system. Indeed, the former must result in a disappearance (or a very strong suppression) of a peak at ≈ 11 Å in Fig. 3-21f, whereas the latter would closely resemble all-molecules PDF in Fig. 3-21e; these, however, was not observed. Interestingly, our analysis also shows a close resemblance in a behavior of PDFs for structural and rotational states (Fig. 3-21f and 3-21g, respectively) within first several coordination ‘spheres’ implying certain correlation between the two associated orders, that is, bowl-up/down switching is associated with a formation of certain rotational (dis-)order in the inverted molecules.

We further explore a nature of disorder in the molecular self-assembly by searching for *local* correlations between molecule bowl inversion and azimuthal rotation of the neighboring molecules. To obtain such an insight, we construct a spatial correlation map describing a possible interplay between these two different orders. Specifically, we adopt a method based on calculation the so-called Moran’s I that can measure a spatial association between the distributions of two variables at nearby locations on the *lattice*. [288, 289] The presence of the spatial weight matrix in the definition of Moran’s I allows us to impose constraints on the number of neighbors to be considered (see Supplementary Note 3-2.6). The results for spatial correlation between bowl-up/down configuration and different rotational classes for the first ‘coordination sphere’ is shown in Fig. 3-21d where a different size of circles reflects different values of the Moran’s I across a field of view. Generally, the map in Fig. 3-21d implies a spatial variation in coupling between the two associated order parameters, which could also be sensitive to presence of defects. The average value of Moran’s I for all molecules is 0.310, whereas the average

value for correlation of rotational classes with bowl-up and bowl-down molecular conformations are 0.246 and 0.426 respectively. This result indicates that a bowl-up-to-bowl-down flip associated with occurrence of an ‘additional’ **D** molecule requires a larger change in a rotational state of the neighboring molecules (compared to a flip in the reversed direction) in order to compensate for a formation of energetically unfavorable (‘extra’) bowl-down state.

Based on the findings above we propose a two-stage ‘reaction path’ that explains a different correlation values of rotational states with neighboring bowl-down and bowl-up structural conformations as schematically depicted in Fig. 3-22.

Specifically, in the first stage, a creation of ‘extra’ bowl-down state elevates the energy of the system, which is then relaxed in the second stage of the ‘reaction’ via adjustment of rotational states in nearby molecule(s). The latter is associated with the obtained values of Moran’s I. It is crucial to note that unlike previous studies which only considered a bowl inversion process for an *isolated* single molecule [271], our analysis allowed to obtain a deeper knowledge of local interaction processes that involve a lateral switching of neighboring molecules in the self-assembled layer. Observation of such an interplay between molecule rotations and its structural conformations provides important clues for understanding local degrees of freedom in the molecular ad-layer which is crucial in terms of its potential applications in multi-level molecular memory devices.

3.4.6 Discussion

To summarize, we have developed a multi-stage pattern recognition approach which encompasses ab-initio simulations, Markov random field and convolutional neural networks for a detailed characterization of surface molecular architectures in the typical field of view ($\sim 10^2$ - 10^3 molecules) of STM experiment.

We now comment on several potential limitations of the methods and possible ways to overcome them. First, the physical priors used for input in both MRF and cNN could be in

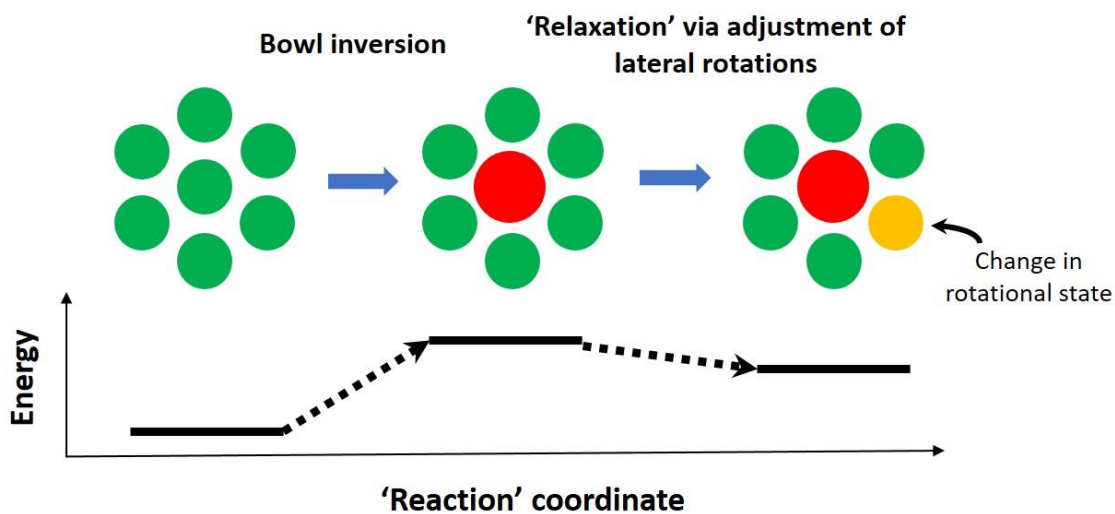


Figure 3-22. Schematics of the proposed reaction path for bowl inversion and its effect on the neighboring molecules.

. In the first stage, the inversion of the bowl-up (green circle) to bowl-down (large red circle) state elevates the energy of the system. In the second stage, the system gets relaxed by adjusting the rotational state of nearby molecule (shown by orange circle).

future extracted (in addition to, or even instead of, FFT and PCA analysis) from state-of-the-art ab-initio analysis and molecular dynamics (MD) simulations thus potentially providing more accurate decoding results. In this regard, it should be noted that low probabilities of class determination for certain molecules, if present, would suggest that some of molecular states in the experimental system were not accounted for by theory. In such case, one must return to the ab-initio modelling stage and reconsider the initial assumptions or adjust parameters. We envision that such process of adjusting (putting constraints on) ab-initio or MD parameters could be automated in future, although this would require an infrastructure capable of performing DFT/MD on-the-fly. Second, it would be also interesting to apply the so-called domain-adversarial training of neural networks[290] which allows to alter theoretically predicted classes based on the observed data. The underlying idea of this approach is that the theoretical and experimental datasets are similar yet different in such a way that traditional neural networks may not capture correct features just from the labeled data. Finally, we foresee that in future a choice of the optimization value p in MRF analysis during an inference of bowl-up/down structural states, as well as during the refinement of cNN results, could be in principle optimized using a statistical distance approach.[132]

Regarding possible further applications of our method we note that in addition to analysis of individual static images of molecular structures, the same analysis can be applied to each individual frame in the STM “movies” of molecular motions (e.g. under external perturbation field) thus providing an invaluable input to molecular dynamics simulations. Furthermore, because our pattern recognition analysis is general in nature, it can be extended to microscopic measurements of structural, electronic, and magnetic orders, as well as their possible spatial correlations, in a variety of condensed matter systems such as, for example, skyrmion lattices.[291]

3.5 Identifying phase transitions in first-principles simulations of relaxor ferroelectrics via variational autoencoders

3.5.1 Introduction

The search for materials with interesting properties has identified relaxor ferroelectrics [292, 293] such as $\text{Pb}(\text{Sc}_{0.5}\text{Nb}_{0.5})\text{O}_3$ (PSN) and $\text{Pb}(\text{Sc}_{0.5}\text{Ta}_{0.5})\text{O}_3$ (PST) as materials of interest due to their dielectric and electrochemical properties, which can be important for transducer and actuator applications. The proportion of Pb-O divacancies in these materials changes the phase diagram of the material significantly. However, the analysis of the simulation data remains mostly analytical through computation of order parameters. [294] In order to expedite the discovery of order parameters that describe the phase transitions, I identified autoencoders as a suitable algorithm for automated exploration of large amounts of data. Autoencoders are a powerful method used for data compression, anomaly detection, data representation transformation, etc. [295, 296] They combine 2 neural networks: an encoder network that takes in the input data and compresses it to a single vector of values, and a decoder, or generative model, that can reconstruct the original data from the vector. However, in many cases, such as images, a single value in a dimension cannot describe the data.

Instead, we need to use a variational autoencoder (VAE), [297] which generates a probability distribution in the latent variable space, which is useful for data such as images and videos where we want to model a probability distribution space. [298] We assume that each data sample is from a distribution $p(x)$ and is explained by a hidden variable t . VAE uses variational inference to approximate $p(t|x)$ with a distribution $q(z|x)$, which can be, for example, a mixture of gaussians. We use the Kullback Leibler (KL) divergence metric to estimate the similarity of the probability distributions [299]. VAE have been successfully used to estimate phase transitions in 2D Ising model and 3D XY model, without introducing any a priori knowledge about the system into the model. [300]

3.5.2 Results

The input data in this case was taken from first-principle simulations of $\text{Pb}_{1-x}(\text{Sc}_{1/2}\text{Nb}_{1/2})\text{O}_{3-x}$ relaxor ferroelectric. Analytically derived order parameters predict transition temperature dependence on the concentration X of nearest neighbor [Pb – O] divacancies in the material. [294] Simulations were recorded at certain X and temperature T . Each simulation represented a 40 by 40 by 40 supercell, with each point recording a local distortion vector which record displacement of lead atom from its ideal perovskite position in three dimensions. Data in each timeframe was transformed into a vector of length 192000 ($40 \times 40 \times 40 \times 3$). The data was agglomerated by concentration X , with no a priori information about X or T introduced into the model. Each frame of the simulation used was considered a data sample. The data was normalized to be in the range of $[-1, 1]$. For this task we chose a fully-connected VAE model with 3 layers in encoder and decoder networks each, representing, correspondingly, 256, 128, and 64 neurons with rectified linear unit activation. [301] Each layer features batch normalization [302] and 10%, 10%, and 20% dropout [303] correspondingly. Final layer of the decoder reconstructs the original 192000-long vector with hyperbolic tangent activation function. The model is shown in Figure 3-23.

For each concentration X , a VAE was trained with 1, 2, and 3 latent variables for 50 epochs. Training with 2 and 3 latent variables resulted in highly-correlated latent variables identification, indicating that 1 latent variable was well suited for describing the underlying physical order parameter.

To identify a phase transition, we examined the plots of latent variables and reconstruction errors for each concentration shown in Fig. 3-24. In a model with 1 latent variable, concentrations resulting in more phase transitions produce distributions with more distinct features. Sharp changes were chosen as indication of transition occurring. Results are plotted on a phase diagram along with analytical order parameter in Fig. 3-25.

We conclude that variational autoencoders are a powerful tool for discovery of order parameters within the simulations data, but more testing is required to be able to accurately identify all present phase transitions.

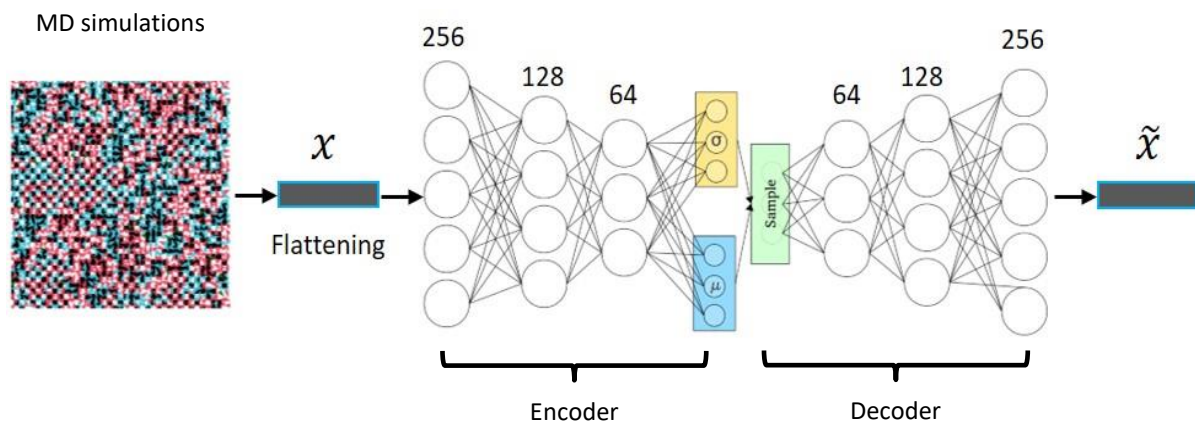


Figure 3-23. Schematic representation of the variational autoencoder model applied to relaxor ferroelectric simulations.

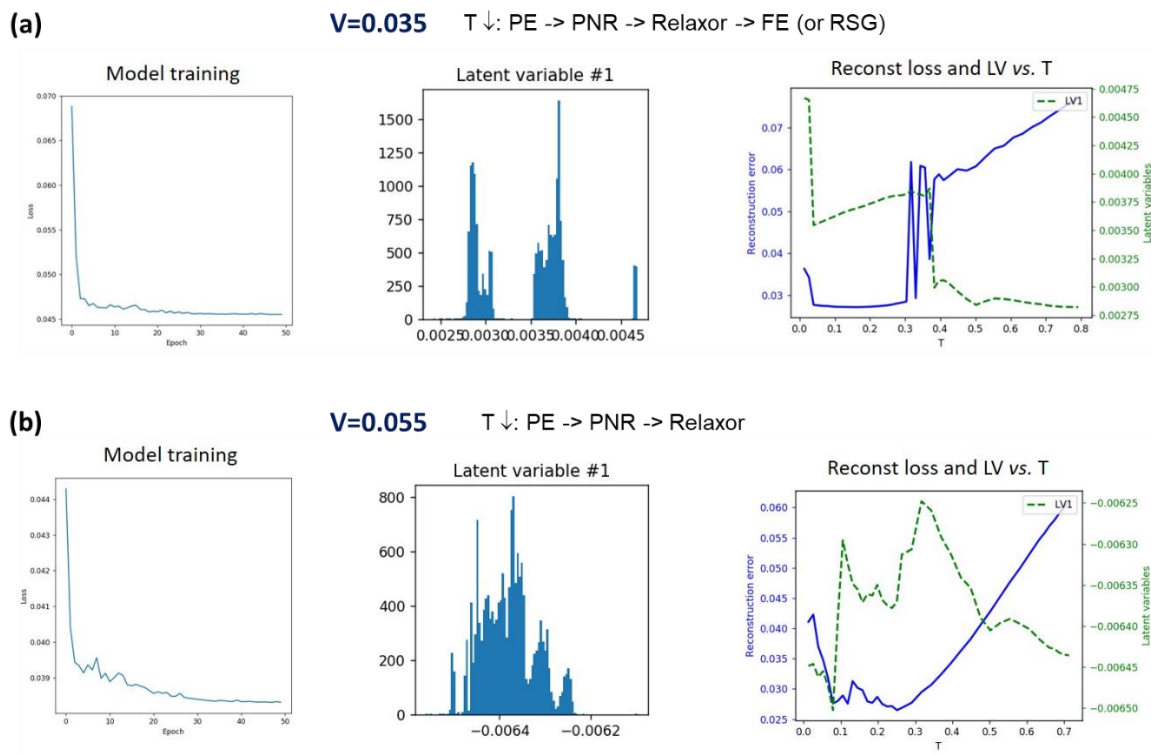


Figure 3-24. Latent Variables and reconstruction errors for different explored concentrations of [Pb – O] divacancies: (a) 0.035 (b) 0.055.

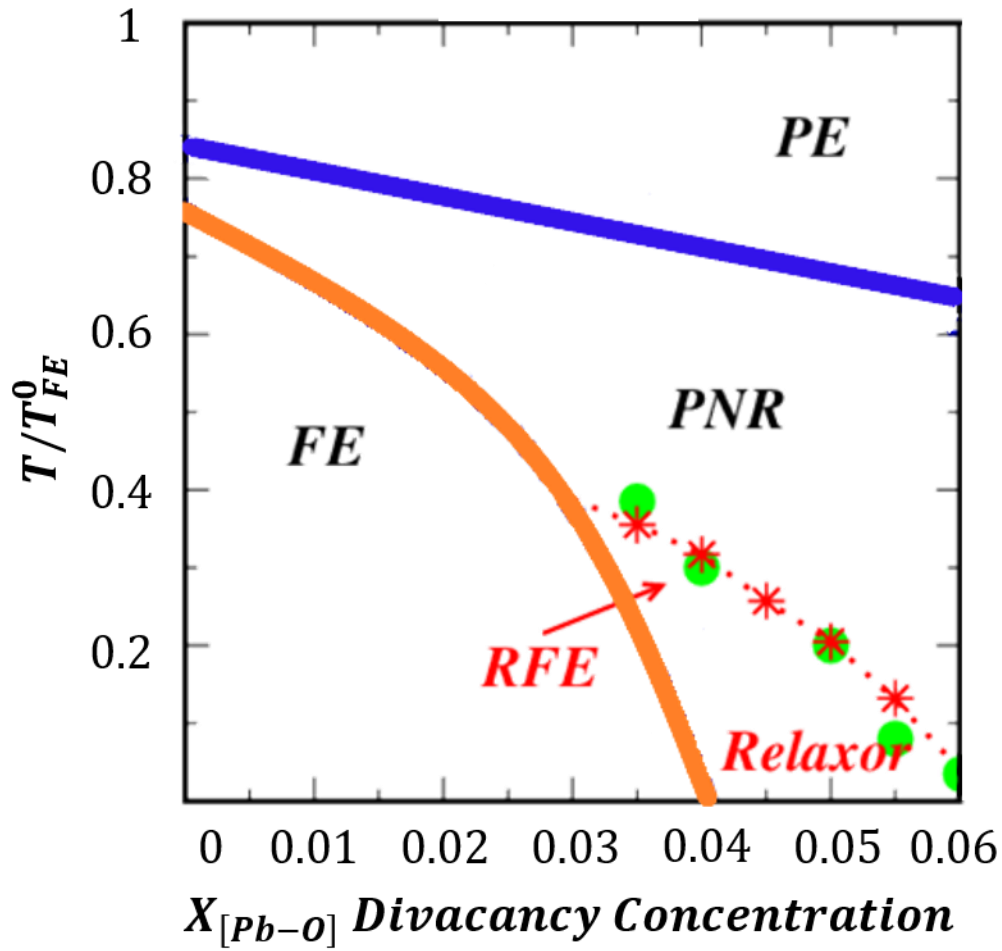


Figure 3-25. Calculated $X[Pb-O]$ vs. T phase diagrams for the system $Pb_{1-x}(Sc_{0.5}Nb_{0.5})O_{3-x}$, with a nano-ordered Sc:Nb-cation configuration.

Labels: PE indicates a normal paraelectric; PNR indicates a system in which chemically ordered regions, with few h_i , have higher polarization than the random matrix; FE indicates a ferroelectric ground-state; RFE indicates a relaxor-region above the FE-ground-state. Blue solid line indicates Burns temperatures. Orange solid line indicates FE-PNR, or FE-RFE transitions. Dotted lines with red asterisk-symbols indicate analytical RFE-PNR or relaxor-PNR transitions. [294] Green dots indicate result from variational autoencoder.

Appendix

Appendix 3-1. Supplementary Information for Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations

Supplementary Note 3-1.1.

To illustrate how the FCN network searches for “dopants” (with higher Z) and “vacancies” we show in Supplementary Figure 3-3 the relevant FCN outputs for graphene and WS₂. In the case of graphene (Supplementary Figure 3-3a), in addition to identifying areas where carbon atoms are missing, the network also identifies areas that break lattice symmetry due to formation of 5-7 reconstruction and are characterized by extended hollow regions for a heptagon (denoted by red arrows).

In the case of WSe₂, the STEM image shows a large asymmetry between two sublattice sites on the honeycomb lattice effectively turning it into a trigonal lattice. Remarkably, the neural network that learned a concept of dopant (with higher Z) and vacancy on one type of the lattice structure, can apply it to a different lattice, which is very similar to the way a human would do it.

Appendix 3-2. Supplementary Information for Learning surface molecular structures via machine vision

Supplementary Note 3-2.1. STM experiment and simulation of STM images

STM experiments were performed with a commercially available STM (JEOL, JSPM-4500S) operated under ultrahigh vacuum at room temperature. Gold (111) substrates were prepared by thermal evaporation of Au onto mica at 300 °C. Sumanene molecules were deposited from a 10 mM solution toluene onto the gold (111) surface at room temperature for 20 min. All STM images were acquired in constant current mode using mechanically cut Au wire as STM tips. To ensure an accurate machine vision-based analysis, the

experimental scan size and image resolution were adjusted in such way that a resolution of a square that encloses an individual molecule was equal or greater than 25px*25px. DFT-based STM simulations were performed within the Tersoff–Haman approximation [304, 305] using PWSCF code.[306]

Supplementary Note 3-2.2. Determination of positions of molecules

The first step in processing the image is extraction of positions of all molecules. We use a template constructed from an average of all ‘bowl-up’ DFT simulated images to perform normalized cross-correlation according to the following formula:

$$\gamma(u, v) = \frac{\sum_{x,y} [f(x, y) - \bar{f}_{u,v}] [t(x - u, y - v) - \bar{t}]}{\left\{ \sum_{x,y} [f(x, y) - \bar{f}_{u,v}]^2 \sum_{x,y} [t(x - u, y - v) - \bar{t}]^2 \right\}^{0.5}}, \quad (\text{s. 3.1})$$

where f is the input image, t is the template, $\bar{f}_{u,v}$ is the mean of $f(x, y)$ in the region under the template, $\bar{t}$ is the mean of the template. For the input image, we used FFT mask with Hamming window to remove a periodic contribution from the substrate (dotted-like-line features in FFT pattern of Fig. 3-16c).

We apply a uniform threshold to the generated correlation surface γ , with cutoff set to 0.35 in order to maximize the number of extracted molecules. As a result we acquire a binary image, in which we identify the connected components and assign their centers as centers of the corresponding molecules.

Supplementary Note 3-2.3. Intensity calculation

Templates for the STM response demonstrate a drastic difference in intensity patterns between ‘bowl up’ and ‘bowl-down’ states: the former have a significant intensity decrease in the center compared to the surrounding regions, while the latter has the opposite. This seems to be visually identifiable in the experimental image, but much less pronounced and more ambiguous. In order to quantify the intensity we extract 15 by 15 pixel image patches

around the center of each molecule identified in previous step. The intensity of the molecule is therefore:

$$I_m = \sum_{x=1}^{15} \sum_{y=1}^{15} i_{x,y}, \quad (s. 3.2)$$

Where $i_{x,y}$ is the intensity of pixel at position x, y in the extracted image patch for molecule m. Since some of the defects are identified as molecules, it is necessary to filter the outliers. We assume that maximum realistic intensity is:

$$I_{max} = \text{mean}(I) + 3 * \text{std}(I), \quad (s. 3.3)$$

where I is the array of calculated intensities for each molecule. Higher intensities are assumed to be outliers due to a defect and set to be equal to I_{max} . We normalize the array by subtracting the minimum value and dividing by maximum value of the result in order to put all intensities into [0, 1] range.

Supplementary Note 3-2.4. Principal component analysis

The aim of PCA can be interpreted as finding a lower dimensional representation of data with a minimum loss of important (relevant) information. Specifically, in PCA one performs an orthogonal linear transformation that maps the data into a new coordinate system such that the greatest variance comes to lie on the first coordinate called the first principal component, the second greatest variance on the second coordinate, and so forth. Hence, the most relevant information (including information on the orientation/rotation of molecules) can be represented by a small number of principal components with the largest variance, whereas the rest of the (low-variance) components correspond to ‘noise’. [186] A simple visual inspection of the data (such as the one in Fig. 3-16d) indicates that one can expect a large variation within the dataset in terms of molecules rotations. Hence such variation should be detected by PCA and corresponding rotational classes can be inferred

through the analysis of PCA eigenvectors. The PCA results for the first 16 eigenvectors can be found in Figure 3-30. Note that since we do not perform mean-centering, the first eigenvector in Fig. 3-16e represents a mean value (i.e. an average state of all extracted molecules). The eigenvectors 3-6 are associated with the rotational classes analyzed manually in Fig. 3-16d. Note that eigenvector 2 can in principle be interpreted as a state influenced by switching between bowl-up and bowl-down conformations induced by the STM tip. Furthermore, eigenvectors 7-16 are likely associated with bowl-up to bowl-down inversion during the scan. Such a switching behavior was not included in the MRF analysis.

Supplementary Note 3-2.5.

As mentioned in the main text, finding exact solution to MRF model is intractable in this case as it would require examining all 2^n combinations of state assignments, where n is the number of molecules, about 1000 per each examined image. We thus need to find an approximate solution, for which we choose to use loopy belief propagation algorithm (max-product version). The detailed description of this method is given, for example, in Ref. [276, 307, 308]. Briefly, from initial configuration, nodes propagate message containing their beliefs about state of the neighboring nodes given all other neighboring nodes messages. This results in an iterative algorithm. All messages start at 1, and are further updated as max-product of potentials and incoming messages:

$$msg(x_j)_{i \rightarrow j} = \max_l \left[\sum_{x_i} \Psi_{ij}(x_i, x_j) \Psi_i(x_i, z_i) * \prod_{k=\text{neighbors of } i \neq j} msg(X)_{k \rightarrow i} \right], (s. 3.4)$$

At each iteration belief is calculated for each node and the state with highest belief is selected, until message update converges:

$$Belief(x_i) = \Psi_i(x_i, z_i) * \prod_{j=\text{neighbors of } i} msg(x_i)_{j \rightarrow i}, \quad (s.3.5)$$

Supplementary Note 3-2.6. Local indicators of spatial association. Local Moran's I analysis

Moran's I is a measure of (overall) spatial clustering of the data in a signal among nearby locations on the lattice in 2D or 3D space. It can be extended to be a measure of spatial correlation between 2 variables distributions over the same spatial structures. The 'correlation coefficient' for global Moran's I is given by [288]

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij} (X_i - \bar{X})(Y_j - \bar{Y})}{\sum_i (Y_i - \bar{Y})^2}, \quad (s.3.6)$$

where N is the number of spatial units, X and Y are variables, $\bar{X}$ and $\bar{Y}$ are corresponding means, and w is the weight matrix defining neighbor interactions. For highly inhomogeneous system, one may use the so-called local indicators of spatial association (or 'LISA') which can evaluate the correlation between two orders at the neighboring points on the lattice for each individual coordination sphere. This is achieved through calculating *local* bivariate Moran's I for each spatial unit [289] such as

$$I_{xy} = \frac{\sum_i \sum_{j \neq i} w_{ij} x_i y_j}{W}, \quad (s.3.7)$$

where x and y are standardized to zero mean and variance of 1. The local Moran's I was used to calculate results presented in Fig. 3-21d.

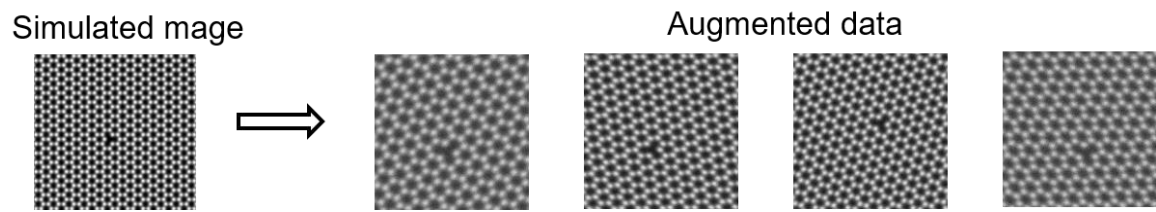


Figure 3-26. Illustration of data augmentation procedure for creating a sufficiently large training set.

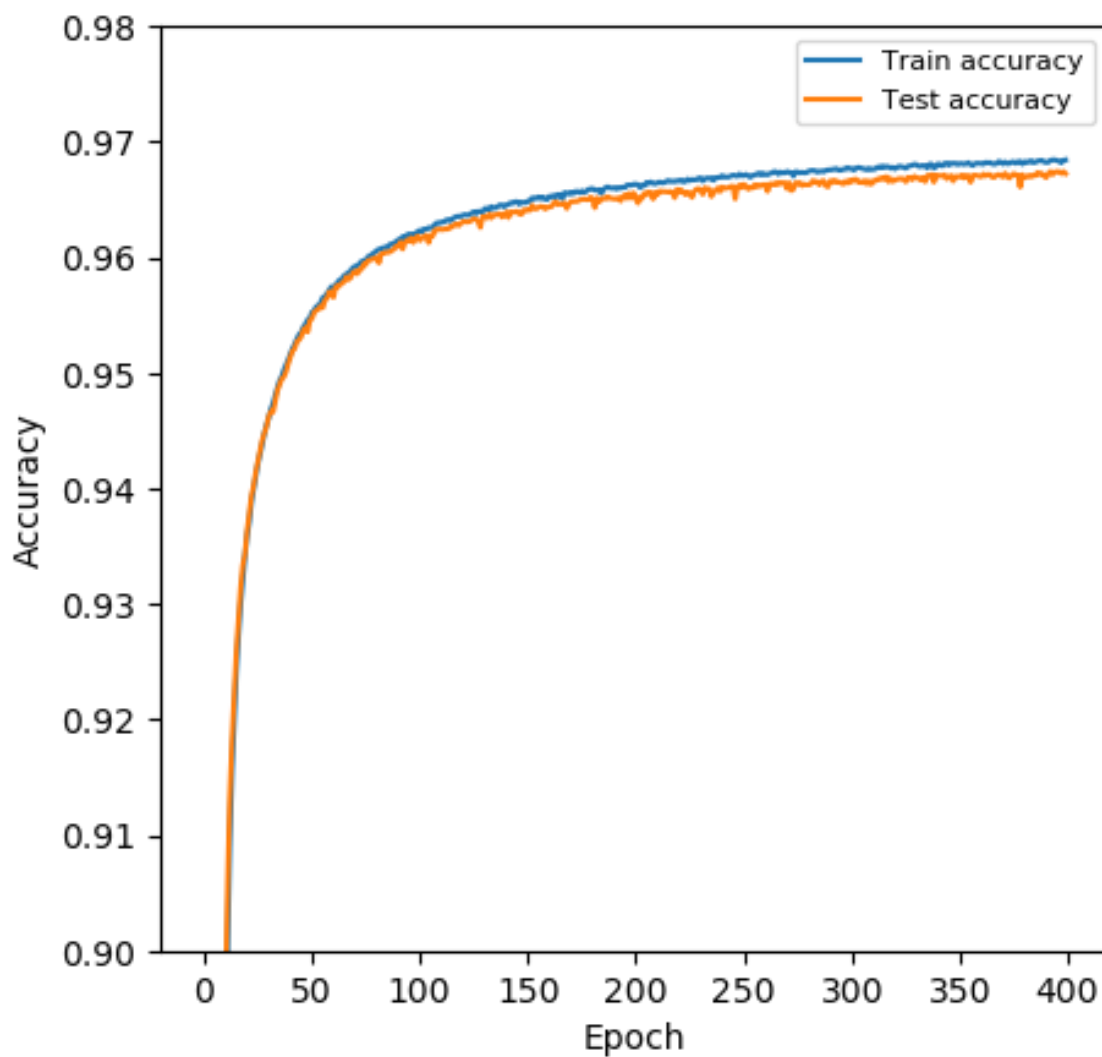


Figure 3-27. Plots of accuracy on the training and validation datasets over training epochs for FCN trained for finding atoms in the lattice. The accuracy on the test set is close to 97 %. The FCNs for finding vacancies and dopants typically achieve about 99 % accuracy.

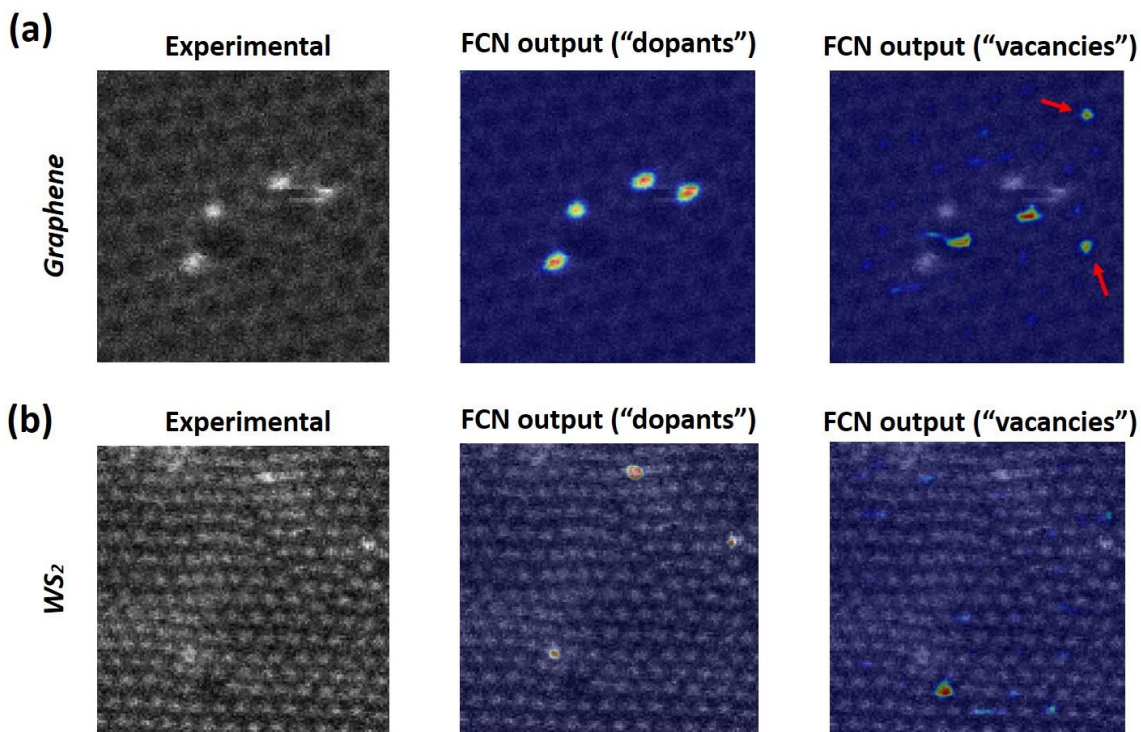


Figure 3-28. Detection of vacancies and vacancy-like structures in graphene and WS₂.

(a) Original experimental image and the softmax outputs overlaid on the experimental image. Arrows denote regions that break lattice symmetry due to formation of 5-7 reconstruction and are characterized by extended darker regions (in a region with heptagon). (b) Same for WS₂. See Supplementary Note 1 for additional details.

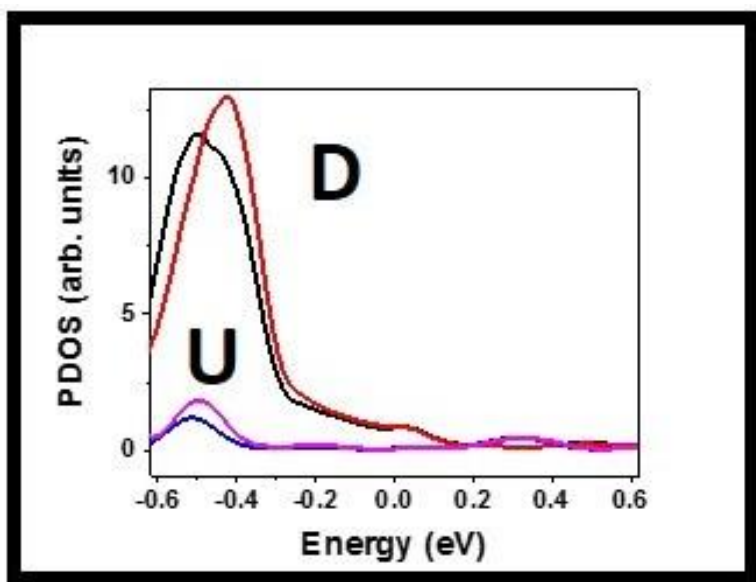


Figure 3-29. DFT calculations of projected density of states (PDOS) for U and D states (summed over all C atoms in the molecule). Strong difference in PDOS at negative energies are associated with enhancement of the STM tunneling current in the experiment at each D molecule.

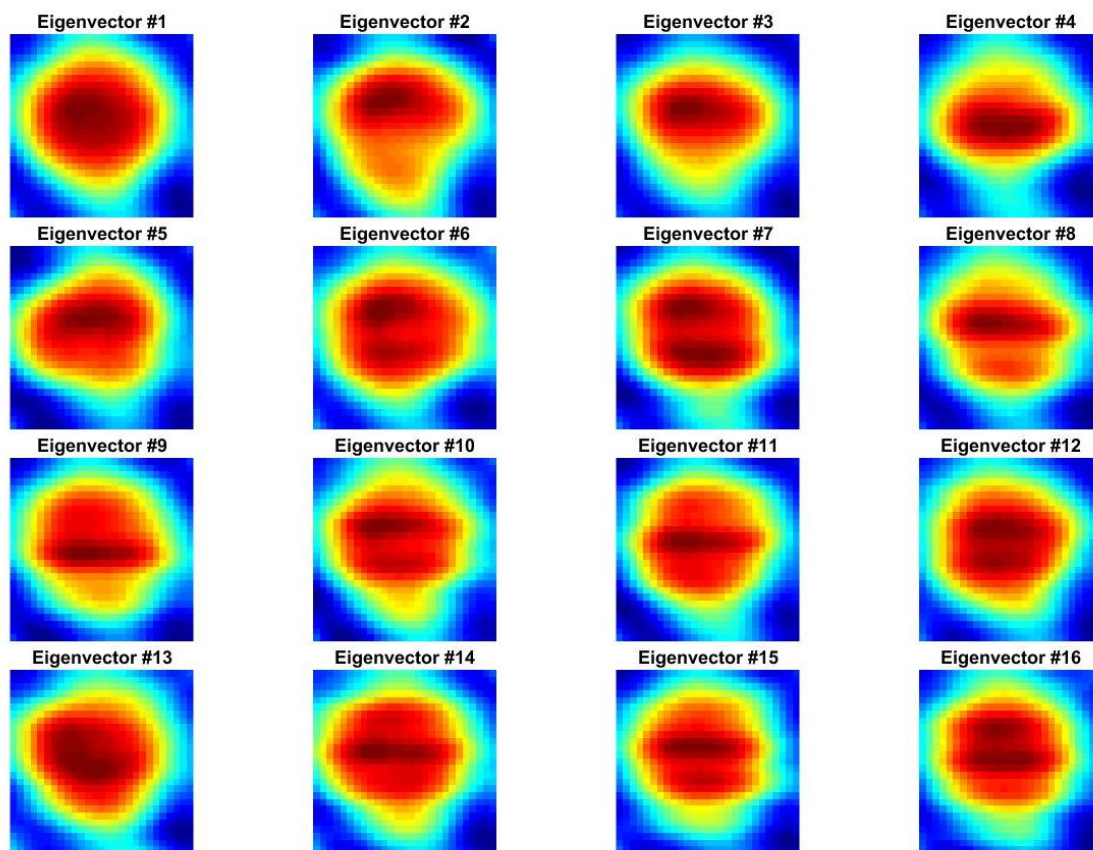
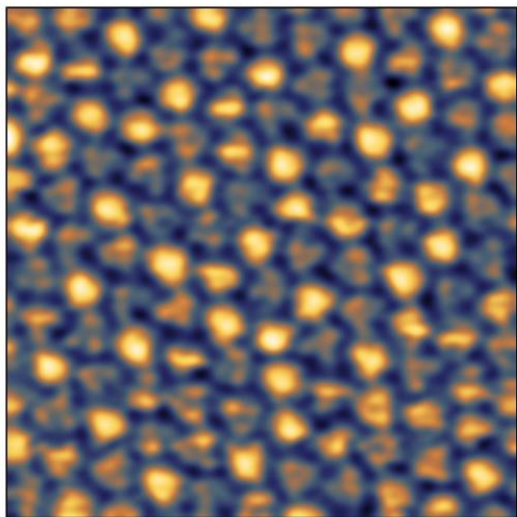


Figure 3-30. First 16 eigenvectors of principal component analysis on the dataset containing all the 938 molecules extracted from Fig. 1c.

(a)



(b)

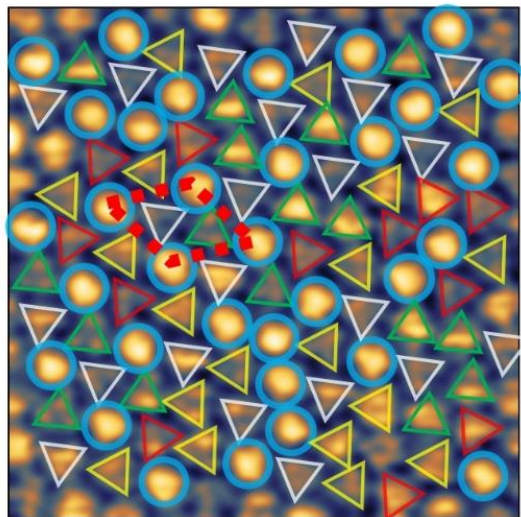


Figure 3-31. The STM image used in analysis in Fig. 1d without (a) and with (b) structure overlay.

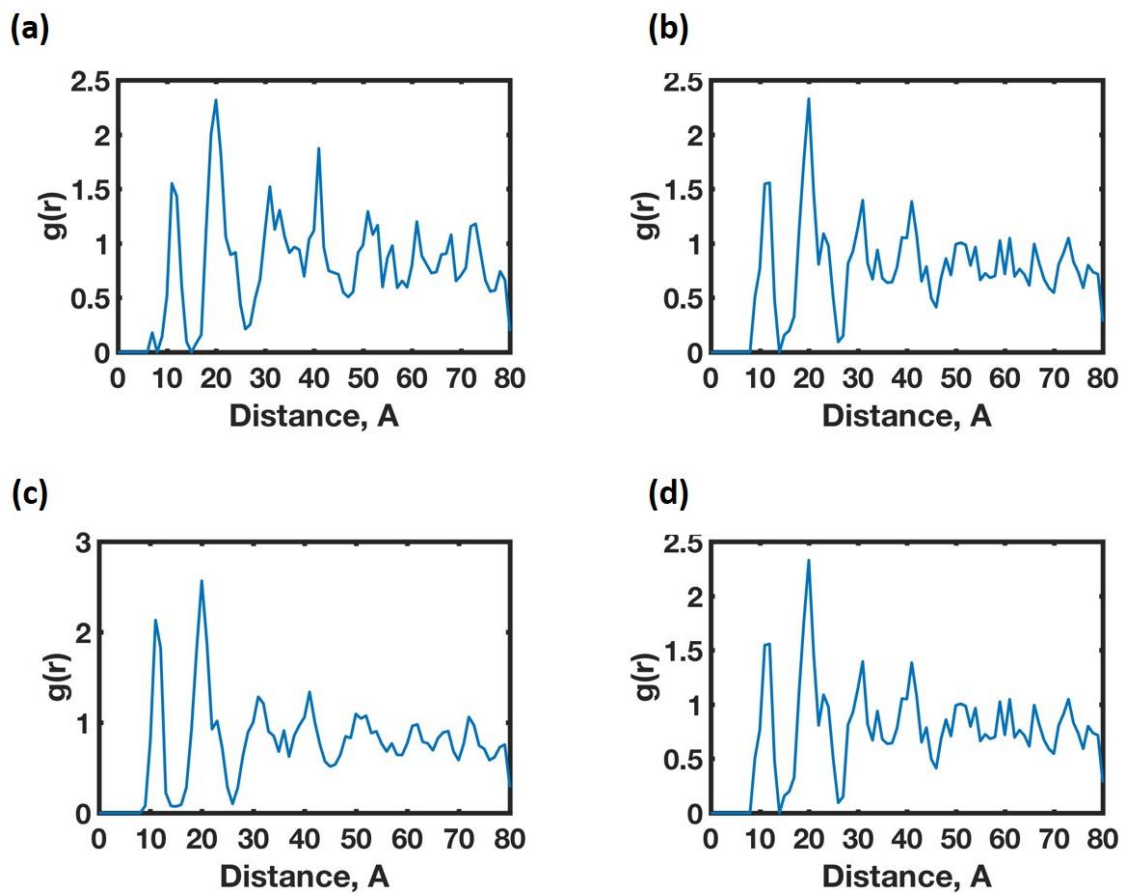


Figure 3-32. PDF functions for all considered rotational classes: (a) 1st class, (b) 2nd class, (c) 3rd class, (d) 4th class.

Table 3-2. Assignment of pairwise potentials based on our knowledge about the molecular system

	Class 1	Class 2	Class 3	Class 4
Class 1	0.01	0.33	0.33	0.33
Class 2	0.33	0.01	0.33	0.33
Class 3	0.33	0.33	0.01	0.33
Class 4	0.33	0.33	0.33	0.01

CHAPTER 4.
TOWARDS ATOMIC SCALE MANUFACTURING

4.1 Introduction

Chapter 3 demonstrated successful recovery of physically-meaningful information from experimental data and simulations in an automated, reproducible fashion. This is a breakthrough in material characterization, however, we want to go further than that and create frameworks that achieve predictive capabilities and can influence experiments, enabling flow of information from data analysis to experiment design and execution. In a way, this can be compared to going from analyzing images from cameras to creation of self-driving automobiles. The ultimate goal is to achieve ability to design and manufacture materials on the atomic scale, and to be able to do it without direct human input. At the same time, such methods must be designed with input from experts, as we cannot just discard such valuable domain knowledge.

First, I describe how to use machine learning to bring together results from an experiment in atomic scale manufacturing and a theory supporting it.

In the second part of the chapter, I describe a combination of deep learning and Bayesian methods that is created by incorporating domain knowledge into the framework and can be further applied to data without repeating the process.

More details about using electron beam for atomic scale manufacturing can be found in the article by Nan Jiang, Eva Zarkadoula, Prineha Narang, Artem Maksov, Ivan Kravchenko, Albina Borisevich, Stephen Jesse, and Sergei V Kalinin originally published in:

Jiang, Nan, Eva Zarkadoula, Prineha Narang, Artem Maksov, Ivan Kravchenko, Albina Borisevich, Stephen Jesse, and Sergei V. Kalinin. "Atom-by-atom fabrication by electron beam via induced phase transformations." *MRS Bulletin* 42, no. 9 (2017): 653-659. [309]

4.2 Direct atomic fabrication and dopant positioning in Si using electron beams with active real time image-based feedback

4.2.1. Publication Information

A version of this article was originally published by Stephen Jesse, Bethany M Hudak, Eva Zarkadoula, Jiaming Song, Artem Maksov, Miguel Fuentes-Cabrera, Panchapakesan Ganesh, Ivan Kravchenko, Paul C Snijders, Andrew R Lupini, Albina Y Borisevich and Sergei V Kalinin in:

Jesse, Stephen, Bethany M. Hudak, Eva Zarkadoula, Jiaming Song, Artem Maksov, Miguel Fuentes-Cabrera, Panchapakesan Ganesh et al. "Direct atomic fabrication and dopant positioning in Si using electron beams with active real-time image-based feedback." *Nanotechnology* 29, no. 25 (2018): 255303. [310]

The article content is reprinted as published with figure numbers modified.

4.2.2. My contribution

In order to better understand the mechanism of atomic fabrication with electron beam, it is necessary to both explore the experimental data and develop simulation methods. In this case, Si was chosen as a model semiconductor material to demonstrate how a combination of experiment and simulation can describe atomic-level manipulation and effects of the electron beam. In this project I worked with STEM images and Molecular dynamics (MD) simulations of Si crystalline/amorphous interface.

I used non-negative matrix factorization (NMF) in order to explore number of phases and their compositions in STEM images of Si before beam-induced crystallization. I generated abundance maps using non-negative least squares analysis in conjunction with extracted endmembers to demonstrate that growth region belongs to the same crystallographic phase as the substrate.

I computed tetrahedral order parameter for each atom within MD simulations and used it in conjunction with nearest neighbor identification to construct descriptor matrices to serve as inputs into k-means clustering algorithm to identify crystallographic phases in time and space. Resulting analysis also demonstrated that the growth region belongs to the same phase as the substrate.

In this case, statistical analysis was able to connect experiment and simulation on qualitative level and confirm effectiveness of the atomic fabrication with electron beams.

4.2.3. Abstract

Semiconductor fabrication is a mainstay of modern civilization, enabling the myriad applications and technologies that underpin everyday life. However, while sub-10 nanometer devices are already entering the mainstream, the end of the Moore's Law roadmap still lacks tools capable of bulk semiconductor fabrication on sub-nanometer and atomic levels, with probe-based manipulation being explored as the only known pathway. Here we demonstrate that the atomic-sized focused beam of a scanning transmission electron microscope can be used to manipulate semiconductors such as Si on the atomic level, inducing growth of crystalline Si from the amorphous phase, reentrant amorphization, milling, and dopant-front motion. These phenomena are visualized in real time with atomic resolution. We further implement active feedback control based on real-time image analytics to automatically control the e-beam motion, enabling shape control and providing a pathway for atom-by-atom correction of fabricated structures in the near future. These observations open a new epoch for atom-by-atom manufacturing in bulk, the long-held dream of nanotechnology.

4.2.4. Introduction

Research on transport phenomena in semiconductors in the late 40's at Bell Labs laid the foundation for many of the technologies that underpin modern civilization[311] and started the incessant drive for integration and miniaturization of electronic devices. Immediately after the demonstration of the solid-state transistor by Brattain, Bardeen, and Shockley[312], it was realized that the future lies in the integration of multiple devices, including transistors and memory elements, on a single crystal. While early strategies pursued the growth of compositionally-graded semiconductor crystals, it was the conceptual breakthrough by Noyes and Kirby that demonstrated the fabrication of in-plane structures in the form of the first integrated circuit, the accomplishment which landed them a Nobel Prize in 2015. Since then, the semiconductor industry has adopted a paradigm for fabrication based on a combination of 1D chemical steps (fabrication in the out of plane, or z-direction) such as oxidation, resist deposition, etching, etc. with 2D patterning steps (patterning in xy plane) using light exposure. The combination of these steps in a predefined sequence, under well-defined conditions, has enabled the modern computer-based civilization, resulting in the present sub-10 nm semiconductor structure.

The undeniable success of present day semiconductor technology is belied by significant limitations. Device processing relies on mesoscopic transport and chemical reactivity, leading to rapid growth of stochastic phenomena and noise during fabrication. Shrinking device size combined with the discrete nature of atomic dopants leads to a large spread in device performance, which can be traced to different distinct (and uncontrollable) atomic configurations .[313] Applications ranging from micro- to nanomechanical systems necessitate the assembly of complex 3D structures, rather than densely integrated layers. These limitations are well-recognized in the semiconductor industry, and the emergence of techniques such as electron beam induced depositions and lithography or focused ion milling[314, 315] is a direct response to these challenges. While lacking the parallel nature of photolithography, all of these techniques have developed into multibillion-dollar industries.

However, electron beam based techniques still lack the capability to fabricate materials down to the atomic level, and the need for such fabrication is by now well realized. In particular, quantum devices for large-scale implementation of quantum computing, single-spin magnetoelectronic devices, and scalable neuromorphic systems all require fabrication at the atomic level, including precise fabrication of crystalline layers down to single atomic planes, positioning of functional dopant atoms, and avoiding atomic-scale defects in the active region of the device and interconnects.[316, 317] In other areas, the impact of these developments can be predicted. For example, in materials science and condensed matter physics the capability to create desired atomic configurations and explore their functional properties (e.g. via local electron spectroscopies) will yield a paradigmatic shift in our understanding of the underlying principles. In other areas, atomic level fabrication can provide pathways towards large scale fabrication of materials with predefined properties – e.g. by providing seed crystals of thermodynamically metastable phases that can be further grown in macroscopic crystals.

Despite this clear incentive, the current methodology for atomic fabrication today is the combined approach based on scanning tunneling microscopy manipulation and surface chemistry, harking back to experiments by D. Eigler and advanced by J. Lyding, M. Simmons,[316-318] and commercially by companies such as Zyvex[170] and NanoFactory[171]. In this approach, the ability of the scanning tunneling microscope to manipulate single atoms is combined with the precise control of surface chemistry (silicon passivation) to position dopants at preselected locations, interface with macroscopic electrodes, and stabilize with surface passivation layers that allow taking the fabricated structure outside of the ultra-high vacuum (UHV) chamber. However, the cost and throughput of this approach remains such that research grade, several qubit devices remain the only viable target application. Hence, the question remains – are there other strategies that can potentially enable atomic scale fabrication of semiconductors avoiding the throughput bottle neck of probe based fabrication?

Here, we demonstrate a novel method for semiconductor manufacturing: the use of the atomically focused beam of a scanning transmission electron microscope to control local material structure of Si in the bulk with atomic precision. Through use of the electron beam,

we can induce a broad variety of phenomena including amorphization, crystallization, and controlled dopant atom and front motion that can be resolved *in-situ*, enabling real-time correction and editing of matter at the atomic level. As an illustration, we implemented a combination of e-beam control electronics and active machine vision based feedback to fabricate predefined crystalline Si patterns. Although recently published work has demonstrated crystallization in oxide materials,[319, 320] here we extend these techniques to the most widespread semiconductor (Si), demonstrate controlled motion of dopants, and incorporate an automatic feedback control system.

The natural limitation of STEM based fabrication is the requirement for beam-transparent systems of ~ 50 nm thickness. Correspondingly, planarization and incorporation of this system into a realistic device structure will be essential for practical application.

4.2.5. Electron matter interactions in STEM

Traditionally, STEMs have been perceived only as imaging or analysis tools, and any beam-induced modifications were considered undesirable beam damage. Yet in the last five years, it has been demonstrated that beam-induced modifications can produce appealing results, including the formation and ordering of oxygen vacancies,[321] single defect formation and motion of extended defects in 2D materials,[322] beam-induced migration of single interstitials in diamond-like lattices,[323] and formation of single chemical bonds.[324] Remarkably, these changes often involve one atom or small groups of atoms, are potentially chemically selective, and can be monitored in real-time with atomic resolution,[325] opening pathways towards control.[211, 326] This combination of atomic manipulation and (sequential) atomic-scale visualization was without precedent until we observed that the e-beam can induce the crystallization of certain amorphous materials including oxides such as SrTiO_3 , a process we refer to as e-beam sculpting.[320] Notably, the interaction between the electron beam and amorphous matter was actively explored in the 1980's and 1990's, and e-beam crystallization of a number of important semiconductors such as Si[327-330] and GaAs[330-332] has been reported. However, these experiments lacked the capability to probe beyond mesoscopic level studies, and no

attempts to actively direct and control the process were reported. Nonetheless, three key factors were established regarding beam-induced processes in semiconducting materials: there is a strong beam energy dependence controlled by the proximity to the knock-on threshold (roughly 145 keV for bulk Si); the interactions generally cannot be reduced to purely thermal processes; and under certain conditions these processes can result in both amorphization and crystallization of material. At that time, this approach did not appear to offer any significant advantages over classical semiconductor processing, and therefore was not extensively pursued.

Here, we use the atomically focused beam of a STEM operating at 200 kV to guide amorphous-crystalline transformation in Si at the atomic-plane level, including both forward and reverse transitions, and demonstrate beam-induced motion of dopant atoms that can be assembled in different configurations. As a model sample, we have chosen amorphous silicon grown on a crystalline Si substrate with Bi dopants. The STEM image prior to e-beam crystallization is shown in Figure 4-1 (a). The boundary (marked with a dotted line) between crystalline and amorphous Si is clearly visible, and dopant atoms can be seen within the amorphous Si matrix.

Figure 4-1 (b) shows the changes in the atomic structure after repeated scans of the image, with the slow scan direction perpendicular to the interface. There is a clear formation of crystalline Si extending into the amorphous region, in apparent epitaxial registry with the substrate. By increasing the beam current to 139 pA, compared to the nominal 30 pA conditions under which Figure 4-1(a-b) was obtained, we can drive the transformation from crystalline Si to amorphous Si, as shown in Figure 4-1(c-d). The increase in beam current results in the amorphization and subsequent drilling-through of crystalline Si.

We have further explored Si patterning over a range of beam current settings and scanning speeds and demonstrated the possibility of transitioning between crystallizing, amorphizing, and drilling/evaporating regimes by moving across these parameters. At 200 kV, the nominal current of the incident electron beam is 30-35 pA. At the nominal current setting and with a reduced scan speeds we are able to crystallize amorphous regions of Si. Increasing to an intermediate current of 75-80 pA and medium to high scan speed results in amorphization of crystalline Si.

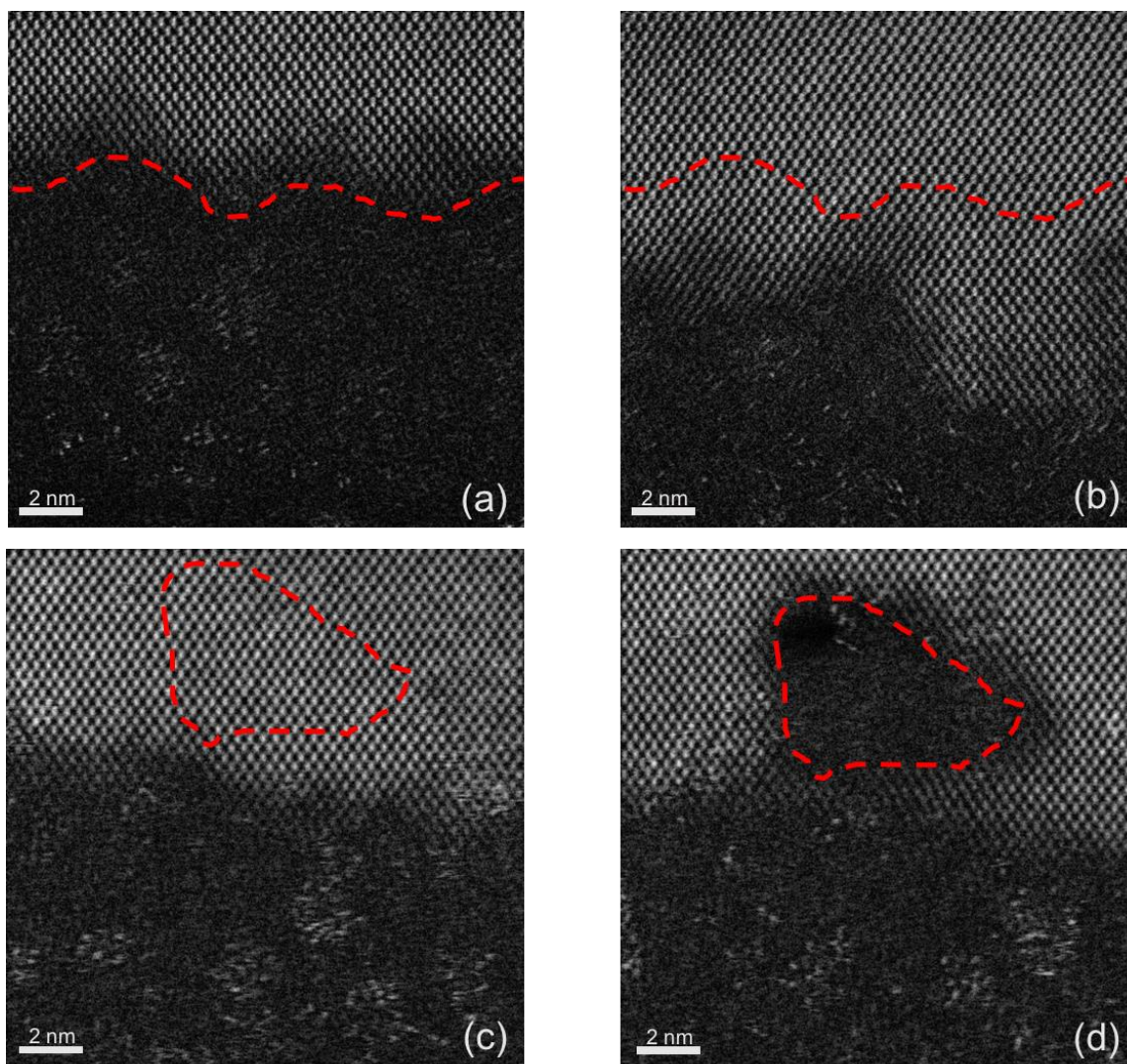


Figure 4-1. Beam induced transformations in Si using a 200 kV beam.

(a) Before growth, (b) Crystallization of amorphous Si along the crystalline-amorphous interface (30 pA beam current, 200 kV). (c, d) Amorphization and subsequent drilling through of crystalline Si (139 pA beam current, 200 kV).

A high current mode of 140 pA and low to medium-high scan speeds leads to drilling or evaporating of the material. While the exploration of quantitative mechanisms behind the observed phenomena is a separate and complex topic requiring detailed studies[333, 334], these observations clearly illustrate that both the fabrication and erasing materials regimes are open for experimental studies. Furthermore, the fact that beam-induced transitions between the phases can be reversible opens a tremendous field for further applications, from memory devices to reconfigurable electronics.

To obtain further insight into the structure of the newly formed crystalline Si, we perform comparative crystallographic image analysis[167, 182]. In this method, a sliding window is scanned across the image, generating a stack of sub-images. The relevant 2D structure factors are calculated, and the resulting data set is linearly unmixed using non-negative matrix factorization (NMF). This procedure is ideally suited for differentiation of dissimilar crystalline phases, so we can apply it here to determine if the beam-crystallized Si grows with the same crystal structure as the crystal Si substrate. Unlike methods based on direct analysis of atomic positions, this method does not require high contrast images, i.e. unmixing is possible for cases where only lowest-order reciprocal lattice peaks are visible. We start by assuming that the initial image (Figure 4-2(a)) contains only two phases and perform unmixing for two endmembers. The resulting abundance maps, along with insets showing unmixed FFT endmembers, are shown in Figure 4-2(b, c). We then use non-negative least squares analysis (NNLS) in conjunction with discovered endmembers on the after-growth image (Figure 4-2(d)). Generated abundance maps clearly show that growth regions consist of the same crystallographic phase as the substrate (Figure 4-2(e, f)).

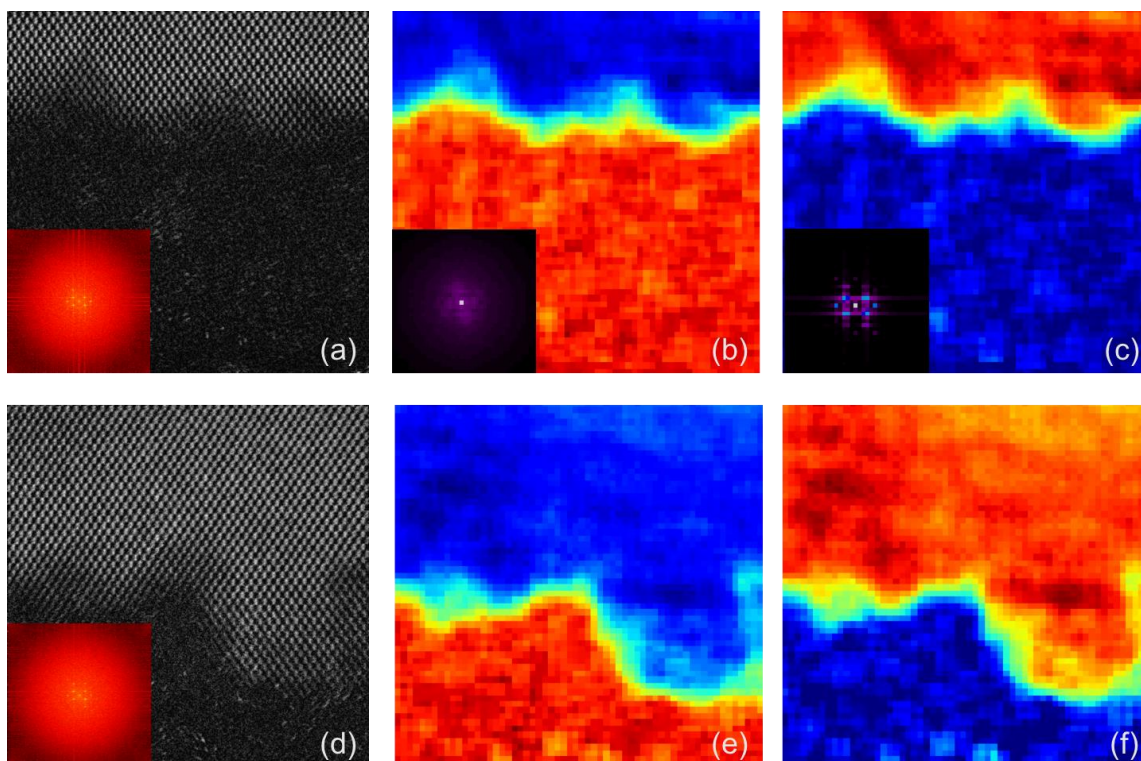


Figure 4-2. Atomic scale tracking of the local crystallinity.

(a) Pre-growth image and its FFT. (b) Abundance map and endmember 1, recovered through NMF, corresponding to the amorphous region (c) Abundance map and endmember 2, recovered through NMF, corresponding to the crystalline region. (d) After-growth image and its FFT. (e) Abundance map for the amorphous region in after-growth image generated using NNLS. (f) Abundance map for the crystalline region in after-growth image generated using NNLS.

4.2.6.. *Directed crystallization*

The combination of e-beam amorphization/crystallization and real-time observation of these transformations suggests the potential for real-time feedback, based on simultaneously monitoring the images to keep track of the beam-induced changes. The level of crystallization can be monitored and controlled to produce structures with a desired geometry. Here, a custom feedback and control system has been developed to guide the atomic layer by atomic layer movement of the crystal-amorphous (CA) interface (either as crystallization into the amorphous region or amorphization in the reverse direction). The system operates by scanning the electron beam parallel to the CA interface and simultaneously capturing the bright field and/or dark field signals of the STEM during these linear scans. A single line scan can be used to determine the degree of local crystallinity by calculating the amplitude portion of the 1D fast Fourier transform (FFT) of the STEM image signal – a line scan across an amorphous region will result in a relatively featureless FFT, whereas a line scan across a crystalline region will yield easily identifiable peaks corresponding to the average spacing between atomic columns. Controlled movement of the CA interface is achieved by using the magnitude and location of these peaks as a feedback signal to move the electron beam appropriately. That is, if one intends to advance the crystalline domain into the amorphous region then: (1) repeated identical line scans across the CA interface are performed to both induce crystallization and assess the degree of local crystallinity, (2) when the degree of crystallinity reaches a pre-determined set-point (i.e. an atomic layer of atoms has transformed from amorphous to crystalline), the line scan is advanced approximately half a unit cell into the amorphous region, and the process is continued. This process is illustrated in Figure 4-3.

A similar procedure can be used to advance the CA interface into the crystalline region for controlled amorphization as well. The differences being that beam conditions are selected to cause the crystalline portions to amorphize, and the feed-back condition becomes that advancement of the beam only occurs when the peak magnitude *drops below* a specified set-point.

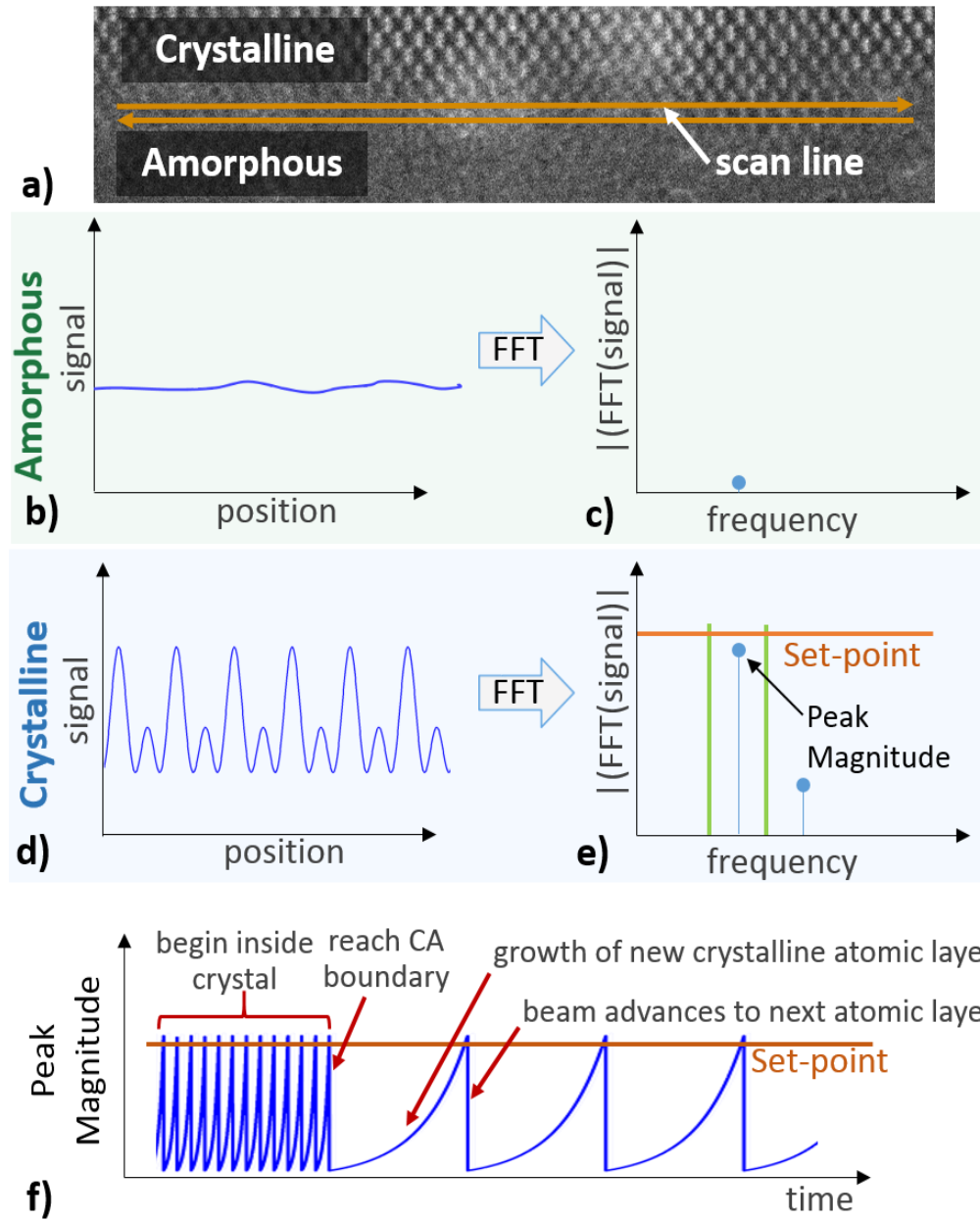


Figure 4-3. Beam directed crystallization process.

a) STEM image of CA interface and location of scan path during crystallization. Schematic depiction of the patterning process. (b) Spatial domain and (c) Fourier domain plots of detector signal in amorphous region. (d) Spatial domain and (e) Fourier domain plots of detector signal in crystalline region. The amplitude of a specific Fourier peak is tracked relative to the set-point. When amplitude exceeds the set-point, the beam is advanced $\frac{1}{2}$ a unit cell into the amorphous region. (f) Plot of the peak amplitude from (e) as a function of time during directed crystal growth from a starting point within the crystal and while the CA front proceeds forward.

4.2.7. Direct growth of doped silicon

To explore the fundamental processes during beam induced crystallization and amorphization, the experiments were repeated for Si containing dopant atoms. For most modern semiconductors, dopants are essential to obtain the desired electrical characteristics; however, direct visualization of single dopants remains a technical challenge. In these samples, a layer of Bi atoms was deposited at the CA interface. Single Bi are highly visible in STEM Z-contrast (Figure 4-4). By scanning the beam in a similar fashion as described above to induce crystal growth, it is possible to induce motion of bismuth atoms perpendicular to the fast-scan direction of the beam. This is demonstrated under various conditions in Figure 4-4. Figures 4-4(a) and (b) show, respectively, the before and after images of crystal growth where the beam-induced growth process was initiated between the CA interface and the region of high concentration dopants, propagating in [110] crystallographic direction. Figure 4-4 (b) shows the resulting crystal growth and the apparent movement of a few dopant atoms along the crystal growth direction. Note here, that at this low concentration of dopants, crystal growth appears to be unimpeded and in fact seems to allow for larger-scale structures than in the undoped case described above (perhaps due to some strain relief). For comparison, in (c) and (d) the beam induced growth is initiated deeper within the crystal than the layer with high dopant concentration, and thus a large number of dopant atoms are displaced. In this case however, crystal growth stopped progressing after several nanometers, presumably due to poor crystallographic compatibility between Bi and Si.

Further investigations of growth and dopant motion was performed by progressing in the [111] direction. Similar behavior was exhibited in (e,f) as in (c,d), namely the crystal growth and dopant motion proceeded together and abruptly stopped when the local dopant concentration reached a critical value. However, with dopant atoms cleared out of the way in (f), it was possible to induce crystal growth perpendicular to the original [111] direction as shown in (g). Figure 4-4(h) and (i) show that the same scanning pattern can also be used to move dopants deeper into the crystal. Notably, the dopant front appears very sharp in both cases, and no fronts form on the sides of the growing crystal.

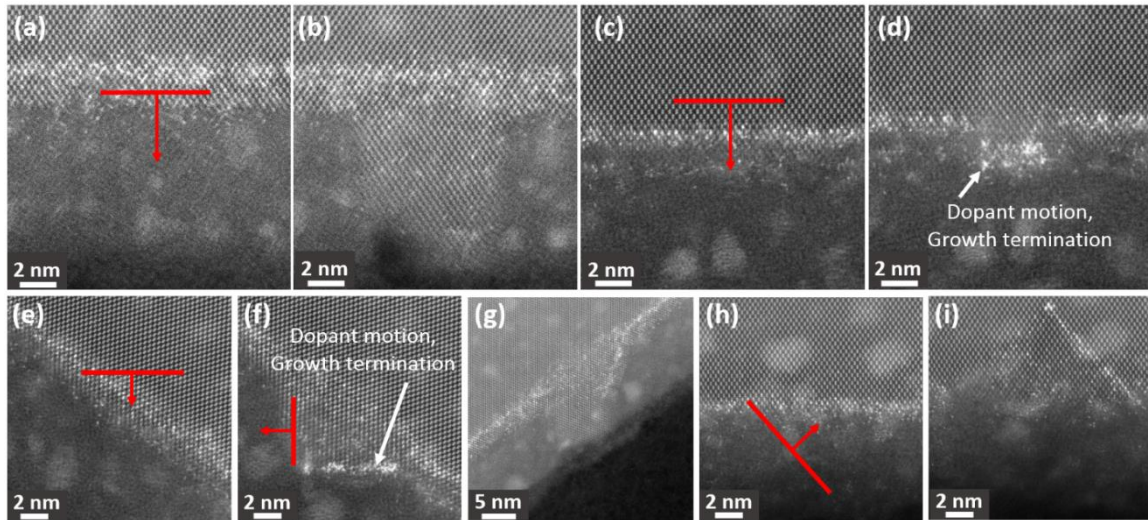


Figure 4-4. Controlled crystallization of amorphous Si and dopant movement.

Red arrows illustrate the direction of the slow scan axis for growth, while the red lines at the base indicate scan width. Panels with red arrows are the “before” images of growth, and the following panels are the “after images; (a,b) crystallization of Si from a dopant front; (c,d) dopant front movement in crystallized Si; panel (f) serves as both for the (e-g) sequence. Note that the field of view is rotated and magnification is changed between panels (f) and (g) to provide better overview of the changes.

This suggests that the hopping/relaxation time for Bi atoms after being activated (“knocked”) by the electron beam is closer to the scan time of an entire line in the fast scan direction rather than a pixel in that line. A process at this timescale can be precisely controlled by the electron beam, directly illustrating the feasibility of controlled atom-by-atom motion. Interestingly, this does not appear to be the case when the slow scan direction is along the [110] crystallographic axis. There is no apparent sharp front, and the resulting dopant profile resembles a dome more than a line. This difference in behavior for different crystallographic axes suggests that [111] is the preferential hopping direction for Bi atoms and highlights the importance of understanding the underlying mechanism and development of a predictive modeling approach for achieving reproducible results.

4.2.8. Modeling electron beam-induced transformations

To gain further insight into the observed phenomena, we consider the effects of the electron beam on the solid. Generally, the energy transfer between a high-energy particle and a solid includes two primary components: losses to the electronic subsystem and direct interactions between high-energy particles and nuclei (knock-on). The knock-on interaction can result in damage when the kinetic energy that can be transferred in a single collision is larger than the energy barrier to displace an atom in the solid. Notably, in amorphous materials the binding energies are broadly distributed, allowing for a broad distribution of knock-on thresholds. For sufficiently high particle energy, multiple event cascades can be initiated. Similarly, in a material with finite thickness, when the knock-on interaction occurs a few layers away from a material surface, ejection of surface atoms is possible. Additionally, the energy barrier will usually be significantly lower for surface atoms, primarily because of the reduced number of bonds[333, 334].

A powerful model for beam-induced changes in materials includes non-equilibrium heating, when the two subsystems – atomic nuclei and electronic – develop different temperatures, thus being in non-equilibrium conditions. Depending on the temperature difference between the two subsystems, energy that is transferred to the electrons can subsequently be transferred to the lattice atoms via the electron-phonon interactions until equilibrium is reached, where it diffuses further through the atoms. This mechanism is

described by the two-temperature (2T) model[335-338]. In the 2T model, the evolution of the electronic and the atomic temperatures are described separately, using a set of heat diffusion equations, one for the electronic and one for the atomic system. The energy exchange between the two subsystems depends on the temperature difference between them, and the strength of this interaction is expressed with the electron-phonon coupling parameter g . For the case of silicon, a combination of results from irradiation experiments³⁶ combined with the inelastic thermal spike model[339, 340], molecular dynamics and DFT computational[341, 342] and model and numerical approaches[342, 343] determine the values of the 2T model parameters. From this, g is calculated to be $1.8 - 5 \times 10^2$ W/cm³/K[339, 340, 343], using the known values for the lattice specific heat and conductivity[344], and the electronic specific heat and diffusivity[342].

Given the uncertainties in these parameters for amorphous solid, here we chose to model the induced crystallization assuming that the electron beam creates a local temperature within a small volume of material. To explore this behavior, we created a model where amorphous Si was deposited on crystalline silicon. The amorphous region was then heated to induced crystallization. The heated region ($20 \text{ \AA} \times 10 \text{ \AA} \times 108 \text{ \AA}$), representing the local volume heated by the beam, was initiated at the CA interface, and slowly moved into the amorphous region. Once the temperature inside the block reached 1300 K, crystallization began in regions close to the interface and moved upwards, terminating approximately at the [111] face, resulting in a pyramidal-like front. After 1 ns, the “beam” was moved $5 \text{ \AA}$ further into the amorphous region, and a block of the same size was again heated to 1300 K. The rest of the system, which now includes half of the previously crystallized block, was kept at 300K. This process was repeated several times until the crystal front reach about half the size of the amorphous sample ($50 \text{ \AA}$).

In order to differentiate between crystalline and amorphous phases we use the tetrahedral order parameter[345] to describe coordination state of each atom at 20 ps intervals according to the formula

$$q_t = 1 - \frac{3}{8} \sum_{i < j=1}^4 (\cos\theta_{ij} + \frac{1}{3})^2, \quad (4.1)$$

where θ_{ij} is the angle between an atom and its two nearest neighbors. The resulting parameter is in the range between zero, indicating an amorphous phase, and one, indicating a crystalline phase. However, within the amorphous phase we also observe multiple small momentarily crystalline regions. For each analyzed frame, we construct the matrix of tetrahedral order parameters for each atom and its corresponding 12 nearest neighbors. We use a k-means clustering algorithm[346] on the first frame to train the classifier, and use it to predict phases in the subsequent simulation frames. Figure 4-5 demonstrates the application of this algorithm to the simulation data, showing only atoms belonging to a crystal. This corroborates evidence from the experiment that the growth belongs to the same crystalline phase as the substrate, since we use pre-growth data as a training set, and the growth is classified by the unsupervised algorithm as the same crystal as the original substrate.

We observe that crystallization does occur mostly inside the heated region, forming characteristic pyramidal growth pattern in the beginning, and becoming slightly wider at the top of the growth region later (Figure 4-5 insets). Along the length of the crystallization front, we observe a characteristic wave-like pattern, which gets amplified as the heated region moves further away from the substrate.

The simulation, despite being simplified, reproduces much of the experimental behaviors, including the tendency to form triangles (pyramids) terminating at [111] planes. This model further offers evidence of roughening, with the brightness of the atomic columns decreasing quickly the further we get from the original boundary, observed both in experiment and theory. Further model development necessitates inclusion of realistic time effects, since presently the time scale is $\sim$ ns, as compared to experimental 10's of seconds. This behavior can be linked to a higher heating rate in MD compared to the experiment, and also allows to compensate for mismatch in timescales. Secondly, we aim to include the contribution of knock-on effects. However, even this simple model provides insight into the morphologies of the experimentally grown structures. The model system was heated up to higher temperature for a short time and produced results indicating crystallization, which hints at the key role of temperature in experiment.

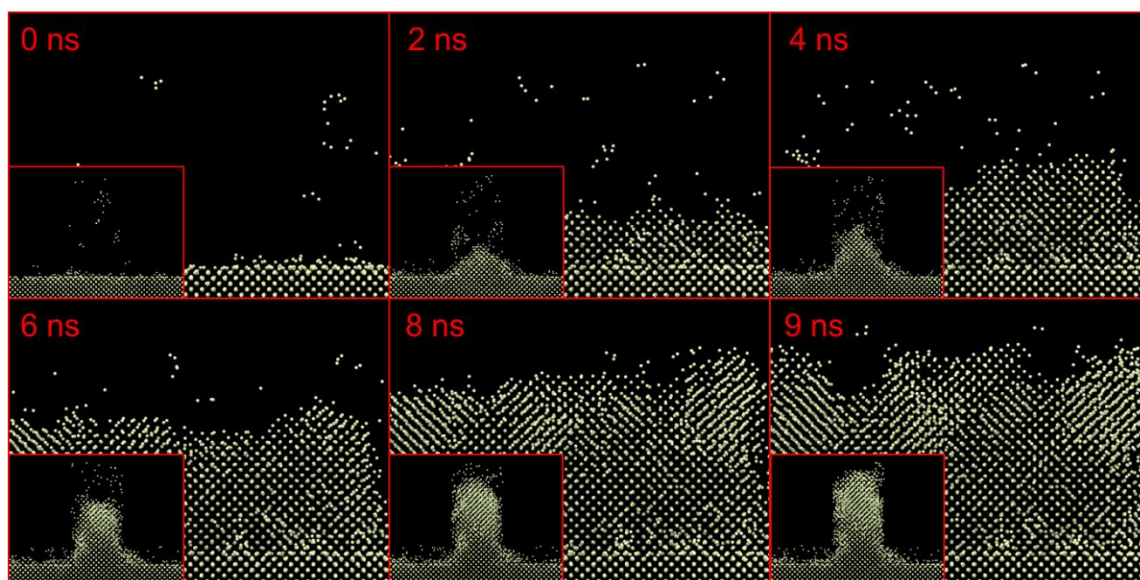


Figure 4-5. Timeline showing views of crystallization front in molecular dynamic simulation of Si crystallization. Main frame shows the side view, inset shows the front view.

4.2.9. Perspectives

Since the early days of nanotechnology revolution, the development of realistic pathways for atom-by-atom fabrication was seen as the key and enabling step to bring its promise into reality. This requirement is most acutely felt now, with the industry pace given by Moore's Law getting to the single-digit nanometer device size, and with new devices based on behaviors of a single atom, such as for quantum computing, rising to the forefront of research and development. The atomic manipulation of Si, the most important industrial semiconductor, demonstrated here, marks a key step in this direction. Remarkably, the capability of the electron beam to crystallize, amorphize, remove material, and controllably move dopant atoms fronts, even under the limitations of microscopes primarily designed to image materials rather manipulate matter, suggests that there may exist enormous potential to shape and direct matter on the atomic level.

While predicting all the opportunities enabled by the potential of STEM to manipulate and control matter at the atomic level will be complex, here we comment on the likely pathways for the development of the field in the next several years. First and foremost, the real-time feedback system implemented here can be expanded to include more complex forms of image analytics, e.g. switching between "modification" and imaging modes. Here, the use of compressed sensing[347] and related approaches could be instrumental in disambiguating low-dose non-invasive and high-dose modification regimes. Secondly, using a full 2D readout from a fast Ronchigram detector instead of simply a HAADF intensity reading can provide a feedback signal that can be used to determine when a desired transformation has occurred while the beams remains at a single location. Third, further development of precise control systems that are capable of high-speed and high-veracity beam positioning by compensating for beam scanning non-idealities (such as phase lag and frequency dependent gains) will be required.

These studies also call for extensive theoretical exploration of beam-induced effects in solids on the atomic level, at time scales spanning ultrafast electron transit times to the seconds of the induced structural relaxations. Evolution of electronic, lattice, and concentration fields and their interdependence need to be considered in detail. We believe

that the results shown above represent an important step towards full experimental control and theoretical understanding of the process.

4.2.10. Materials and methods

STEM experiment:

A Nion UltraSTEM aberration-corrected scanning transmission electron microscope, operating at 200 kV was used in this work. The nominal convergence angle was 30 mrad. High angle annular dark field (HAADF) images were acquired on a detector with an inner angle of 63 mrad. The nominal probe current is 30 pA, and the current was varied during experiments to direct crystallization and amorphization.

Modelling:

A crystalline Si sample with 64000 atoms was relaxed at 0 K using the Stillinger-Weber potential given in Sastry et al.[348] Heating up this system to a temperature of 2000 K at constant volume using the canonical, NVT, ensemble and the Nosé-Hoover thermostat, created an amorphous sample with 64000 atoms. (NVT was chosen here so as to ensure that both the amorphous and the crystal sample upon which the former will be deposited, see below, had the same dimensions.) The pair-distribution function confirmed amorphization. The amorphous sample was then relaxed at 0 K and subsequently put on top of the crystalline sample. The amorphous and crystalline samples were brought together at an initial distance of about 2 Å, and minimized, while the bottom 3 layers of the crystalline sample, i.e. those layers further apart from the interface, were kept immobile. Subsequently, except for the bottom three layers, the whole system was heated up to 300K and let to equilibrate for 500ps at this temperature using the Nosé-Hoover thermostat. At this point, a block of 20 Å width, 10 Å height, and 108 Å length of the amorphous sample, which was placed right at the interface, was heated up to temperatures high enough for crystallization to take place during a period of 1ns with a timestep of 1fs; the rest of the system was kept at 300 K (except for the bottom 3 layers of the crystalline sample, which remained immobile). Raising the temperature of the block is taken here to mimic the effect

that a beam focused in the block might produce. Once crystallization was observed, the heated region was moved up by 5 Å and the heating process was repeated until the crystallization front reached approximately half the height of the amorphous sample. All the calculations were performed with the software LAMMPS[349, 350] and run in the supercomputer *Titan*.

Amorphous Si growth:

Two sets of samples were used in this work. For the first experiment, the amorphous silicon layer was deposited on a single crystal silicon wafer shortly after its surface was RCA cleaned. The amorphous silicon deposition process is described elsewhere.[351] For the second, the Si substrates were prepared in ultrahigh vacuum with a base pressure of 4.5×10^{-11} Torr. To prepare a Si(100)-(2×1) reconstruction and atomically flat morphology, the substrates were degassed and flash-annealed according to commonly established procedures. The surface quality was examined by STM and LEED. A 12 nm thick amorphous Si film was deposited onto these Si(100) substrates at room temperature using an e-beam evaporator in ultrahigh vacuum. The sample was exposed to the ambient conditions before STEM sample preparation. The Bi-doped Si heterostructure was grown in multiple steps to reduce the Bi segregation from the Si. Specifically, we first synthesized Bi nanolines according to Ref. [[352]] by evaporating Bi from an effusion held cell at 485 °C onto a (2×1)-reconstructed Si(100) substrate held at 570 °C. A thin crystalline Si layer was subsequently grown by solid phase epitaxy: a thin amorphous Si film was deposited at room temperature and then annealed at 434 °C for 5 s during which it crystallized. Subsequently, a 24 nm thick amorphous Si was deposited at room temperature.

Samples suitable for electron microscopy were prepared from the Si material by cutting the as-grown samples in half and gluing the halves together to produce a sandwich. These are then sliced into thin sections that are glued onto support grids and mechanically polished to approximately 10 microns thickness. Ion-milling is used to produce to electron transparent samples, which are then baked at a nominal 140 °C in vacuum overnight to reduce surface contamination immediately before imaging.

4.3. Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS₂

4.3.1 Publication information

The original article was submitted for review to Nature Communications.

It was authored by Artem Maksov, Ondrej Dyck, Kai Wang, Kai Xiao, David B. Geohegan, Bobby G. Sumpter, Rama K. Vasudevan, Stephen Jesse, Sergei V. Kalinin, and Maxim Ziatdinov, and is available as preprint at:

Maksov, Artem, Ondrej Dyck, Kai Wang, Kai Xiao, David B. Geohegan, Bobby G. Sumpter, Rama K. Vasudevan, Stephen Jesse, Sergei V. Kalinin, and Maxim Ziatdinov. "Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS₂." *arXiv preprint arXiv:1803.05381* (2018). [353]

The article is reproduced as in the preprint, with figure numbers and citations altered to fit the numbering of the dissertation. Supplementary material added at the end of the chapter as an appendix.

4.3.2 My contribution

Atomic-scale manufacturing will require utilizing automation. In this case, the goal was to study the effects of electron beam irradiation on 2D WS₂ material, imaged by STEM as movie data. I performed FFT subtraction procedure to study the applicability of the method to large amounts of data. I found out that method requires manual tuning and is poorly generalizable to large datasets such as movies. In order to address this issue, I trained a custom deep fully-convolutional neural network on a set generated by performing FFT subtraction and data augmentation on the first image of a movie. The trained network was able to identify locations of defects in the rest of the frames.

I further utilized a Gaussian mixture model (GMM) to explore the possible state space of the identified defects and achieved classification that produced physically meaningful results.

Knowing the positions and the classes of defects, I was able to study their evolution in time as a Markov process, assigning probabilities for class transitions. Finally, I explored applicability of class activation maps based deep learning techniques to analysis of atomic images.

4.3.3. Abstract

Understanding elementary mechanisms behind solid-state phase transformations and reactions is the key to optimizing desired functional properties of many technologically relevant materials. Recent advances in scanning transmission electron microscopy (STEM) allow the real-time visualization of solid-state transformations in materials, including those induced by an electron beam and temperature, with atomic resolution. However, despite the ever-expanding capabilities for high-resolution data acquisition, the inferred information about kinetics and thermodynamics of the process and single defect dynamics and interactions is minima, due to the inherent limitations of manual *ex-situ* analysis of the collected volumes of data. To circumvent this problem, we developed a deep learning framework for dynamic STEM imaging that is trained to find the structures (defects) that break a crystal lattice periodicity and apply it for mapping solid state reactions and transformations in layered WS₂ doped with Mo. This framework allows extracting thousands of lattice defects from raw STEM data (single images and movies) in a matter of seconds, which are then classified into different categories using unsupervised clustering methods. We further expanded our framework to extract parameters of diffusion for the sulfur vacancies and analyzed transition probabilities associated with switching between different configurations of defect complexes consisting of Mo dopant and sulfur vacancy, providing insight into point defect dynamics and reactions. This approach is universal and its application to beam induced reactions allows mapping chemical transformation pathways in solids at the atomic level.

4.3.4. Introduction

Chemical reactions and phase transformations underpin phenomena ranging from cosmological processes, to the emergence of life on Earth, to modern technological progress, and are therefore of tremendous interest for both basic and applied sciences. Simple gas phase reactions of small molecules can be readily studied using well-established spectroscopy methods (infrared,[354, 355] mass,[356] NMR[357]), utilizing spatial homogeneity of reaction volumes when the same process occurs multiple times. In conjunction with first-principles-based modelling,[358, 359] a reliable picture of molecular reactivity is being built. For studies of more complex organic and biochemical reactions, time-resolved cryogenic microscopy[360] and femtosecond x-ray pump-probes[361] provide a reliable investigative framework, again relying on the statistical similarity between multiple orientations of the same molecule.

The situation is far more complicated for solid state reactions involving continuous solids. Traditionally, solid-state phase transformations and reactions were explored by bulk measurements and x-ray techniques. However, such techniques may not be able to provide sufficient spatial resolution for understanding elementary mechanisms behind the observed transformations. This problem can be partially solved by direct *ex-situ* visualization of reaction zones,[362, 363] providing information on the geometry and, in certain cases, atomic configurations at the reaction fronts. Similarly, utilization of colloid models[364] allows for the development of model systems, albeit the nature of local interactions is significantly different from atomic systems.

In recent years, the advances in scanning transmission electron microscopy ((S)TEM) have enabled the direct visualization of dynamic phenomena at the atomic level.[209, 309, 365-371] The physical/chemical phenomena studied by in-situ STEM are wide ranging and now include e-beam induced defect evolution,[207, 208, 372-379] dislocation migration,[380-382] observation of e-beam induced production of single layer Fe and ZnO membranes in graphene nanopores,[383, 384] e-beam induced chemical etching and growth from nanoparticle and single atom catalysts,[385-390] sub-10 nm scale lithography,[391] graphene healing,[392] conductive nanowire formation,[393]

crystallization and amorphization at 2D interfaces,[205, 394] formation of fullerenes,[395] and graphene edge dynamics.[396, 397] This list can hardly be considered comprehensive but it serves to illustrate well the vast array of dynamic changes that are being observed and rapidly explored via in situ (S)TEM techniques. A tantalizing development which was published just last year (2017) is the introduction of a single dopant atom into a graphene lattice, the controlled movement of the atom through the lattice, and the assembly of a few primitive structures atom-by-atom.[210, 244, 246, 398] Such efforts harken back to the work of Don Eigler who first demonstrated controlled atomic motion via scanned probe techniques.[399] However, given the colorful array of other atomic, chemical process observed in (S)TEM and the continuously growing portfolio of commercially available in-situ equipment (heating, electrical biasing, gas and liquid cells etc.), it seems likely many more processes can be brought under our direct control, turning the (S)TEM into an atomic scale fabrication platform.[400]

Successes in e-beam atom-by-atom fabrication and atom-by-atom mapping of solid state reactions will not only require explorational research and instrumental improvements. The key piece of the puzzle will involve successfully grappling with the enormous amount of data which can be generated by these machines to infer material specific information describing kinetics and thermodynamics of point and extended defects, reaction paths, and mechanisms for extended defect and second phase nucleation and growth. The “by hand” analysis of years past is no longer a tractable solution considering the dimensionality and number of datasets which are now routinely obtained. This necessitates the creation of methods which allow for automated analysis of dynamic transformations to extract relevant materials descriptors and reconstruct reaction pathways from various sources of detector readouts, such as the variety imaging and spectroscopic modes. In this article we attempt to forge an inroad in one aspect of this challenge, namely automated image analysis for the detection and tracking of defects in STEM video of 2D materials and further proceed to extract microscopic point defect reaction mechanisms from these observations.

Here, we analyze the phase evolution of Mo-doped WS₂ during electron beam irradiation. In this process, the electron irradiation results in removal of the sulfur atoms, rendering the system oversaturated with respect to low-valence tungsten-sulfur moieties. We develop a

deep learning network for rapid analysis of this dynamic data, analyze transformation pathways, create a library of defects, and explore minute distortions in local atomic environment around the defects of interest, ultimately building a complete framework for exploring point defect dynamics and reactions.

4.3.5. Results and Discussion

Figure 4-6 shows several selected frames from the STEM “movie” of lattice transformations in the Mo-doped WS₂ monolayer under 100 kV electron beam irradiation. The full “movie” is available in the Supplementary Material. It can be clearly seen that the system evolves with time, evolving numerous point defects. On accumulation of non-stoichiometry, the latter start to segregate, forming extended defect, nucleating secondary phases, and resulting in the segmentation and rearrangement of the 2D layer. The key task is get out the information of interest about the defects. Unfortunately, most of the methods for localizing and identifying/classifying defects available to date are slow, inefficient and require frequent manual inputs.

To overcome the limitations of the available approaches, we developed physics-based machine learning method for localizing and identifying defects. We exploit the fact that each defect is associated with violation of ideal periodicity of the lattice. Therefore, we train a convolutional neural network (cNN) using a *single* image at the early stage of the beam-induced transformation, when macroscopic periodicity is still maintained, and each defect can be readily discovered providing the “ground truth” for network training. Thus, a trained network relies only on the local characteristics of the image, and hence can identify defects on the later stages of system evolution when the long range periodicity of the lattice is broken due to a second phase evolution and displacement and rotation of unreacted WS₂ fragments. Furthermore, we find that the network can discover via “extrapolation” other defects which may not necessarily be a part of the initial training set. Such “extrapolation” is possible due to generalization abilities of deep learning models. Indeed, we have recently demonstrated[194] that a deep cNN trained on the simulated

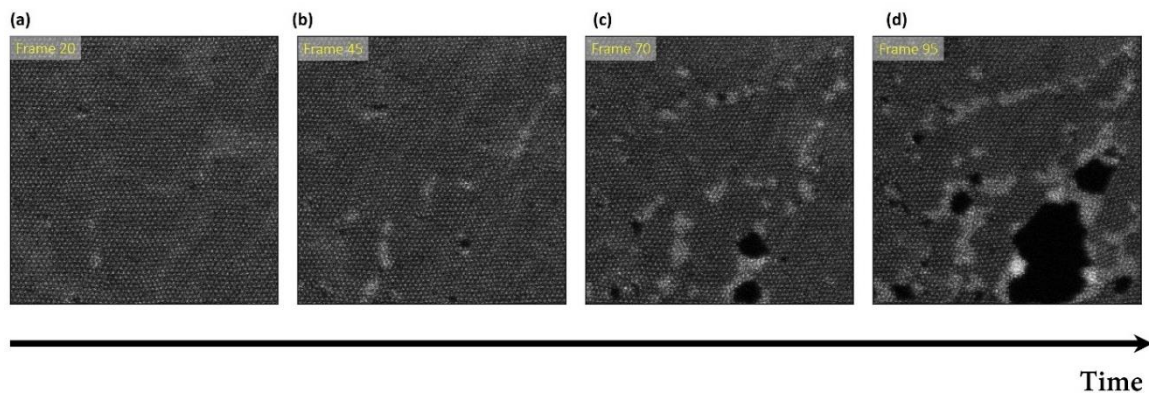


Figure 4-6. Defects evolution under e-beam irradiation in Mo-doped WS₂.

(a-d) Four selected frames from the STEM “movie” of Mo-doped WS₂ obtained at 100 kV illustrating formation of defects and lattice transformations as a function of time. The full movie can be found in the supplementary material.

images of an idealistic lattice vacancy structure can in principle generalize well enough to detect larger and more complex lattice vacancy structures in the system (*e.g.*, double and triple vacancies, as well as reconstructed vacancies). The extracted defect structures can be identified/classified using unsupervised clustering and unmixing techniques. Finally, the selected defects can be studied further using local crystallography techniques, [78] such as a combination of atom finder and principal component analysis for analyzing minute atomic distortions in their vicinity in the “movies”, as well as with a Markov analysis for identifying transition probabilities between different defect configurations.

As a first step of analysis, we define the topology of a neural network to target specific physics of beam-induced transformations. The network must be able i) to separate atomic scale lattice disorder from the rest of the lattice, ii) to return the precise location of the detected defects, iii) to be able to generalize to previously unseen defect structures. One possible candidate is the class activation maps based deep learning analysis, in which a model trained on image-level labels is capable in principle of discriminating the image regions used to identify the specific class [247] (defect). The disadvantage of such approach is that one must start with manually selecting the isolated single defect structures to create a training set. In addition, we found that while this approach allows certain atomic defect structures to be located with a sufficient accuracy, it does show relatively poor generalization ability. The alternative approach is to use a fully convolutional neural network model, [401] which can be trained to output a pixel-wise classification map, with the same size/resolution as the original input image, that shows a probability of each pixel belonging to certain type of object (defect). This type of model has been recently successfully applied to finding lattice atoms in raw STEM data [194] and we therefore chose it for the current problem.

The next task is to create a training set that will be used to “teach” a model to find lattice defects in STEM “movies”, allowing for sufficient flexibility to discover all the defects but at the same time avoiding over-classification for classes that physically cannot be present in the data. We found that it is possible to train a network using only the first frame of such a “movie” or a single image obtained before recording a movie, and then let the trained network analyze the remaining part of the movie. This approach utilizes the fact that in

macroscopically (i.e. on the length scale of the image) system the defects can be trivially discovered via the Fourier method, [402] providing the ground truth for network training. However, when trained, the network relies solely on local edge properties for identification and is thus stable towards formation of extended defects, rotations, and fragmentations of the lattice.

To identify the defects, we select a single image (frame) at the beginning of transformations (Fig. 4-7a). Once the image is selected, we perform global Fast Fourier Transform (FFT) on the selected experimental image and apply a high pass filter in reciprocal space in order to remove non-periodic components of the lattice (Fig. 4-7b). We then perform inverse FFT to obtain periodic image and subtract the original image from it such that only the deviations from the ideal periodic lattice remain.[402] In this image difference, vacancies show up as bright spots. Next, the image difference is thresholded to find locations of the single defects (Fig. 4-7c). Note that the thresholded images represent ‘ground truth’ which will be used to train a convolutional neural network. The training set is created by performing data augmentation of the selected experimental images and the corresponding ground truth images. This augmented dataset can be used to train a neural network to return positions of atomic lattice disorder from raw experimental data (Fig. 4-7d). Once trained, not only this cNN-based method for finding defects is faster and more efficient than the method based on FFT subtraction, but it also allows, unlike the FFT method, to find position of defects in the images of fragmented atomic lattice where multiple (joint and/or disjoint) lattice domains can be rotated by different angles with respect to each other. Because our model allows finding defects that break lattice periodicity irrespectively of the exact type of the defect, we consider it to be a ‘universal’ defect finder for a given material. We now use the cNN model trained according to the method described above to locate atomic defects in dynamic STEM data on Mo-doped WS₂. Interestingly, although our model was trained using only the 1st frame (out of 100) of the movie, it was able to accurately identify the positions of atomic defects in the remaining 99 frames (see Figure 4-12; full movie of the defects found can be downloaded from Supplementary Material). Once a sufficient number of defects ($\sim 10^4$ in this case) is extracted via the cNN model it becomes possible to categorize them into different classes.

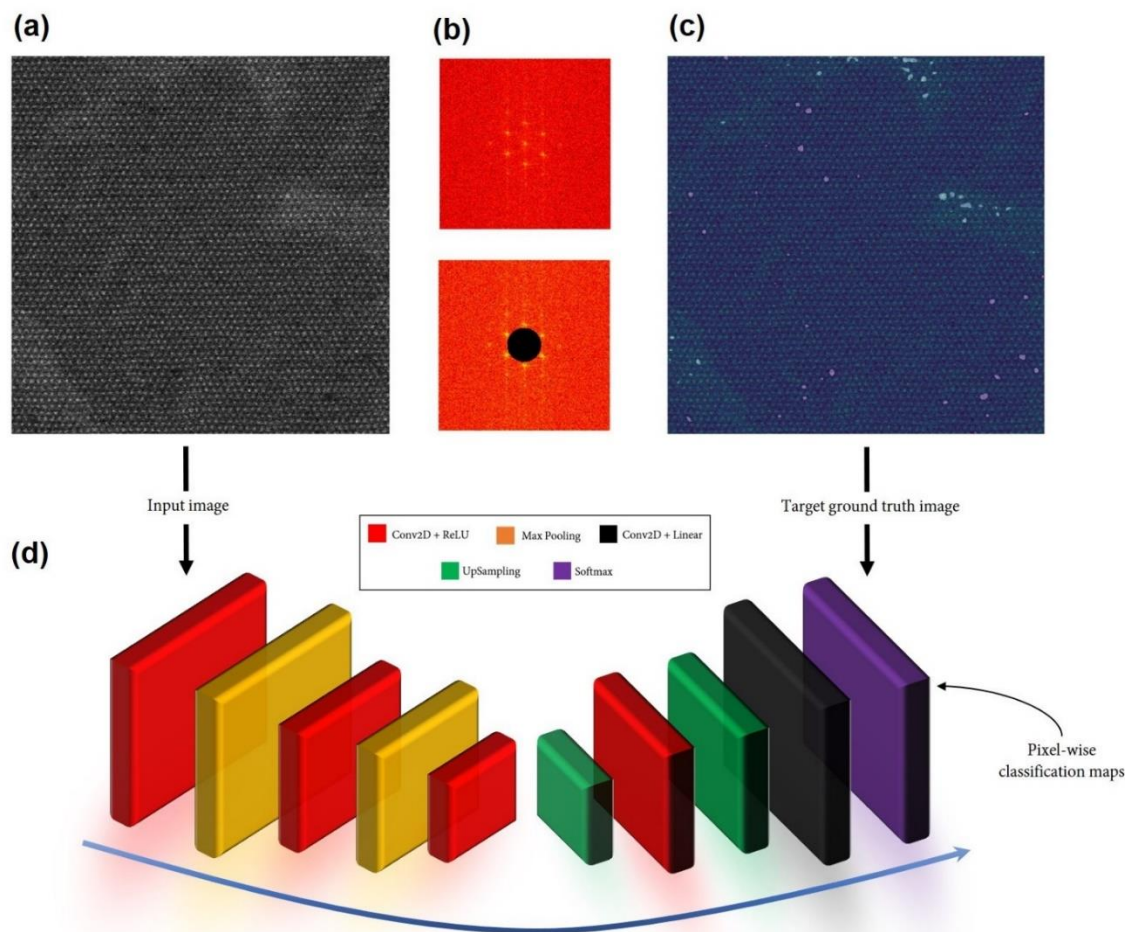


Figure 4-7. Locating atomic defects in dynamic STEM data using a cNN model trained on a single STEM movie frame.

(a) Experimental STEM images of Mo-doped WS_2 . (b) Global FFT of the image in (a) and (c) global FFT with high-pass filter applied. (d) Binary masks for image differences between the original data in (a) and inverse FFT of data in (c). The image in (a) is a training image and the data in (d) serves as 'ground truth' (pixel-wise labeling). (e) Schematics of cNN with an encoder-decoder type of structure.

To perform such a defect classification in unsupervised fashion, we adapted a Gaussian mixture model (GMM). [403] The GMM is applied to a stack of defect “windows” (images of the identical size, usually 32px*32px cropped around the center of each defect) extracted using the pixel-wise classification maps in the cNN output. Here, we chose the number of GMM components to be five as it appears to be an optimal number of components for understanding the type of defect structures present in the data. Indeed, an increase in the number of components resulted in fine (sub-) structures of the detected defects, while decrease in the number of components produced some physically meaningless structures (see Figure 4-13). We also note that the number of components past purely exploratory stage can be adjusted based on known defect chemistry of material (either from common physics principles, DFT calculations, or combinatorial analysis)

The defect structures associated with the unmixed components of GMM are shown in Fig. 4-8a-e. The class 1 and class 3 (Fig. 4-8a,c) were found to correspond to a substitutional atom in W sublattice with a lower Z number, which we interpret as Mo dopant (**Mo_w**). Note that Mo atom does not occupy a symmetric central spot in these structures as one would expect for a lone Mo dopant. This suggests that there are additional distortions present in the defects that form classes 1 and 3, likely associated with a disorder in S sublattice. Interestingly, presence of a coupling between distinct defect species has been recently observed in static STEM images from Mo-doped WS₂ system and attributed to merging of defects during growth and post-growth procedures.[404] This comparison illustrates that that as in other cases, systematic application of statistical and machine learning methods allows both to recover earlier observations and, as we show next, derive new information about underlying physical and chemical processes. The class 4 and class 5 (Fig. 4-8d, e) are associated with vacancy in W sublattice (**V_w**) and in S sublattice (**V_s**), respectively. Presence of adatoms / “contaminations” created during the e-beam surface transformations (e.g. chemical species from initial WS₂ material deposited back on to the surface in combination with carbon atoms) can explain a defect structure from the class 2 (Fig. 4-8b).

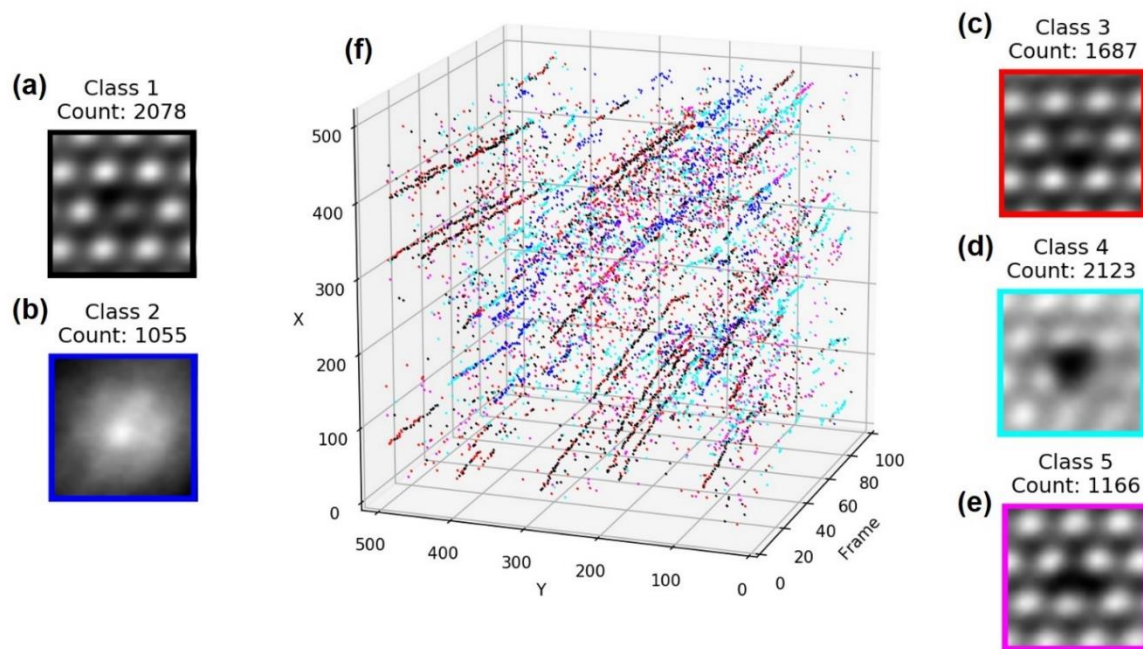


Figure 4-8. Unsupervised classification of defects located by the deep convolutional neural network.

(a-e) Results of GMM unmixing into 5 components (classes). (f) Spatio-temporal trajectories of the detected defects. Color scheme is the same as in (a).

Figure 4-8f shows spatio-temporal trajectories (‘brush diagram’ [405]) for the identified defects. Based on the analysis of the diagram, we identify three characteristic statistical behaviors: weakly moving trajectories, stronger diffusion, and “uncorrelated events” / “flickering”. Presence of more than one characteristic behavior of the atomic defects may be potentially connected to complex spatial character of strain fields during the material transformation, which may impact diffusion properties as well as create certain “localization regions”, in which the motion of defects is restrained. [406, 407] In the following we will focus on the analysis of the continuous and quasi-continuous trajectories only. The most well-defined trajectories are associated with Mo dopants (class 1 and class 3). These Mo defects show different diffusion behaviors depending on their location in the lattice and are characterized by reversible switching between two configurations (class 1 and class 3) along their trajectories. The defects associated with S and W vacancies typically form shorter (compared to Mo defects) trajectories. One possible explanation is that these vacancies are getting filled by the W and S species from the extended clusters of the deposited WS₂ material (although we did not find any associated correlations with point defects of class 2).

We now demonstrate that, based on the results produced by a combination of cNN and GMM, it becomes possible to estimate diffusion characteristics of the selected defect species. Particularly, we studied diffusion properties of S vacancies. We first collapse the 3-d spatiotemporal diagram for a chosen class of defect into the 2-d representation. For this purpose, we project the ‘windows’ of specific class of defect, which allows to separate defects that are continuous in time from the randomly occurring ones (see Figure 4-14). This analysis is complemented by a density-based clustering algorithm, [408] which yields similar results. After extracting defect coordinates for each selected defect “flow” (Fig. 4-9), we can obtain variance of each distribution and estimate a diffusion coefficient within a framework of a random walk model in two dimensions. This yields values of diffusion coefficient between $3 \times 10^{-4} \text{ nm}^2/\text{s}$ and $6 \times 10^{-4} \text{ nm}^2/\text{s}$.

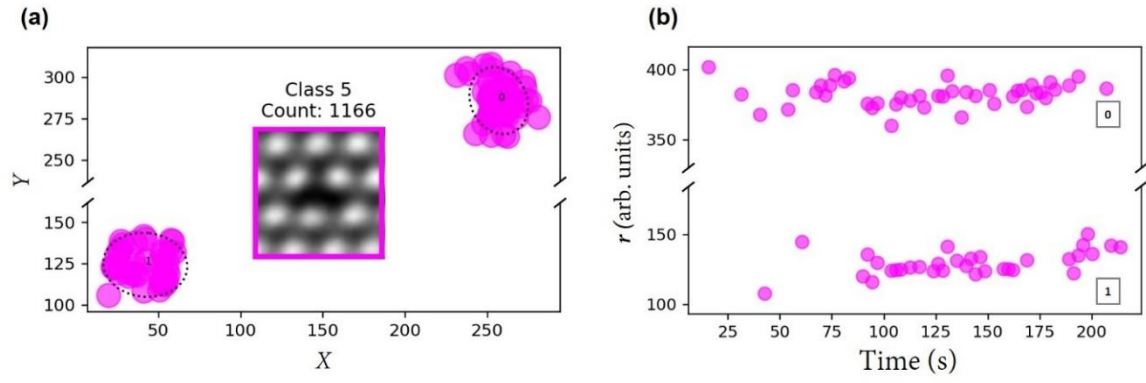


Figure 4-9. Defect flows.

(a) 2-d projections (X - Y coordinates) of the 3-d defect “flow” of the S vacancies (inset) with 95% prediction ellipses overlaid. (b) 1-d $r(t)$ representation of the same data.

We further proceed to the analysis of another type of defect, namely, the defect associated with Mo dopant (class 1 and 3). Here it is worth noting that while the GMM-based decomposition into 5 components provides us with a good understanding of the major types of defects present in the system, it doesn't allow studying the fine details (variations) of the detected structures. Performing such an analysis is especially important for classes 1 and 3 that show peculiar switching behavior in Fig. 4-8f. We therefore investigated the “internal” structures of classes 1 and 3 using the so-called local crystallography analysis.[78] We first employ a deep-learning-based 'atom finder'[235] that allows extracting positions of atoms from thousands of noisy images of defects in a matter of seconds (note that S atoms cannot be reliably identified at the current experimental resolution and hence we omit them). The first 2 PCA components associated with displacements from the averaged structure of the central Mo atom and six W neighbor atoms for each defect class are plotted in Fig. 4-10a,b. Since Mo dopant does not considerably distort the WS₂ lattice,[404] the structural variations in PCA analysis must be associated with a disorder in S sublattice. In general, one must exercise caution in assigning a specific physical meaning to the PCA components. However, the results shown in Fig. 4-10 strongly suggest a presence of strong variations in a relative position of central Mo atom with respect to neighbor W atoms, thus it is possible that these variations originate from the presence of S vacancies next to Mo dopant.

Based on the PCA analysis of the atomic displacements in Fig. 4-10a,b and general lattice symmetry considerations we use GMM to split the defect structures from classes 1 and 3 into four subclasses (Fig. 4-11b) associated with undistorted **Mo_w** defect (no coupling to S vacancy) and three (**Mo_w + V_s**) complexes (it is worth noting that the same result can be achieved by splitting the entire stack of all the defect images into >12 classes, see Figure 4-13). Our next goal is to analyze the switching behavior between different states. Using the same approach as described for the analysis of diffusion parameters we first identified continuous in time defect trajectories for all the 4 subclasses from Fig. 6b, isolated them, and then converted into $r(t)$ 1-d representation (Fig. 6a). In this case, each “flow” represents a sequence (in time) of defect structures undergoing a switching between four different states. This observation suggests that the switching

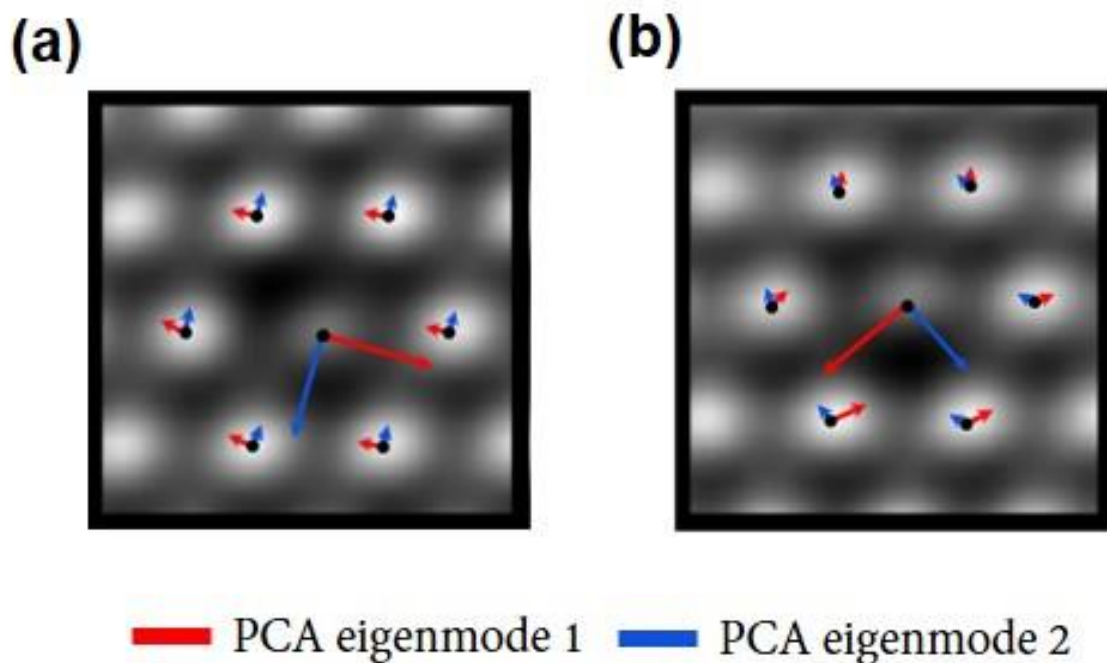


Figure 4-10. Local crystallography analysis.

(a, b) The principal component analysis-derived first two eigenmodes of atomic displacements for defects associated with a Mo dopant (class 1 and class 3 in Fig. 4-8a and 4-8c, respectively) presented as vectors of deformation from atom positions in the averaged defect structure.

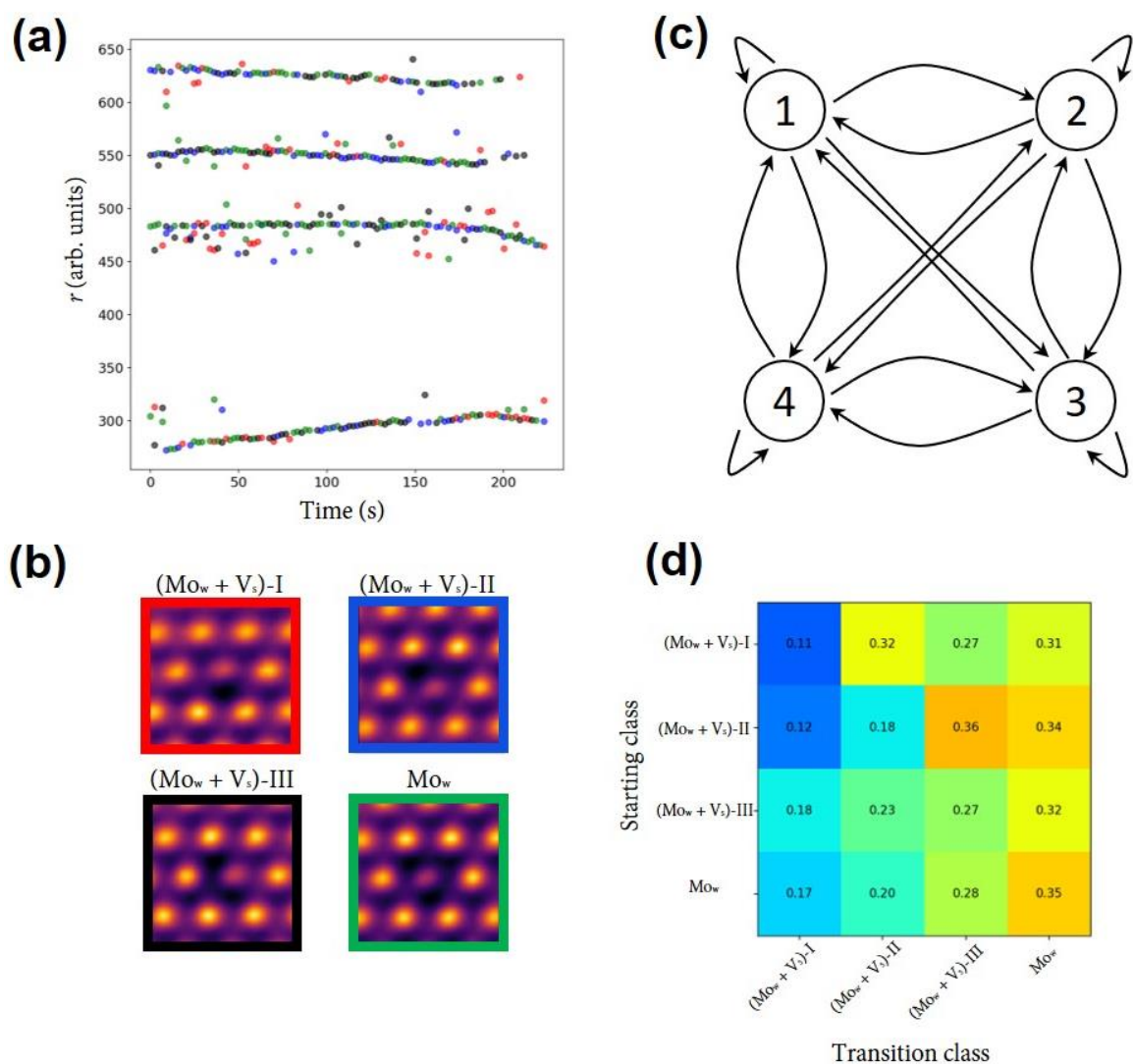


Figure 4-11. Analysis of transition probabilities between different defect states.

(a) 1-d representation of defect “flows” for the 4 subclasses of defect associated with Mo dopant in (b). Color scheme is the same as in (a) and (b). (c) Illustration of Markov transition processes between 4 states. (d) Markov transition matrix for the 4 subclasses (lone Mo dopant and 3 complexes of Mo dopant with S vacancy) based on the analysis of trajectories in (a).

between different states can be analyzed as a Markov process, defining corresponding reaction constants on a single defect level.

The corresponding Markov transition matrix is depicted in Fig. 4-11d (see also the schematics of transitions in Fig. 4-11c). This analysis suggests the **Mo_w** defect may couple to S vacancy in the dynamic STEM experiment. To explain transitions between **Mo_w** and (**Mo_w + V_s**) we argue that, due to a lower diffusion barrier of a S vacancy, as well as higher probability of S sublattice atoms being knocked-out during the e-beam irradiation, it is likely that the S vacancy created in the vicinity of Mo dopant can get “captured” by it. Interestingly, we also found transitions between different (**Mo_w + V_s**) structures. While the detailed explanation of such a behavior would require rigorous first-principles calculations and additional experiments, one can argue that the (**Mo_w + V_s + V_s**) structures are not stable and/or have a very short “lifetime” compared to the experimental time resolution (to the best of our knowledge such structures have not been observed even when (**Mo_w + V_s**) defects are abundant) and therefore attachment of the second S vacancy leads to “pushing” one of the two S vacancies out of the structure. The noticeably different values of transition probabilities for (**Mo_w + V_s**) - **I** structure can be explained by a different rate of “supply” of S vacancies from different lattice directions, for example due to non-trivial distribution of strain fields during e-beam induced transformation and their effect on diffusion characteristics in different lattice directions.

4.3.6. Conclusions and Outlook

In summary, we have presented a deep-learning-based approach for analysis of dynamic transformation of the lattice structure in STEM “movies” from Mo-doped WS₂. We started by teaching a deep neural network how the defects that break lattice periodicity appear in STEM data using a single experimental image (frame 0) and then used the generalization abilities of the network to find various types of atomic defects in the rest of the experimental data. We then performed unsupervised classification of the detected defect structures using Gaussian mixture model and showed that the classification results can be linked to specific physical structures. We were then able to i) identify dominant point

defects and their characteristic statistical behaviors in the spatiotemporal diagrams, ii) analyze diffusion for the selected defect species (S vacancy), and iii) study transformation pathways for Mo-S complexes, including detailed transition probabilities. In this manner, point-defect dynamics and solid state reactions in material are studied on atomic level, and corresponding reaction constants are determined for just one point defect.

As far as the future studies are concerned, we believe that one particularly promising direction is incorporating specific physics-based constraints into the machine learning based analysis of STEM videos. Indeed, the current approaches treat observed lattice defects as collections of pixels, without “understanding” the physics behind the observations. One possible way of overcoming such physics-agnostic classification is by integrating a Markov model into the initial search and identification/classification scheme. The Markov model can be guided by the theoretical calculations of interaction potentials on the atomic level, enforcing physical constraints to transition probabilities of atoms and molecules, effects of electron beam on the matter, operating both in space (hidden Markov random field) and time (hidden Markov model) domains. For example, one may incorporate transition probabilities between certain types of defects (*e.g.*, reconstructed vs. non-reconstructed defect), as well as a maximum diffusion length of a defect for a given time scale calculated from first principles, and with a Markov model use it to refine the results of the initial classification. This would be an important step towards creating a fully-autonomous, AI microscope that is making decisions based on the knowledge of physics that it was “taught”.

4.3.7. Methods

Sample preparation

The Mo doped WS₂ monolayers were grown on SiO₂/Si substrate at 800°C by a low-pressure chemical vapor deposition. [392] To prepare STEM samples, poly(methyl methacrylate), PMMA (A4), was first spun onto the SiO₂/Si substrate with monolayer crystals at 3500 rpm for 60 s. After being cured at 100 °C for 15 min, the PMMA/W_{1-x}Mo_xS₂ sample was detached from the substrate with a 30% KOH solution (100 °C and

0.5–1.0 h). The sample was then transferred to DI water to remove the KOH residue. The washed film was scooped onto a QUANTIFOIL TEM grid. The PMMA was then removed with acetone, and the samples were soaked in methanol for 12 h to achieve a clean surface with flakes. To remove the polymer, the TEM grids were then annealed in an Ar flow (90 sccm, 10 Torr) at 350 °C for 3 h.

STEM experiment

STEM imaging was performed using a Nion UltraSTEM U100 STEM operated at 100kV. The images were acquired in high angle annular dark field (HAADF) imaging mode and were introduced to the DL network without any post processing.

Data Analysis

The DL network was implemented using Keras 2.0 (<https://keras.io>) Python deep learning library, with the TensorFlow backend. The convolutional neural network had an encoder-decoder type of structure. The encoder part had alternating convolutional layers for feature extraction and max-pooling layers to account for translational invariance as well as for reducing the size of processed data. The filters (“kernels”) in each convolutional layer had size (3×3) and stride 1, activated by a rectified linear unit (ReLU) function. The decoder part of the network, whose role was to map the encoded low-resolution feature maps to full input-resolution feature maps, consisted of the same filters but in reverse order and upsampling layers. The feature maps from the final convolutional layer of the network were fed into a softmax classifier for pixel-wise classification, providing us with information on the probability of each pixel being a defect. The *Adam optimizer*[234] was used for training. The Gaussian mixture model was implemented with scikit-learn machine learning library (<http://scikit-learn.org>). The complete workflow for studying defects in dynamic STEM data, which includes creation and training/testing of DL model, unsupervised defect classification, analysis of diffusion characteristics, local crystallography analysis and Markov transition matrix analysis, is available in a form of Jupyter notebooks in the supplementary material.

Appendix

Appendix 4-1. Supplementary Information for Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS₂

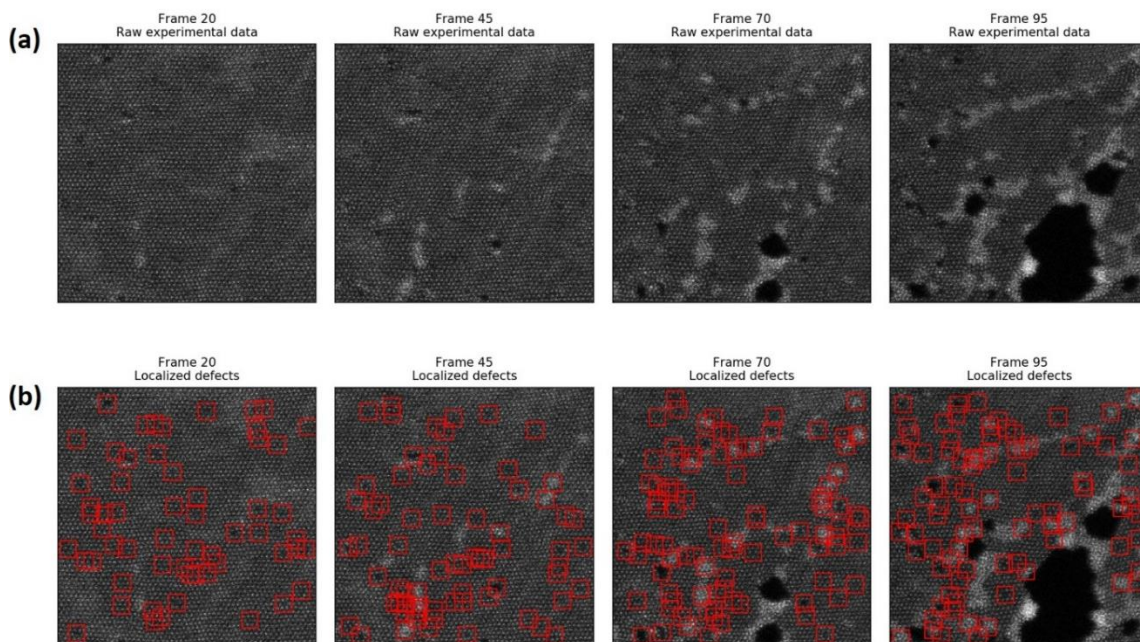


Figure 4-12. (a) Raw experimental data (selected frames). (b) Same data with *point-like defects* localized. Note that here we do not consider extended defects.

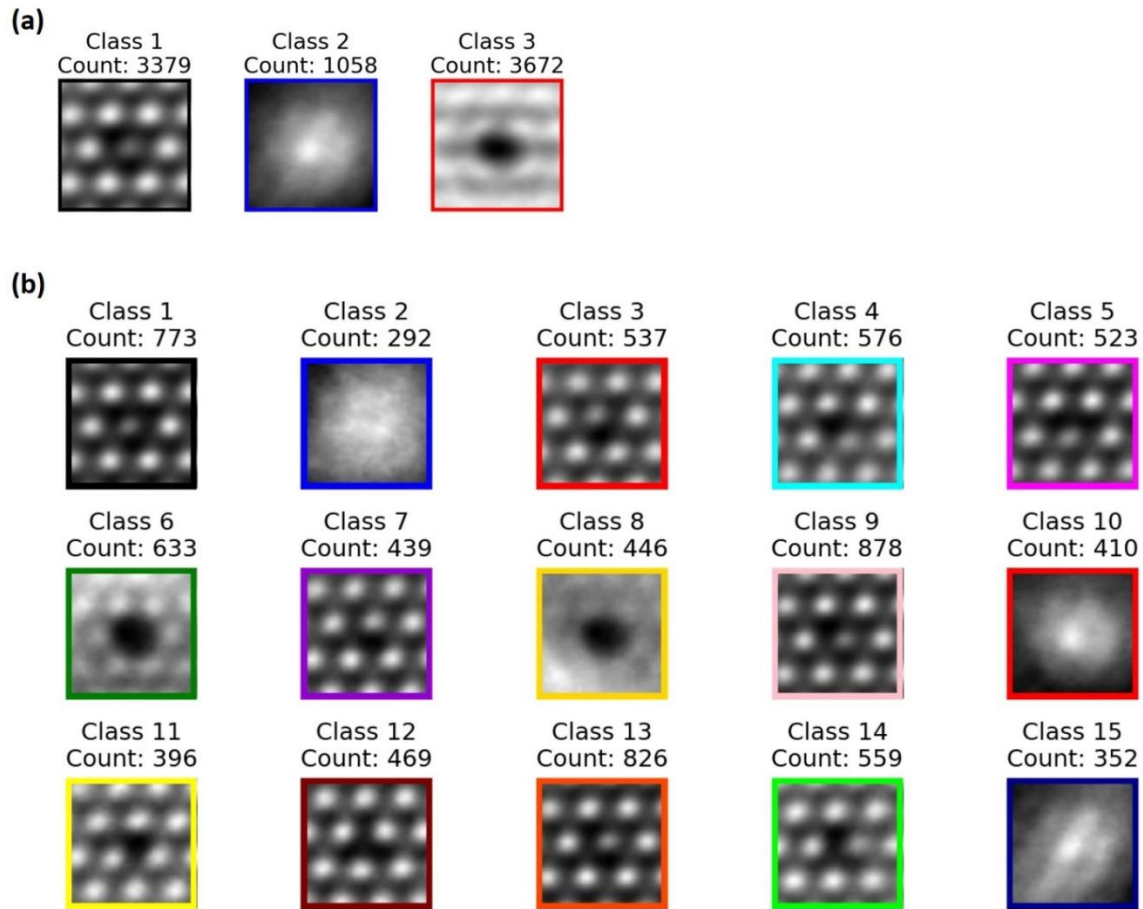


Figure 4-13. (a) GMM splitting into 3 clusters. The class 3 structure is clearly distorted due to insufficient number of components (b) GMM splitting into 15 clusters.

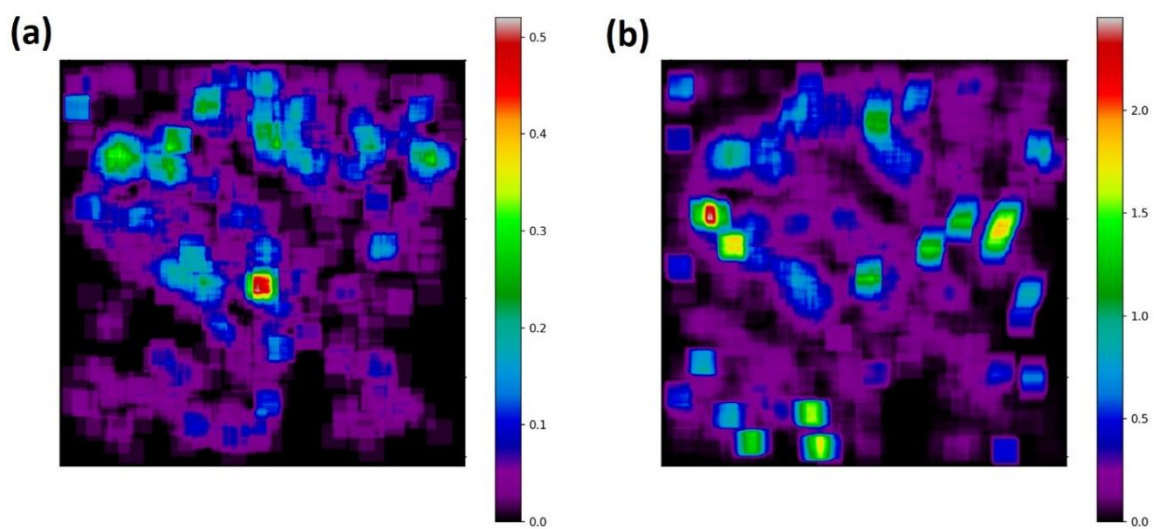


Figure 4-14. (a) Projection of sub-windows associated with S vacancy (class 5). (b) Projection of sub-windows associated with Mo defects (class 1 and class 3)

CHAPTER 5.

CONCLUSION

5.1 Summary

My graduate research started from investigating relatively simple image processing algorithms for identification of atomic positions in microscopy images and has grown to utilize a variety of advanced machine learning methods to extract physically meaningful information from both experimental and simulated images, videos, 3D simulations, and hyperspectral data. In this chapter I present a brief summary of each project I worked on and further elaborate on my vision of the future of machine learning in materials science.

Chapter 2 concentrates on examples of using machine learning algorithms and image processing for extraction of relevant statistical information. In section 2.2, I began with identifying atomic positions and creating atomic lattice on STM and STEM images of RuCl_3 . I performed PCA on extracted data and identified real-space modes of displacement, which were not predicted by either reciprocal space imaging or DFT simulations.

I continued to use identification of atomic positions in STM to create input for statistical matching of theory and experiment in case of the $\text{FeSe}_x\text{Te}_{1-x}$ superconductor. Section 2.3 describes the results and the future implications of this approach.

In section 2.4, I used PCA, K-means clustering, Bayesian Linear Unmixing, Non-negative Matrix Factorization, and Local Indicators of Spatial Association to analyze hyperspectral data of superconductor materials.

Chapter 3 describes utilization of deep learning algorithms in combination with theory-based simulations and refinement methods to achieve physically-meaningful results that were previously impossible to get. In section 3.2, I used deep fully convolutional neural networks trained on first-principle simulations to find and classify atoms in STEM images of graphene, which was not possible using previous image processing methods. I expanded the training set to include different types of lattices and successfully applied the new networks to other types of lattices and even STM images, with methods and results summarized in section 3.3.

In section 3.4, I used deep convolutional neural networks trained on simulated STM images from DFT calculations to identify states of sumanene molecules in STM images. I further improved accuracy of such state decoding by using network outputs as an input to a Markov random field model involving state probabilities derived from DFT simulations. Finally, I used Local Indicators of Spatial Association to explore the process of relaxation in the self-assembly.

Finally, section 3.5 describes the use of variational autoencoders for exploring the underlying order parameters in 3D simulations of relaxor ferroelectric materials.

In Chapter 4 I discuss the perspectives of utilizing electron microscopy for matter manipulation on the atomic scale and how machine learning can help us with it. In section 4.2, I demonstrate proof-of-concept results in a study of crystallization and amorphization of silicon explored by both experiment and molecular dynamics simulations, where k-means clustering and non-negative matrix factorization were used to study the homogeneity of resulting material.

In section 4.3, I further create a framework to study defect evolution in a STEM movie of WS₂. I created a physically-meaningful training set through utilization of FFT subtraction and data augmentation methods. I used it with a fully-convolutional deep neural network to achieve pixel level segmentation of the images and localization of individual defects. I extracted identified defects and use a Gaussian mixture model to explore the resulting data and to classify the defects. Knowing both the positions and types of the defects, I track their evolution in time as a Markov process and extract underlying transition probabilities.

My research has paved the way for automatic microscopy image understanding, which is the first significant step towards automated atomic scale material design and manufacturing. I am leaving the blueprints for methods that allow to study the physical properties of imaged/simulated materials. Deep learning frameworks can now look at the images/movies and know what they contain.

5.2 Future outlook

There are considerable opportunities for further research in the area of using artificial intelligence methods in materials sciences. In this part of the chapter I will briefly outline the most pressing tasks in automation of material characterization and manufacturing on the atomic scale.

5.2.1 Automating electron microscope for atomic scale manufacturing

The most pressing improvement that can be achieved is to create a feedback loop between the electron microscope and the deep learning frameworks trained for specific applications. Such a feedback loop would understand what is imaged through methods described in this dissertation. This would result in an automated characterization of the materials, speeding up the process of material discovery significantly.

A second step is to begin transforming the experiment design itself. Recent advances in reinforcement learning has shown that properly trained neural networks can, for example, play video games better than humans (and without breaks) [409, 410], and this has been noticed in the past as a great perspective for microscopy applications [411]. With automated extraction of information from images, and abilities to create simulated datasets, we can train a reinforcement learning model to be able to modify the imaged material to match a desired template. This can be done non-stop, in an automated fashion, which would lead to creation of nano-factories for atomic scale manufacturing in the nearest future.

5.2.2 Quantifying the uncertainty of algorithms.

One of the most significant concerns about deep learning in the scientific community right now is regarding the accuracies of the results. Indeed, uncertainty quantification is incredibly important with the small amounts of data scientific experiments produce. However, right now it is very difficult to achieve due to the lack of labeled experimental

data. While automating microscopy can help us with data collection, it would still be great to be able to achieve better statistical understanding of the frameworks we utilize. Some of the methods described in this dissertation already used Bayesian inference for materials understanding, and there is a significant body of work on using Gaussian processes in materials science. [38, 412]. One of the most promising approaches is to start utilizing a combination of deep learning and Gaussian processes – Neural Processes. [413, 414]

5.2.3 Expanding machine understanding from image to material.

A natural way to represent crystalline or molecular materials is a graph model, in which atoms represent nodes and edges represent interactions between the materials. While this property has been used for classifying molecular states in this dissertation, neural networks in the framework was not directly exposed to such understanding. With recent advances in research of graph networks [415, 416], it should become possible for a network to process imaged materials not on the level of images/videos, but on the level of understanding the actual material representation as a graph model. Combined with continued development of materials databases [14, 417], such frameworks will significantly improve characterization of materials and matter manipulation on the atomic scale.

5.3 AI in Materials Science

The world is on the verge of being transformed by wide adoption of AI methods, and materials science will not remain untouched. AI will perform automated image analysis, and control microscopes, allowing scientists to concentrate on more fundamental tasks. Automated manufacturing will become widespread, from buildings down to the atomic structures. I hope that my research will be a stepping stone towards achieving this future.

REFERENCES

1. Cyberstates 2018. 2018 [cited 2018 May 5]; Available from: http://cyberstates.org/pdf/CompTIA_Cyberstates_2018.pdf.
2. Basu, S., J.G. Fernald, and M.D. Shapiro. *Productivity growth in the 1990s: technology, utilization, or adjustment?* in *Carnegie-Rochester conference series on public policy*. 2001. Elsevier.
3. Mathuros, F., *Major Push to End the Hidden Human Toll and Pollution behind Smartphone and Electric Car Batteries*. 2017, World Economic Forum.
4. Moore, G.E., *Cramming more components onto integrated circuits*. Proceedings of the IEEE, 1998. **86**(1): p. 82-85.
5. Waldrop, M.M., *The chips are down for Moore's law*. Nature, 2016. **530**(7589): p. 144-147.
6. Clark, D., *Intel Rechisels the Tablet on Moore's Law*, in *Wall Street Journal Digits Tech News and Analysis*. 2015.
7. Fuechsle, M., et al., *A single-atom transistor*. Nature nanotechnology, 2012. **7**(4): p. 242.
8. Vugmeister, B., *Polarization dynamics and formation of polar nanoregions in relaxor ferroelectrics*. Physical Review B, 2006. **73**(17): p. 174117.
9. Woodward, D.I., J. Knudsen, and I.M. Reaney, *Review of crystal and domain structures in the $PbZr_xTi_{1-x}O_3$ solid solution*. Physical Review B, 2005. **72**(10): p. 104110.
10. Fiebig, M., *Revival of the magnetoelectric effect*. Journal of Physics D: Applied Physics, 2005. **38**(8): p. R123.
11. Binder, K. and A.P. Young, *Spin glasses: Experimental facts, theoretical concepts, and open questions*. Reviews of Modern physics, 1986. **58**(4): p. 801.
12. Dagotto, E., *Complexity in strongly correlated electronic systems*. Science, 2005. **309**(5732): p. 257-262.
13. Dagotto, E., T. Hotta, and A. Moreo, *Colossal magnetoresistant materials: the key role of phase separation*. Physics reports, 2001. **344**(1-3): p. 1-153.
14. Jain, A., et al., *Commentary: The Materials Project: A materials genome approach to accelerating materials innovation*. Apl Materials, 2013. **1**(1): p. 011002.
15. Holdren, J.P., *Materials genome initiative for global competitiveness*. National Science and technology council OSTP. Washington, USA, 2011.
16. Spaldin, N.A. and M. Fiebig, *The renaissance of magnetoelectric multiferroics*. Science, 2005. **309**(5733): p. 391-392.
17. Adler, S.B., *Factors governing oxygen reduction in solid oxide fuel cell cathodes*. Chemical reviews, 2004. **104**(10): p. 4791-4844.
18. Young, S.R., et al., *Data mining for better material synthesis: The case of pulsed laser deposition of complex oxides*. Journal of Applied Physics, 2018. **123**(11): p. 115303.
19. *The First Five Years of the Materials Genome Initiative: Accomplishments and Technical Highlights*. 2016, Materials Genome Initiative.
20. Curtarolo, S., et al., *The high-throughput highway to computational materials design*. Nature materials, 2013. **12**(3): p. 191.

21. McCorduck, P., *Machines who think: A personal inquiry into the history and prospects of artificial intelligence*. 2009: AK Peters/CRC Press.
22. Lorica, B. *The state of AI adoption*. 2017 December 18, 2017 [cited 2018 June 5]; Available from: <https://www.oreilly.com/ideas/the-state-of-ai-adoption>.
23. Oh, K.-S. and K. Jung, *GPU implementation of neural networks*. Pattern Recognition, 2004. **37**(6): p. 1311-1314.
24. Jouppi, N.P., et al. *In-datacenter performance analysis of a tensor processing unit*. in *Proceedings of the 44th Annual International Symposium on Computer Architecture*. 2017. ACM.
25. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning*. nature, 2015. **521**(7553): p. 436.
26. Goodfellow, I., et al., *Deep learning*. Vol. 1. 2016: MIT press Cambridge.
27. He, K., et al. *Deep residual learning for image recognition*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.
28. Hinton, G., et al., *Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups*. IEEE Signal processing magazine, 2012. **29**(6): p. 82-97.
29. Amodei, D., et al. *Deep speech 2: End-to-end speech recognition in english and mandarin*. in *International Conference on Machine Learning*. 2016.
30. Martin, J., et al., *Observation of electron-hole puddles in graphene using a scanning single-electron transistor*. Nature Physics, 2008. **4**(2): p. 144.
31. Zhang, Y., et al., *Origin of spatial charge inhomogeneity in graphene*. Nature Physics, 2009. **5**(10): p. 722.
32. Xu, G., et al., *Phase instability induced by polar nanoregions in a relaxor ferroelectric system*. Nature materials, 2008. **7**(7): p. 562.
33. Gofryk, K., et al., *Local inhomogeneity and filamentary superconductivity in Pr-doped CaFe₂As₂*. Physical review letters, 2014. **112**(4): p. 047005.
34. Gomes, K.K., et al., *Visualizing pair formation on the atomic scale in the high-T_c superconductor Bi₂Sr₂CaCu₂O_{8+δ}*. Nature, 2007. **447**(7144): p. 569.
35. Williams, D.B. and C.B. Carter, *The transmission electron microscope*, in *Transmission electron microscopy*. 1996, Springer. p. 3-17.
36. Crewe, A.V., *Scanning transmission electron microscopy*. Journal of microscopy, 1974. **100**(3): p. 247-259.
37. Binnig, G. and H. Rohrer, *Scanning tunneling microscopy*. IBM Journal of research and development, 2000. **44**(1/2): p. 279.
38. Mueller, T., A.G. Kusne, and R. Ramprasad, *Machine learning in materials science: Recent progress and emerging applications*. Reviews in Computational Chemistry, 2016. **29**: p. 186-273.
39. Baldi, P., P. Sadowski, and D. Whiteson, *Searching for exotic particles in high-energy physics with deep learning*. Nature communications, 2014. **5**: p. 4308.
40. Baldi, P., et al., *Jet substructure classification in high-energy physics with deep neural networks*. Physical Review D, 2016. **93**(9): p. 094034.
41. Guest, D., et al., *Jet flavor classification in high-energy physics with deep neural networks*. Physical Review D, 2016. **94**(11): p. 112002.

42. Angermueller, C., et al., *Deep learning for computational biology*. Molecular systems biology, 2016. **12**(7): p. 878.
43. Litjens, G., et al., *A survey on deep learning in medical image analysis*. Medical image analysis, 2017. **42**: p. 60-88.
44. Chen, C., *CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature*. Journal of the American Society for information Science and Technology, 2006. **57**(3): p. 359-377.
45. *Web of Science*. 2018, Thomson Reuters.
46. Krizhevsky, A., I. Sutskever, and G. Hinton, *ImageNet classification with deep convolutional neural networks.*, in *Proc. Advances in Neural Information Processing Systems*. 2012. p. 1090-1098.
47. Simonyan, K. and A. Zisserman, *Very deep convolutional networks for large-scale image recognition*. arXiv preprint arXiv:1409.1556, 2014.
48. He, K., et al. *Identity mappings in deep residual networks*. in *European conference on computer vision*. 2016. Springer.
49. Zagoruyko, S. and N. Komodakis, *Wide residual networks*. arXiv preprint arXiv:1605.07146, 2016.
50. Pearl, J., *Theoretical Impediments to Machine Learning With Seven Sparks from the Causal Revolution*. arXiv preprint arXiv:1801.04016, 2018.
51. Marcus, G., *Deep Learning: A Critical Appraisal*. arXiv preprint arXiv:1801.00631, 2018.
52. Ziatdinov, M., et al., *Atomic-scale observation of structural and electronic orders in the layered compound α -RuCl₃*. Nature communications, 2016. **7**: p. 13774.
53. Sears, J.A., et al., *Magnetic order in α -RuCl₃: A honeycomb-lattice quantum magnet with strong spin-orbit coupling*. Physical Review B, 2015. **91**(14): p. 144420.
54. Sandilands, L.J., et al., *Scattering Continuum and Possible Fractionalized Excitations in α -RuCl₃*. Physical Review Letters, 2015. **114**(14): p. 147201.
55. Banerjee, A., et al., *Proximate Kitaev quantum spin liquid behaviour in a honeycomb magnet*. Nat Mater, 2016. **15**(7): p. 733-740.
56. Rau, J.G., E.K.-H. Lee, and H.-Y. Kee, *Spin-Orbit Physics Giving Rise to Novel Phases in Correlated Systems: Iridates and Related Materials*. Annual Review of Condensed Matter Physics, 2016. **7**(1): p. 195-221.
57. Hwan Chun, S., et al., *Direct evidence for dominant bond-directional interactions in a honeycomb lattice iridate Na₂IrO₃*. Nat Phys, 2015. **11**(6): p. 462-466.
58. Shitade, A., et al., *Quantum Spin Hall Effect in a Transition Metal Oxide Na₂IrO₃*. Physical Review Letters, 2009. **102**(25): p. 256403.
59. Witczak-Krempa, W., et al., *Correlated Quantum Phenomena in the Strong Spin-Orbit Regime*. Annual Review of Condensed Matter Physics, 2014. **5**(1): p. 57-82.
60. Jackeli, G. and G. Khaliullin, *Mott Insulators in the Strong Spin-Orbit Coupling Limit: From Heisenberg to a Quantum Compass and Kitaev Models*. Physical Review Letters, 2009. **102**(1): p. 017205.

61. Petrova, O., P. Mellado, and O. Tchernyshyov, *Unpaired Majorana modes in the gapped phase of Kitaev's honeycomb model*. Physical Review B, 2013. **88**(14): p. 140405.
62. Singh, Y., et al., *Relevance of the Heisenberg-Kitaev Model for the Honeycomb Lattice Iridates $A_2\text{IrO}_3$* . Physical Review Letters, 2012. **108**(12): p. 127203.
63. Jiang, H.-C., et al., *Possible proximity of the Mott insulating iridate Na_2IrO_3 to a topological phase: Phase diagram of the Heisenberg-Kitaev model in a magnetic field*. Physical Review B, 2011. **83**(24): p. 245104.
64. Rachel, S., L. Fritz, and M. Vojta, *Landau Levels of Majorana Fermions in a Spin Liquid*. Physical Review Letters, 2016. **116**(16): p. 167201.
65. Sarma, S.D., M. Freedman, and C. Nayak, *Majorana zero modes and topological quantum computation*. Npj Quantum Information, 2015. **1**: p. 15001.
66. Vamshi, M.K., et al., *Kitaev interactions between $j = 1/2$ moments in honeycomb Na_2IrO_3 are large and ferromagnetic: insights from ab initio quantum chemistry calculations*. New Journal of Physics, 2014. **16**(1): p. 013056.
67. Katukuri, V.M., et al., *Strong magnetic frustration and anti-site disorder causing spin-glass behavior in honeycomb Li_2RhO_3* . Scientific Reports, 2015. **5**: p. 14718.
68. Nishimoto, S., et al., *Strongly frustrated triangular spin lattice emerging from triplet dimer formation in honeycomb Li_2IrO_3* . Nat Commun, 2016. **7**.
69. Plumb, K.W., et al., *$\alpha\text{-RuCl}_3$: A spin-orbit assisted Mott insulator on a honeycomb lattice*. Physical Review B, 2014. **90**(4): p. 041112.
70. Sandilands, L.J., et al., *Spin-orbit excitations and electronic structure of the putative Kitaev magnet $\alpha\text{-RuCl}_3$* . Physical Review B, 2016. **93**(7): p. 075144.
71. Kubota, Y., et al., *Successive magnetic phase transitions in $\alpha\text{-RuCl}_3$: XY-like frustrated magnet on the honeycomb lattice*. Physical Review B, 2015. **91**(9): p. 094422.
72. Cantow, H.-J., et al., *Atomic Structure and Superstructure of $\alpha\text{-RuCl}_3$ by Scanning Tunneling Microscopy*. Angewandte Chemie International Edition in English, 1990. **29**(5): p. 537-541.
73. Johnson, R.D., et al., *Monoclinic crystal structure of $\alpha\text{-RuCl}_3$ and the zigzag antiferromagnetic ground state*. Physical Review B, 2015. **92**(23): p. 235119.
74. Cao, H.B., et al., *Low-temperature crystal and magnetic structure of $\alpha\text{-RuCl}_3$* . Physical Review B, 2016. **93**(13): p. 134423.
75. Stroganov, E.V. and K.V. Ovchinnikov, Ser. Fiz. Khim., 1957. **12**: p. 152.
76. Kim, H.-S. and H.-Y. Kee, *Crystal structure and magnetism in $\alpha\text{-RuCl}_3$: An ab initio study*. Physical Review B, 2016. **93**(15): p. 155143.
77. Bonnet, N., *Artificial intelligence and pattern recognition techniques in microscope image processing and analysis*, in *Advances in Imaging and Electron Physics*, W.H. Peter, Editor. 2000, Elsevier. p. 1-77.
78. Belianinov, A., et al., *Identification of phases, symmetries and defects through local crystallography*. Nat Commun, 2015. **6**.

79. Belianinov, A., et al., *Big data and deep data in scanning and electron microscopies: deriving functionality from multidimensional data sets*. Adv. Struct. Chem. Imaging., 2015. **1**.
80. Kalinin, S.V., B.G. Sumpter, and R.K. Archibald, *Big-deep-smart data in imaging for guiding materials design*. Nat Mater, 2015. **14**(10): p. 973-980.
81. Xu, G., et al., *Phase instability induced by polar nanoregions in a relaxor ferroelectric system*. Nat Mater, 2008. **7**(7): p. 562-566.
82. Gomes, K.K., et al., *Visualizing pair formation on the atomic scale in the high-Tc superconductor Bi₂Sr₂CaCu₂O₈+*[dgr]**. Nature, 2007. **447**(7144): p. 569-572.
83. Rau, J.G., E.K.-H. Lee, and H.-Y. Kee, *Generic Spin Model for the Honeycomb Iridates beyond the Kitaev Limit*. Physical Review Letters, 2014. **112**(7): p. 077204.
84. Yadav, R., et al. *Spin-orbit excitation energies, anisotropic exchange, and magnetic phases of honeycomb RuCl₃*. ArXiv e-prints, 2016. **1604**.
85. Sela, E., et al., *Order-by-disorder and spin-orbital liquids in a distorted Heisenberg-Kitaev model*. Physical Review B, 2014. **90**(3): p. 035113.
86. Rousochatzakis, I., et al., *Phase Diagram and Quantum Order by Disorder in the Kitaev K₁-K₂ Honeycomb Magnet*. Physical Review X, 2015. **5**(4): p. 041035.
87. Banerjee, A., et al. *Neutron tomography of magnetic Majorana fermions in a proximate quantum spin liquid*. ArXiv e-prints, 2016. **1609**.
88. Pollini, I., *Electronic properties of the narrow-band material α -RuCl₃*. Physical Review B, 1996. **53**(19): p. 12769-12776.
89. Comin, R., et al., *Na₂IrO₃ as a Novel Relativistic Mott Insulator with a 340-meV Gap*. Physical Review Letters, 2012. **109**(26): p. 266406.
90. Kim, H.-S., et al., *Kitaev magnetism in honeycomb RuCl₃ with intermediate spin-orbit coupling*. Physical Review B, 2015. **91**(24): p. 241110.
91. Zhou, X., et al. *ARPES study of the Kitaev Candidate α -RuCl₃*. ArXiv e-prints, 2016. **1603**.
92. Binotto, L., I. Pollini, and G. Spinolo, *Optical and transport properties of the magnetic semiconductor α -RuCl₃*. physica status solidi (b), 1971. **44**(1): p. 245-252.
93. Lüpke, F., et al., *Highly unconventional surface reconstruction of Na₂IrO₃ with persistent energy gap*. Physical Review B, 2015. **91**(4): p. 041405.
94. Chaloupka, J., G. Jackeli, and G. Khaliullin, *Zigzag Magnetic Order in the Iridium Oxide Na₂IrO₃*. Physical Review Letters, 2013. **110**(9): p. 097204.
95. Soumyanarayanan, A., et al., *Quantum phase transition from triangular to stripe charge order in NbSe₂*. Proceedings of the National Academy of Sciences, 2013. **110**(5): p. 1623-1627.
96. Weber, D., et al., *Magnetic Properties of Restacked 2D Spin 1/2 honeycomb RuCl₃ Nanosheets*. Nano Letters, 2016. **16**(6): p. 3578-3584.
97. Perdew, J.P., K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*. Physical Review Letters, 1996. **77**(18): p. 3865-3868.
98. Blöchl, P.E., *Projector augmented-wave method*. Physical Review B, 1994. **50**(24): p. 17953-17979.

99. Kresse, G. and J. Furthmüller, *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*. Physical Review B, 1996. **54**(16): p. 11169-11186.
100. Kresse, G. and D. Joubert, *From ultrasoft pseudopotentials to the projector augmented-wave method*. Physical Review B, 1999. **59**(3): p. 1758-1775.
101. Monkhorst, H.J. and J.D. Pack, *Special points for Brillouin-zone integrations*. Physical Review B, 1976. **13**(12): p. 5188-5192.
102. Vlcek, L., et al., *Knowledge Extraction from Atomically Resolved Images*. ACS nano, 2017. **11**(10): p. 10313-10320.
103. Cherns, D., *Direct resolution of surface atomic steps by transmission electron microscopy*. The Philosophical Magazine: A Journal of Theoretical Experimental and Applied Physics, 1974. **30**(3): p. 549-556.
104. Wall, J., et al., *Scanning Transmission Electron Microscopy at High Resolution*. Proceedings of the National Academy of Sciences, 1974. **71**(1): p. 1-5.
105. Isaacson, M., et al., *Direct observations of atomic diffusion by scanning transmission electron microscopy*. Proceedings of the National Academy of Sciences, 1977. **74**(5): p. 1802-1806.
106. Pennycook, S.J., *A Scan Through the History of STEM*, in *Scanning Transmission Electron Microscopy: Imaging and Analysis*, S.J. Pennycook and P.D. Nellist, Editors. 2011, Springer New York: New York, NY. p. 1-90.
107. Schmid, M., H. Stadler, and P. Varga, *Direct observation of surface chemical order by scanning tunneling microscopy*. Physical Review Letters, 1993. **70**(10): p. 1441-1444.
108. Jia, C.-L., et al., *Atomic-scale study of electric dipoles near charged and uncharged domain walls in ferroelectric films*. Nature Materials, 2007. **7**: p. 57.
109. Jia, C.-L., et al., *Unit-cell scale mapping of ferroelectricity and tetragonality in epitaxial ultrathin ferroelectric films*. Nature Materials, 2006. **6**: p. 64.
110. Chisholm, M.F., et al., *Atomic-Scale Compensation Phenomena at Polar Interfaces*. Physical Review Letters, 2010. **105**(19): p. 197602.
111. Chang Hye, J., et al., *Atomically Resolved Mapping of Polarization and Electric Fields Across Ferroelectric/Oxide Interfaces by Z-contrast Imaging*. Advanced Materials, 2011. **23**(21): p. 2474-2479.
112. Kim, Y.-M., et al., *Probing oxygen vacancy concentration and homogeneity in solid-oxide fuel-cell cathode materials on the subunit-cell level*. Nature Materials, 2012. **11**: p. 888.
113. Kim, Y.-M., et al., *Direct observation of ferroelectric field effect and vacancy-controlled screening at the BiFeO₃/La_xSr_{1-x}MnO₃ interface*. Nature Materials, 2014. **13**: p. 1019.
114. Borisevich, A., et al., *Mapping Octahedral Tilts and Polarization Across a Domain Wall in BiFeO₃ from Z-Contrast Scanning Transmission Electron Microscopy Image Atomic Column Shape Analysis*. ACS Nano, 2010. **4**(10): p. 6071-6079.

115. Borisevich, A.Y., et al., *Suppression of Octahedral Tilts and Associated Changes in Electronic Properties at Epitaxial Oxide Heterostructure Interfaces*. Physical Review Letters, 2010. **105**(8): p. 087204.
116. Kim, Y.M., et al., *Interplay of Octahedral Tilts and Polar Order in BiFeO₃ Films*. Advanced Materials, 2013. **25**(17): p. 2497-2504.
117. Jia, C.L., et al., *Oxygen octahedron reconstruction in the SrTiO₃/LaAlO₃ heterointerfaces investigated using aberration-corrected ultrahigh-resolution transmission electron microscopy*. Physical Review B, 2009. **79**(8): p. 081405.
118. Brune, H., et al., *Effect of strain on surface diffusion and nucleation*. Physical Review B, 1995. **52**(20): p. R14380-R14383.
119. Pedersen, M.Ø., et al., *Diffusion of N Adatoms on the Fe(100) Surface*. Physical Review Letters, 2000. **84**(21): p. 4898-4901.
120. Anthony, G.G., et al., *The Ehrlich–Schwoebel barrier on an oxide surface: a combined Monte-Carlo and in situ scanning tunneling microscopy approach*. Nanotechnology, 2015. **26**(45): p. 455705.
121. Voyles, P.M., *Informatics and data science in materials microscopy*. Current Opinion in Solid State and Materials Science, 2017. **21**(3): p. 141-158.
122. Ishikawa, R., et al., *Three-Dimensional Location of a Single Dopant with Atomic Precision by Aberration-Corrected Scanning Transmission Electron Microscopy*. Nano Letters, 2014. **14**(4): p. 1903-1908.
123. Cossio, P., et al., *BioEM: GPU-accelerated computing of Bayesian inference of electron microscopy images*. Computer Physics Communications, 2017. **210**: p. 163-171.
124. Ophus, C., et al., *Large-scale experimental and theoretical study of graphene grain boundary structures*. Physical Review B, 2015. **92**(20): p. 205402.
125. Whitlam, S., et al., *Common Physical Framework Explains Phase Behavior and Dynamics of Atomic, Molecular, and Polymeric Network Formers*. Physical Review X, 2014. **4**(1): p. 011044.
126. Hefei, H., et al., *Phase separation in the iron chalcogenide superconductor Fe 1+ y Te x Se 1− x*. New Journal of Physics, 2011. **13**(5): p. 053031.
127. Liu, Y., et al., *In situ observation of phase separation and hierarchical microstructure of KxFe2-ySe2 single crystals*. arXiv preprint arXiv:1502.04335, 2015.
128. Lin, W., et al., *Direct Probe of Interplay between Local Structure and Superconductivity in FeTe0.55Se0.45*. ACS Nano, 2013. **7**(3): p. 2634-2641.
129. Hansen, J.-P. and I.R. McDonald, *Theory of simple liquids*. 1990: Elsevier.
130. Wootters, W.K., *Statistical distance and Hilbert space*. Physical Review D, 1981. **23**(2): p. 357-362.
131. Braunstein, S.L. and C.M. Caves, *Statistical distance and the geometry of quantum states*. Physical Review Letters, 1994. **72**(22): p. 3439-3443.
132. Vlcek, L. and A.A. Chialvo, *Rigorous force field optimization principles based on statistical distance minimization*. The Journal of Chemical Physics, 2015. **143**(14): p. 144110.

133. Aherne, F.J., N.A. Thacker, and P.I. Rockett, *The Bhattacharyya metric as an absolute similarity measure for frequency coded data*. Kybernetika, 1998. **34**(4): p. [363]-368.
134. Thacker, N.A., P.A. Bromiley, and E. Trucco. *The Equal Variance Domain: Issues surrounding the use of probability densities in algorithm design*. in *Presented at the Joint Meeting of the BMVA and the Royal Statistical Society*. 2004.
135. Chialvo, A.A., *Excess properties of liquid mixtures from computer simulation: a coupling-parameter approach to the determination of their dependence on molecular asymmetry*. Molecular Physics, 1991. **73**(1): p. 127-140.
136. Hestenes, M.R. and E. Stiefel, *Methods of conjugate gradients for solving linear systems*. Vol. 49. 1952: NBS Washington, DC.
137. Oka, K., et al., *Dual Defects of Cation and Anion in Memristive Nonvolatile Memory of Metal Oxides*. Journal of the American Chemical Society, 2012. **134**(5): p. 2535-2538.
138. Nakayama, M. and M. Martin, *First-principles study on defect chemistry and migration of oxide ions in ceria doped with rare-earth cations*. Physical Chemistry Chemical Physics, 2009. **11**(17): p. 3241-3249.
139. Shin, B. and M.J. Aziz, *Modeling RHEED intensity oscillations in multilayer epitaxy: Determination of the Ehrlich-Schwoebel barrier in Ge(001) homoepitaxy*. Physical Review B, 2007. **76**(16): p. 165408.
140. Wang, X., et al., *In Situ Studies of the Active Sites for the Water Gas Shift Reaction over Cu–CeO₂ Catalysts: Complex Interaction between Metallic Copper and Oxygen Vacancies of Ceria*. The Journal of Physical Chemistry B, 2006. **110**(1): p. 428-434.
141. Chipman, H., E.I. George, and R.E. McCulloch, *The Practical Implementation of Bayesian Model Selection*, in *Model selection*, P. Lahiri, Editor. 2001, Institute of Mathematical Statistics: Beachwood, OH. p. 65-116.
142. Shirts, M.R. and J.D. Chodera, *Statistically optimal analysis of samples from multiple equilibrium states*. The Journal of Chemical Physics, 2008. **129**(12): p. 124105.
143. Ziatdinov, M., et al., *Deep data mining in a real space: separation of intertwined electronic responses in a lightly doped BaFe₂As₂*. Nanotechnology, 2016. **27**(47): p. 475706.
144. Martin, J., et al., *Observation of electron–hole puddles in graphene using a scanning single-electron transistor*. Nature Physics, 2007. **4**: p. 144.
145. Gofryk, K., et al., *Local Inhomogeneity and Filamentary Superconductivity in Pr-Doped CaFe₂As₂*. Physical Review Letters, 2014. **112**(4): p. 047005.
146. Gomes, K.K., et al., *Visualizing pair formation on the atomic scale in the high-*T_c* superconductor Bi₂Sr₂CaCu₂O_{8+δ}*. Nature, 2007. **447**: p. 569.
147. *Scanning Tunneling Microscopy*, ed. J.A. Stroscio and W.J. Kaiser. 1993, San Diego: Academic Press. 459.
148. Kiraly, B., et al., *Graphene–Silicon Heterostructures at the Two-Dimensional Limit*. Chemistry of Materials, 2015. **27**(17): p. 6085-6090.

149. Ugeda, M.M., et al., *Missing Atom as a Source of Carbon Magnetism*. Physical Review Letters, 2010. **104**(9): p. 096804.
150. Howald, C., P. Fournier, and A. Kapitulnik, *Inherent inhomogeneities in tunneling spectra of $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8-x}$ crystals in the superconducting state*. Physical Review B, 2001. **64**(10): p. 100504.
151. Pasupathy, A.N., et al., *Electronic Origin of the Inhomogeneous Pairing Interaction in the High- T_c Superconductor $\text{Bi}_2\text{Sr}_2\text{CaCu}_2\text{O}_{8+\delta}$* . Science, 2008. **320**(5873): p. 196-201.
152. Beidenkopf, H., et al., *Spatial fluctuations of helical Dirac fermions on the surface of topological insulators*. Nature Physics, 2011. **7**: p. 939.
153. Jennifer, E.H., *Spectroscopic scanning tunneling microscopy insights into Fe-based superconductors*. Reports on Progress in Physics, 2011. **74**(12): p. 124513.
154. Fischer, Ø., et al., *Scanning tunneling spectroscopy of high-temperature superconductors*. Reviews of Modern Physics, 2007. **79**(1): p. 353-419.
155. Kun, Z., et al., *Scanning tunneling microscopy studies of topological insulators*. Journal of Physics: Condensed Matter, 2014. **26**(39): p. 394003.
156. Stewart, G.R., *Superconductivity in iron compounds*. Reviews of Modern Physics, 2011. **83**(4): p. 1589-1652.
157. Li, L., et al., *Role of magnetism in superconductivity of BaFe_2As_2 : Study of 5d Au-doped crystals*. Physical Review B, 2015. **92**(9): p. 094504.
158. Rosenthal, E.P., et al., *Visualization of electron nematicity and unidirectional antiferroic fluctuations at high temperatures in NaFeAs* . Nature Physics, 2014. **10**: p. 225.
159. Chuang, T.-M., et al., *Nematic Electronic Structure in the “Parent” State of the Iron-Based Superconductor $\text{Ca}(\text{Fe}_{1-x}\text{Co}_x)_2\text{As}_2$* . Science, 2010. **327**(5962): p. 181-184.
160. Kresin, V.Z., Y.N. Ovchinnikov, and S.A. Wolf, *Inhomogeneous superconductivity and the “pseudogap” state of novel superconductors*. Physics Reports, 2006. **431**(5): p. 231-259.
161. Kalinin, S.V., B.G. Sumpter, and R.K. Archibald, *Big-deep-smart data in imaging for guiding materials design*. Nature Materials, 2015. **14**: p. 973.
162. Niestemski, F.C., et al., *Unveiling the Atomic and Electronic Structure at the Surface of the Parent Pnictide SrFe_2As_2* . ArXiv e-prints, 2009.
163. Strelcov, E., et al., *Deep Data Analysis of Conductive Phenomena on Complex Oxide Interfaces: Physics from Data Mining*. ACS Nano, 2014. **8**(6): p. 6449-6457.
164. Haykin, S.S., *Neural networks and learning machines*. Vol. 3. 2009: Pearson Upper Saddle River, NJ, USA:.
165. Stephen, J. and V.K. Sergei, *Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy*. Nanotechnology, 2009. **20**(8): p. 085714.
166. Bonnet, N., *Multivariate statistical methods for the analysis of microscope image series: applications in materials science*. Journal of Microscopy, 1998. **190**(1-2): p. 2-18.

167. Belianinov, A., et al., *Big data and deep data in scanning and electron microscopies: deriving functionality from multidimensional data sets*. Advanced Structural and Chemical Imaging, 2015. **1**(1): p. 6.
168. Yi, M., et al., *Dynamic competition between spin-density wave order and superconductivity in underdoped $Ba_{1-x}KxFe_2As_2$* . Nature Communications, 2014. **5**: p. 3711.
169. Zhou, T., et al., *Quasiparticle states around a nonmagnetic impurity in electron-doped iron-based superconductors with spin-density-wave order*. Physical Review B, 2011. **83**(21): p. 214502.
170. Dobigeon, N., et al., *Joint Bayesian Endmember Extraction and Linear Unmixing for Hyperspectral Imagery*. IEEE Transactions on Signal Processing, 2009. **57**(11): p. 4355-4368.
171. Strelcov, E., et al., *Constraining Data Mining with Physical Models: Voltage- and Oxygen Pressure-Dependent Transport in Multiferroic Nanostructures*. Nano Letters, 2015. **15**(10): p. 6650-6657.
172. Keshava, N. and J.F. Mustard, *Spectral unmixing*. IEEE Signal Processing Magazine, 2002. **19**(1): p. 44-57.
173. Winter, M.E. *N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data*. in *Imaging Spectrometry V*. 1999. International Society for Optics and Photonics.
174. Winter, M.E. *A proof of the N-FINDR algorithm for the automated detection of endmembers in a hyperspectral image*. in *Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X*. 2004. International Society for Optics and Photonics.
175. Nascimento, V.B., et al., *Surface Geometric and Electronic Structures of $BaFe_2As_2(001)$* . Physical Review Letters, 2009. **103**(7): p. 076104.
176. Eley, S., et al., *Approaching zero-temperature metallic states in mesoscopic superconductor-normal-superconductor arrays*. Nature Physics, 2011. **8**: p. 59.
177. Xu, Y.M., et al., *Fermi surface dichotomy of the superconducting gap and pseudogap in underdoped pnictides*. Nature Communications, 2011. **2**: p. 392.
178. Ahilan, K., et al., *^{19}F NMR investigation of the iron pnictide superconductor $LaFeAsO_{0.89}F_{0.11}$* . Physical Review B, 2008. **78**(10): p. 100501.
179. Mertelj, T., et al., *Distinct Pseudogap and Quasiparticle Relaxation Dynamics in the Superconducting State of Nearly Optimally Doped $SmFeAsO_{0.8}F_{0.2}$ Single Crystals*. Physical Review Letters, 2009. **102**(11): p. 117002.
180. Zhou, X., et al., *Evolution from Unconventional Spin Density Wave to Superconductivity and a Pseudogaplike Phase in $NaFe_{1-x}Co_xAs$* . Physical Review Letters, 2012. **109**(3): p. 037002.
181. Hoffman, J.E., *A Search for Alternative Electronic Order in the High Temperature Superconductor $Bi_2Sr_2CaCu_2O_{+[\delta]}$ by Scanning Tunneling Microscopy*. 2003, Citeseer.
182. Vasudevan, R.K., et al., *Big data in reciprocal space: Sliding fast Fourier transforms for determining periodicity*. Applied Physics Letters, 2015. **106**(9): p. 091601.

183. Pennycook, S., *Z-contrast transmission electron microscopy: direct atomic imaging of materials*. Annual review of materials science, 1992. **22**(1): p. 171-195.
184. Li, Z., et al., *Three-dimensional atomic-scale structure of size-selected gold nanoclusters*. Nature, 2008. **451**(7174): p. 46.
185. LeBeau, J.M., et al., *Standardless atom counting in scanning transmission electron microscopy*. Nano Letters, 2010. **10**(11): p. 4405-4408.
186. Jesse, S. and S.V. Kalinin, *Principal component and spatial correlation analysis of spectroscopic-imaging data in scanning probe microscopy*. Nanotechnology, 2009. **20**(8): p. 085714.
187. Ren, J., et al., *Electronic origin of the low-symmetry scanning tunneling microscopy image of the layered transition-metal halide. alpha.-ruthenium (III) chloride*. The Journal of Physical Chemistry, 1993. **97**(18): p. 4764-4768.
188. Grimme, S., *Semiempirical GGA-type density functional constructed with a long-range dispersion correction*. Journal of computational chemistry, 2006. **27**(15): p. 1787-1799.
189. Zhou, Y., et al., *Evidencing the existence of exciting half-metallicity in two-dimensional TiCl₃ and VCl₃ sheets*. Scientific reports, 2016. **6**: p. 19407.
190. Nasu, J., M. Udagawa, and Y. Motome, *Vaporization of Kitaev spin liquids*. Physical review letters, 2014. **113**(19): p. 197205.
191. Nasu, J., et al., *Fermionic response from fractionalization in an insulating two-dimensional magnet*. Nature Physics, 2016. **12**(10): p. 912.
192. Kass, R.E., *The Geometry of Asymptotic Inference*. Stat. Sci., 1989. **4**(3): p. 188-234.
193. Jin, R., et al., *Electronic, magnetic and optical properties of two Fe-based superconductors and related parent compounds*. Superconductor Science and Technology, 2010. **23**(5): p. 054005.
194. Ziatdinov, M., et al., *Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations*. ACS Nano, 2017. **11**(12): p. 12742-12752.
195. Pennycook, S.J. and P.D. Nellist, *Scanning Transmission Electron Microscopy*, ed. S.J. Pennycook and P.D. Nellist. 2011: Springer-Verlag New York.
196. Moore, A.M. and P.S. Weiss, *Functional and Spectroscopic Measurements with Scanning Tunneling Microscopy*. Annual Review of Analytical Chemistry, 2008. **1**(1): p. 857-882.
197. Zandvliet, H.J.W. and A. van Houselt, *Scanning Tunneling Spectroscopy*. Annual Review of Analytical Chemistry, 2009. **2**(1): p. 37-55.
198. Muller, D.A., *Structure and bonding at the atomic scale by scanning transmission electron microscopy*. Nat Mater, 2009. **8**(4): p. 263-270.
199. Mishra, R., et al., *Oxygen-Vacancy-Induced Polar Behavior in (LaFeO₃)₂/(SrFeO₃) Superlattices*. Nano Letters, 2014. **14**(5): p. 2694-2701.
200. Yankovich, A.B., et al., *Picometre-precision analysis of scanning transmission electron microscopy images of platinum nanocatalysts*. 2014. **5**: p. 4155.

201. Cantoni, C., et al., *Orbital Occupancy and Charge Doping in Iron-Based Superconductors*. Advanced Materials, 2014. **26**(35): p. 6193-6198.
202. Li, X., et al., *Suppression of Defects and Deep Levels Using Isoelectronic Tungsten Substitution in Monolayer MoSe₂*. Advanced Functional Materials, 2017. **27**(19): p. 1603850-n/a.
203. Li, Y., et al., *Complex structural dynamics of nanocatalysts revealed in Operando conditions by correlated imaging and spectroscopy probes*. 2015. **6**: p. 7583.
204. Chi, M., et al., *Surface faceting and elemental diffusion behaviour at atomic scale for alloy nanoparticles during in situ annealing*. 2015. **6**: p. 8925.
205. Jesse, S., et al., *Atomic-Level Sculpting of Crystalline Oxides: Toward Bulk Nanofabrication with Single Atomic Plane Precision*. Small, 2015. **11**(44): p. 5895-5900.
206. Zhang, W.M., et al., *Controllable shrinking and shaping of silicon nitride nanopores under electron irradiation*. Applied Physics Letters, 2007. **90**(16): p. 163102.
207. Ishikawa, R., et al., *Direct Observation of Dopant Atom Diffusion in a Bulk Semiconductor Crystal Enhanced by a Large Size Mismatch*. Physical Review Letters, 2014. **113**(15): p. 155501.
208. Yang, Z., et al., *Direct Observation of Atomic Dynamics and Silicon Doping at a Topological Defect in Graphene*. Angewandte Chemie, 2014. **126**(34): p. 9054-9058.
209. Kalinin, S.V. and S.J. Pennycook, *Single-atom fabrication with electron and ion beams: From surfaces and two-dimensional materials toward three-dimensional atom-by-atom assembly*. MRS Bulletin, 2017. **42**(9): p. 637-643.
210. Dyck, O., et al., *Placing single atoms in graphene with a scanning transmission electron microscope*. Applied Physics Letters, 2017. **111**(11): p. 113104.
211. Kalinin, S.V., Borisevich, A., Jesse, S., *Fire up the atom forge*. Nature, 2016. **539**(7630): p. 485-487.
212. de Oteyza, D.G., et al., *Direct Imaging of Covalent Bond Structure in Single-Molecule Chemical Reactions*. Science, 2013. **340**(6139): p. 1434.
213. Wang, Y., et al., *Observing Atomic Collapse Resonances in Artificial Nuclei on Graphene*. Science, 2013. **340**(6133): p. 734.
214. Zeljkovic, I. and J.E. Hoffman, *Interplay of chemical disorder and electronic inhomogeneity in unconventional superconductors*. Physical Chemistry Chemical Physics, 2013. **15**(32): p. 13462-13478.
215. Zhang, Y., et al., *Mechanochemistry Induced Using Force Exerted by a Functionalized Microscope Tip*. Angewandte Chemie International Edition, 2017. **56**(39): p. 11769-11773.
216. Komeda, T., et al., *Lateral Hopping of Molecules Induced by Excitation of Internal Vibration Mode*. Science, 2002. **295**(5562): p. 2055.
217. Sommer, C. and D.W. Gerlich, *Machine learning in cell biology – teaching computers to recognize phenotypes*. Journal of Cell Science, 2013. **126**(24): p. 5529.

218. de Bruijne, M., *Machine learning approaches in medical image analysis: From detection to diagnosis*. Medical Image Analysis, 2016. **33**(Supplement C): p. 94-97.
219. LeCun, Y., Y. Bengio, and G. Hinton, *Deep learning*. Nature, 2015. **521**(7553): p. 436-444.
220. Rosenblatt, F., *The perception: a probabilistic model for information storage and organization in the brain*. Psychological Review, 1958. **65**(6): p. 386-408.
221. Carrasquilla, J. and R.G. Melko, *Machine learning phases of matter*. Nat Phys, 2017. **13**(5): p. 431-434.
222. van Nieuwenburg, E.P.L., Y.-H. Liu, and S.D. Huber, *Learning phase transitions by confusion*. Nat Phys, 2017. **13**(5): p. 435-439.
223. Badrinarayanan, V., A. Kendall, and R. Cipolla, *SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation*. CoRR, 2015. **abs/1511.00561**.
224. Shelhamer, E., J. Long, and T. Darrell, *Fully Convolutional Networks for Semantic Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015. **39**: p. 640-651.
225. Laanait, N., et al., *Identifying local structural states in atomic imaging by computer vision*. Advanced Structural and Chemical Imaging, 2016. **2**(1): p. 14.
226. Banhart, F., J. Kotakoski, and A.V. Krasheninnikov, *Structural Defects in Graphene*. ACS Nano, 2011. **5**(1): p. 26-41.
227. Zhou, W., et al., *Direct Determination of the Chemical Bonding of Individual Impurities in Graphene*. Physical Review Letters, 2012. **109**(20): p. 206803.
228. Ramasse, Q.M., et al., *Probing the Bonding and Electronic Structure of Single Atom Dopants in Graphene with Electron Energy Loss Spectroscopy*. Nano Letters, 2013. **13**(10): p. 4989-4995.
229. Susi, T., et al., *Computational insights and the observation of SiC nanograin assembly: towards 2D silicon carbide*. Scientific Reports, 2017. **7**(1): p. 4399.
230. Vlassiuk, I., et al., *Large scale atmospheric pressure chemical vapor deposition of graphene*. Carbon, 2013. **54**(Supplement C): p. 58-67.
231. Garcia, A.G.F., et al., *Effective Cleaning of Hexagonal Boron Nitride for Graphene Devices*. Nano Letters, 2012. **12**(9): p. 4449-4454.
232. Dyck, O., et al., *Mitigating E-beam-induced Hydrocarbon Deposition on Graphene for Atomic-Scale Scanning Transmission Electron Microscopy Studies*. arXiv preprint arXiv:1709.00470, 2017.
233. Li, X., et al., *Isoelectronic Tungsten Doping in Monolayer MoSe₂ for Carrier Type Modulation*. Advanced Materials, 2016. **28**(37): p. 8240-8247.
234. Kingma, D.P. and J. Ba, *Adam: A Method for Stochastic Optimization*. CoRR, 2014. **abs/1412.6980**.
235. Ziatdinov, M., et al., *Deep analytics of atomically-resolved images: manifest and latent features*. arXiv preprint arXiv:1801.05133, 2018.
236. Ishikawa, R., et al., *Single atom visibility in STEM optical depth sectioning*. Applied Physics Letters, 2016. **109**(16): p. 163102.

237. Wei, J., et al., *Strong Coupling between ZnO Excitons and Localized Surface Plasmons of Silver Nanoparticles Studied by STEM-EELS*. Nano Letters, 2015. **15**(9): p. 5926-5931.
238. Tizei, L.H.G., et al., *Exciton Mapping at Subwavelength Scales in Two-Dimensional Materials*. Physical Review Letters, 2015. **114**(10): p. 107601.
239. Collins, L., et al., *G-mode magnetic force microscopy: Separating magnetic and electrostatic interactions using big data analytics*. Applied Physics Letters, 2016. **108**(19): p. 193103.
240. Kholkin, A., et al., *Review of ferroelectric domain imaging by piezoresponse force microscopy*, in *Scanning probe microscopy*. 2007, Springer. p. 173-214.
241. Stroh, C., et al., *Single-molecule recognition imaging microscopy*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(34): p. 12503-12507.
242. Kazuma, E., et al., *Direct Pathway to Molecular Photodissociation on Metal Surfaces Using Visible Light*. Journal of the American Chemical Society, 2017. **139**(8): p. 3115-3121.
243. Dyck, O., et al., *Single atom manipulation and control in a scanning transmission electron microscope*. arXiv preprint arXiv:1708.01523, 2017.
244. Susi, T., J.C. Meyer, and J. Kotakoski, *Manipulating low-dimensional materials down to the level of single atoms with electron irradiation*. Ultramicroscopy, 2017. **180**: p. 163-172.
245. Susi, T., et al., *Electron-beam manipulation of Si dopants in graphene*. arXiv preprint arXiv:1712.08755, 2017.
246. Dyck, O., et al., *Assembling Di-and Multiatomic Si Clusters in Graphene via Electron Beam Manipulation*. arXiv preprint arXiv:1710.09416, 2017.
247. Zhou, B., et al. *Learning deep features for discriminative localization*. in *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*. 2016. IEEE.
248. Artetxe, M., et al., *Unsupervised neural machine translation*. arXiv preprint arXiv:1710.11041, 2017.
249. Long, J., E. Shelhamer, and T. Darrell. *Fully convolutional networks for semantic segmentation*. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.
250. Badrinarayanan, V., A. Handa, and R. Cipolla, *Segnet: A deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling*. arXiv preprint arXiv:1505.07293, 2015.
251. Kishida, T., et al., *Magnetic Ordering in Sr₃YCo₄O_{10+x}*. Scientific reports, 2016. **6**: p. 19762.
252. Tokura, Y., *Critical features of colossal magnetoresistive manganites*. Reports on Progress in Physics, 2006. **69**(3): p. 797.
253. Renner, C., et al., *Atomic-scale images of charge ordering in a mixed-valence manganite*. Nature, 2002. **416**(6880): p. 518.
254. Tselev, A., et al., *Surface control of epitaxial manganite films via oxygen pressure*. ACS nano, 2015. **9**(4): p. 4316-4327.

255. Fujii, S., et al., *Bowl Inversion and Electronic Switching of Buckybowls on Gold*. Journal of the American Chemical Society, 2016. **138**(37): p. 12142-12149.
256. Kingma, D.P. and J. Ba, *Adam: A method for stochastic optimization*. arXiv preprint arXiv:1412.6980, 2014.
257. Barthel, J., *Dr. Probe - High-resolution (S)TEM image simulation software*. 2017.
258. Keen, D.A. and A.L. Goodwin, *The crystallography of correlated disorder*. Nature, 2015. **521**(7552): p. 303-309.
259. Overly, A.R., et al., *Design of crystal-like aperiodic solids with selective disorder-phonon coupling*. Nature Communications, 2016. **7**: p. 10445.
260. Petrović, A.P., et al., *A disorder-enhanced quasi-one-dimensional superconductor*. Nature Communications, 2016. **7**: p. 12262.
261. Guo, H., et al., *Strain Doping: Reversible Single-Axis Control of a Complex Oxide Lattice via Helium Implantation*. Physical Review Letters, 2015. **114**(25): p. 256801.
262. Bennett, T.D., et al., *Interplay between defects, disorder and flexibility in metal-organic frameworks*. Nat Chem, 2017. **9**(1): p. 11-16.
263. Kalinin, S.V. and S.J. Pennycook, *Microscopy: Hasten high resolution*. Nature, 2014. **515**.
264. Freysoldt, C., et al., *First-principles calculations for point defects in solids*. Reviews of Modern Physics, 2014. **86**(1): p. 253-305.
265. Rabe, K.M., *First-Principles Calculations of Complex Metal-Oxide Materials*. Annual Review of Condensed Matter Physics, 2010. **1**(1): p. 211-235.
266. Friesner, R.A., *Ab initio quantum chemistry: Methodology and applications*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(19): p. 6648-6653.
267. Jia, C.L., et al., *Atomic-scale study of electric dipoles near charged and uncharged domain walls in ferroelectric films*. Nature Materials, 2008. **7**(1): p. 57-61.
268. Jia, C.L., et al., *Unit-cell scale mapping of ferroelectricity and tetragonality in epitaxial ultrathin ferroelectric films*. Nature Materials, 2007. **6**(1): p. 64-69.
269. Nelson, C.T., et al., *Spontaneous Vortex Nanodomain Arrays at Ferroelectric Heterointerfaces*. Nano Letters, 2011. **11**(2): p. 828-834.
270. Sakurai, H., T. Daiko, and T. Hirao, *A Synthesis of Sumanene, a Fullerene Fragment*. Science, 2003. **301**(5641): p. 1878.
271. Jaafar, R., et al., *Bowl Inversion of Surface-Adsorbed Sumanene*. Journal of the American Chemical Society, 2014. **136**(39): p. 13666-13671.
272. Stephen, J. and V.K. Sergei, *Principal component and spatial correlation analysis of spectroscopic-imaging data in scanningB probe microscopy*. Nanotechnology, 2009. **20**(8): p. 085714.
273. Dmitry, A.O., et al., *The manipulation of C 60 in molecular arrays with an STM tip in regimes below the decomposition threshold*. Nanotechnology, 2013. **24**(5): p. 055302.

274. Cross, G.R. and A.K. Jain, *Markov Random Field Texture Models*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1983. **PAMI-5**(1): p. 25-39.
275. Blake, A., P. Kohli, and C. Rother, *Markov Random Fields for Vision and Image Processing*. 2011: The MIT Press. 480.
276. Weiss, Y., *Comparing the mean field method and belief propagation for approximate inference in MRFs*. Advanced mean field methods: theory and practice, 2001: p. 229-240.
277. Murphy, K.P., Y. Weiss, and M.I. Jordan, *Loopy belief propagation for approximate inference: an empirical study*, in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. 1999, Morgan Kaufmann Publishers Inc.: Stockholm, Sweden. p. 467-475.
278. Schmidt, M., *UGM: A Matlab toolbox for probabilistic undirected graphical models*. <http://www.cs.ubc.ca/~schmidtm/Software/UGM.html>, 2007.
279. Nielsen, M.A., *Neural Networks and Deep Learning*. 2015: Determination Press.
280. Jean, N., et al., *Combining satellite imagery and machine learning to predict poverty*. Science, 2016. **353**(6301): p. 790.
281. Litjens, G., et al., *Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis*. Scientific Reports, 2016. **6**: p. 26286.
282. Cyphersmith, A., et al., *Defocused Emission Patterns from Chiral Fluorophores: Application to Chiral Axis Orientation Determination*. The Journal of Physical Chemistry Letters, 2011. **2**(6): p. 661-665.
283. Palm, R.B., *Prediction as a candidate for learning deep hierarchical models of data*. 2012, Technical University of Denmark.
284. Wachowiak, A., et al., *Visualization of the Molecular Jahn-Teller Effect in an Insulating K_4C_{60} Monolayer*. Science, 2005. **310**(5747): p. 468.
285. Lu, X., et al., *Charge transfer and screening in individual C_{60} molecules on metal substrates: A scanning tunneling spectroscopy and theoretical study*. Physical Review B, 2004. **70**(11): p. 115418.
286. Amara, H., et al., *Scanning tunneling microscopy fingerprints of point defects in graphene: A theoretical prediction*. Physical Review B, 2007. **76**(11): p. 115423.
287. El-Barbary, A.A., et al., *Structure and energetics of the vacancy in graphite*. Physical Review B, 2003. **68**(14): p. 144107.
288. Getis, A. and J.K. Ord, *The analysis of spatial association by use of distance statistics*. Geographical analysis, 1992. **24**(3): p. 189-206.
289. Anselin, L., *Local Indicators of Spatial Association—LISA*. Geographical Analysis, 1995. **27**(2): p. 93-115.
290. Ganin, Y., et al. *Domain-Adversarial Training of Neural Networks*. ArXiv e-prints, 2015. **1505**.
291. Matsumoto, T., et al., *Direct observation of $N=7$ domain boundary core structure in magnetic skyrmion lattice*. Science Advances, 2016. **2**(2).
292. Smolensky, G., G. Smolensky and AI Agranovskaya, *Sov. Phys. Solid State 1, 1429 (1959)*. Sov. Phys. Solid State, 1959. **1**: p. 1429.
293. Cross, L.E., *Relaxor ferroelectrics*. Ferroelectrics, 1987. **76**(1): p. 241-267.

294. Burton, B., et al., *The Polar Nano Regions Relaxor Transition in $Pb_{1-x}Sc_{1/2}Nb_{1/2}O_{3-x}$; X = bulk concentration of nearest neighbor [Pb-O] divacancies*. arXiv preprint arXiv:1612.00398, 2016.
295. Hinton, G.E. and R.R. Salakhutdinov, *Reducing the dimensionality of data with neural networks*. science, 2006. **313**(5786): p. 504-507.
296. Ngiam, J., et al. *Multimodal deep learning*. in *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011.
297. Kingma, D.P. and M. Welling, *Auto-encoding variational bayes*. arXiv preprint arXiv:1312.6114, 2013.
298. Pu, Y., et al. *Variational autoencoder for deep learning of images, labels and captions*. in *Advances in neural information processing systems*. 2016.
299. Goldberger, J., S. Gordon, and H. Greenspan. *An efficient image similarity measure based on approximations of KL-divergence between two Gaussian mixtures*. in null. 2003. IEEE.
300. Wetzel, S.J., *Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders*. Physical Review E, 2017. **96**(2): p. 022140.
301. Nair, V. and G.E. Hinton. *Rectified linear units improve restricted boltzmann machines*. in *Proceedings of the 27th international conference on machine learning (ICML-10)*. 2010.
302. Ioffe, S. and C. Szegedy, *Batch normalization: Accelerating deep network training by reducing internal covariate shift*. arXiv preprint arXiv:1502.03167, 2015.
303. Srivastava, N., et al., *Dropout: A simple way to prevent neural networks from overfitting*. The Journal of Machine Learning Research, 2014. **15**(1): p. 1929-1958.
304. Tersoff, J. and D. Hamann, *Theory and application for the scanning tunneling microscope*. Physical review letters, 1983. **50**(25): p. 1998.
305. Tersoff, J. and D. Hamann, *Theory of the scanning tunneling microscope*. Physical Review B, 1985. **31**(2): p. 805.
306. Giannozzi, P., et al., *QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials*. Journal of physics: Condensed matter, 2009. **21**(39): p. 395502.
307. Murphy, K.P., Y. Weiss, and M.I. Jordan. *Loopy belief propagation for approximate inference: An empirical study*. in *Proceedings of the Fifteenth conference on Uncertainty in artificial intelligence*. 1999. Morgan Kaufmann Publishers Inc.
308. Schmidt, M., *UGM: Matlab code for undirected graphical models*. URL <http://www.di.ens.fr/mschmidt/Software/UGM.html>, 2012.
309. Jiang, N., et al., *Atom-by-atom fabrication by electron beam via induced phase transformations*. MRS Bulletin, 2017. **42**(9): p. 653-659.
310. Jesse, S., et al., *Direct atomic fabrication and dopant positioning in Si using electron beams with active real-time image-based feedback*. Nanotechnology, 2018. **29**(25): p. 255303.

311. Hoddeson, M.R.a.L., *Crystal Fire: The Invention of the Transistor and the Birth of the Information Age*. 1998: W. W. Norton & Company.
312. Gertner, J., *The Idea Factory: Bell Labs and the Great Age of American Innovation*. 2013: Penguin Books.
313. Asenov, A., *Random dopant induced threshold voltage lowering and fluctuations in sub 50 nm MOSFETs: a statistical 3D atomistic simulation study*. *Nanotechnology*, 1999. **10**(2): p. 153.
314. Carballo, J.-A., et al. *ITRS 2.0: Toward a re-framing of the Semiconductor Technology Roadmap*. in *Computer Design (ICCD), 2014 32nd IEEE International Conference on*. 2014. IEEE.
315. Keyes, R.W., *Physical limits of silicon transistors and circuits*. *Reports on Progress in Physics*, 2005. **68**(12): p. 2701.
316. Schofield, S.R., et al., *Atomically precise placement of single dopants in Si*. *Physical Review Letters*, 2003. **91**(13).
317. Fuechsle, M., et al., *A single-atom transistor*. *Nature Nanotechnology*, 2012. **7**(4): p. 242-246.
318. Weber, B., et al., *Ohm's Law Survives to the Atomic Scale*. *Science*, 2012. **335**(6064): p. 64-67.
319. Lee, G., et al., *Electron beam induced epitaxial crystallization in a conducting and insulating α -LaAlO₃/SrTiO₃ system*. *RSC Advances*, 2017. **7**(64): p. 40279-40285.
320. Jesse, S., He, Q., Lupinin, A.R., Leonard, D., Oxley, M.P., Ovchinnikov, O., Unocic, R.R., Tselev, A., Fuentes-Cabrera, M., Sumpter, B.G., Pennycook, S.J., Kalinin, S.V., and Borisevich, A.Y., *Atomic-Level Sculpting of Crystalline Oxides: Toward Bulk Nanofabrication with Single Atomic Plane Precision*. *Small*, 2015: p. 1.
321. Jang, J.H., et al., *In Situ Observation of Oxygen Vacancy Dynamics and Ordering in the Epitaxial LaCoO₃ System*. *ACS Nano*, 2017.
322. Kurasch, S., et al., *Atom-by-Atom Observation of Grain Boundary Migration in Graphene*. *Nano Letters*, 2012. **12**(6): p. 3168-3173.
323. Ishikawa, R., et al., *Direct Observation of Dopant Atom Diffusion in a Bulk Semiconductor Crystal Enhanced by a Large Size Mismatch*. *Physical Review Letters*, 2014. **113**(15).
324. Markevich, A., et al., *Electron beam controlled covalent attachment of small organic molecules to graphene*. *Nanoscale*, 2016. **8**(5): p. 2711-2719.
325. Yang, Z.Q., et al., *Direct Observation of Atomic Dynamics and Silicon Doping at a Topological Defect in Graphene*. *Angewandte Chemie-International Edition*, 2014. **53**(34): p. 8908-8912.
326. Toma Susi, J.C.M., Jani Kotakoski, *Manipulating low-dimensional materials down to the level of single atoms with electron irradiation*. *Ultramicroscopy*, 2017. **180**: p. 163-172.
327. Dai, S., et al., *Electron-Beam-Induced Elastic-Plastic Transition in Si Nanowires*. *Nano Letters*, 2012. **12**(5): p. 2379-2385.

328. Amkreutz, D., et al., *Electron-beam crystallized large grained silicon solar cell on glass substrate*. Progress in Photovoltaics, 2011. **19**(8): p. 937-945.
329. Wang, Z.L., et al., *ION-INDUCED CRYSTALLIZATION AND AMORPHIZATION AT CRYSTAL/AMORPHOUS INTERFACES OF SILICON*. Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms, 1995. **100**(4): p. 493-501.
330. Jencic, I., et al., *ELECTRON-BEAM-INDUCED CRYSTALLIZATION OF ISOLATED AMORPHOUS REGIONS IN SI, GE, GAP, AND GAAS*. Journal of Applied Physics, 1995. **78**(2): p. 974-982.
331. Yang, X.X., et al., *Low energy electron-beam-induced recrystallization of continuous GaAs amorphous foils*. Materials Science and Engineering B-Solid State Materials for Advanced Technology, 1997. **49**(1): p. 5-13.
332. Li, Z.C., et al., *Electron-beam induced nucleation and growth in amorphous GaAs*. Acta Metallurgica Sinica, 2003. **39**(1): p. 13-16.
333. Egerton, R., *Mechanisms of radiation damage in beam-sensitive specimens, for TEM accelerating voltages between 10 and 300 kV*. Microscopy research and technique, 2012. **75**(11): p. 1550-1556.
334. Egerton, R., P. Li, and M. Malac, *Radiation damage in the TEM and SEM*. Micron, 2004. **35**(6): p. 399-409.
335. Seitz, F. and J. Koehler, *Displacement of atoms during irradiation*. Solid State Physics-Advances In Research And Applications, 1956. **2**: p. 305-448.
336. Kaganov, M., I. Lifshitz, and L. Tanatarov, *Relaxation between electrons and the crystalline lattice*. Sov. Phys. JETP, 1957. **4**(2): p. 173-178.
337. Meftah, A., et al., *Track formation in SiO₂ quartz and the thermal-spike mechanism*. Physical Review B, 1994. **49**(18): p. 12457.
338. Wong, B.T. and P.M. Mengüç, *Thermal Transport for Applications in Micro/Nanomachining*. 2008: Springer Science & Business Media.
339. Chettah, A., et al., *Behavior of crystalline silicon under huge electronic excitations: A transient thermal spike description*. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2009. **267**(16): p. 2719-2724.
340. Osmani, O., et al., *Damage in crystalline silicon by swift heavy ion irradiation*. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2012. **282**: p. 43-47.
341. Daraszewicz, S. and D. Duffy, *Extending the inelastic thermal spike model for semiconductors and insulators*. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2011. **269**(14): p. 1646-1649.
342. Khara, G.S., et al., *The influence of the electronic specific heat on swift heavy ion irradiation simulations of silicon*. Journal of Physics: Condensed Matter, 2016. **28**(39): p. 395201.
343. Akkerman, A. and M. Murat, *Electron-phonon interactions in silicon: Mean free paths, related distributions and transport characteristics*. Nuclear Instruments

- and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 2015. **350**: p. 49-54.
344. Bell, R., M. Toulemonde, and P. Siffert, *Calculated temperature distribution during laser annealing in silicon and cadmium telluride*. Applied Physics A: Materials Science & Processing, 1979. **19**(3): p. 313-319.
 345. Ganesh, P. and M. Widom, *Liquid-Liquid Transition in Supercooled Silicon Determined by First-Principles Simulation*. Physical Review Letters, 2009. **102**(7): p. 075701.
 346. Hartigan, J.A. and M.A. Wong, *Algorithm AS 136: A k-means clustering algorithm*. Journal of the Royal Statistical Society. Series C (Applied Statistics), 1979. **28**(1): p. 100-108.
 347. Kovarik, L., et al., *Implementing an accurate and rapid sparse sampling approach for low-dose atomic resolution STEM imaging*. Applied Physics Letters, 2016. **109**(16).
 348. Sastry, S. and C. Austen Angell, *Liquid-liquid phase transition in supercooled silicon*. Nat Mater, 2003. **2**(11): p. 739-743.
 349. Plimpton, S., *Fast parallel algorithms for short-range molecular dynamics*. Journal of computational physics, 1995. **117**(1): p. 1-19.
 350. Plimpton, S., P. Crozier, and A. Thompson, *LAMMPS-large-scale atomic/molecular massively parallel simulator*. Sandia National Laboratories, 2007. **18**.
 351. Yang, Y., et al., *Dielectric meta-reflectarray for broadband linear polarization conversion and optical vortex generation*. Nano letters, 2014. **14**(3): p. 1394-1399.
 352. Owen, J.H.G., et al., *One-dimensional Si-in-Si(001) template for single-atom wire growth*. Applied Physics Letters, 2010. **97**(9).
 353. Maksov, A., et al., *Deep Learning Analysis of Defect and Phase Evolution During Electron Beam Induced Transformations in WS₂*. arXiv preprint arXiv:1803.05381, 2018.
 354. Attygalle, A.B., et al., *Gas-Phase Infrared Spectroscopy for Determination of Double Bond Configuration of Monounsaturated Compounds*. Analytical Chemistry, 1994. **66**(10): p. 1696-1703.
 355. Iskra, A., et al., *Infrared Spectroscopy of Gas-Phase M+(CO₂)_n (M = Co, Rh, Ir) Ion–Molecule Complexes*. The Journal of Physical Chemistry A, 2017. **121**(1): p. 133-140.
 356. Holman, R.W., et al., *Mass spectrometry for investigations of gas-phase radical cation chemistry: The two step cycloaddition of the benzene radical cation and 1,3-butadiene*. Tetrahedron, 1986. **42**(22): p. 6235-6244.
 357. Jameson, C.J., *Gas-phase NMR spectroscopy*. Chemical Reviews, 1991. **91**(7): p. 1375-1395.
 358. Frank, I., M. Parrinello, and A. Klamt, *Insight into Chemical Reactions from First-Principles Simulations: The Mechanism of the Gas-Phase Reaction of OH Radicals with Ketones*. The Journal of Physical Chemistry A, 1998. **102**(20): p. 3614-3617.

359. Snyder, J.A., et al., *A Density Functional Theory Study of the Gas-Phase Hydrolysis of Dinitrogen Pentoxide*. The Journal of Physical Chemistry A, 1999. **103**(46): p. 9355-9358.
360. Frank, J., *Time-resolved cryo-electron microscopy: Recent progress*. Journal of Structural Biology, 2017. **200**(3): p. 303-306.
361. Pande, K., et al., *Femtosecond structural dynamics drives the trans/cis isomerization in photoactive yellow protein*. Science, 2016. **352**(6286): p. 725.
362. Barmak, K., et al., *Ex situ characterization of phase transformations and associated microstructures in polycrystalline thin films*. Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, 1999. **17**(4): p. 1950-1957.
363. Lucadamo, G., et al., *Microstructure evolution during solid state reactions of Nb/Al multilayers*. Acta Materialia, 2001. **49**(14): p. 2813-2826.
364. Suresh, S., *Colloid model for atoms*. Nature Materials, 2006. **5**: p. 253.
365. Taheri, M.L., et al., *Current status and future directions for in situ transmission electron microscopy*. Ultramicroscopy, 2016. **170**: p. 86-95.
366. Mishra, R., et al., *Single-atom dynamics in scanning transmission electron microscopy*. MRS Bulletin, 2017. **42**(9): p. 644-652.
367. Zhao, X., et al., *Engineering and modifying two-dimensional materials by electron beams*. MRS Bulletin, 2017. **42**(9): p. 667-676.
368. Susi, T., et al., *Towards atomically precise manipulation of 2D nanostructures in the electron microscope*. 2D Materials, 2017. **4**(4): p. 042004.
369. Zheng, H. and Y. Zhu, *Perspectives on in situ electron microscopy*. Ultramicroscopy, 2017. **180**: p. 188-196.
370. Jesse, S., et al., *Directing Matter: Toward Atomic-Scale 3D Nanofabrication*. ACS Nano, 2016. **10**(6): p. 5600-5618.
371. Gonzalez-Martinez, I., et al., *Electron-beam induced synthesis of nanostructures: a review*. Nanoscale, 2016. **8**(22): p. 11340-11362.
372. He, Z., et al., *Atomic Structure and Dynamics of Metal Dopant Pairs in Graphene*. Nano Letters, 2014. **14**(7): p. 3766-3772.
373. Robertson, A.W., et al., *Stability and Dynamics of the Tetravacancy in Graphene*. Nano Letters, 2014. **14**(3): p. 1634-1642.
374. Robertson, A.W., et al., *Dynamics of Single Fe Atoms in Graphene Vacancies*. Nano Letters, 2013. **13**(4): p. 1468-1475.
375. Lee, J., et al., *Direct visualization of reversible dynamics in a Si₆ cluster embedded in a graphene pore*. 2013. **4**: p. 1650.
376. Kotakoski, J., et al., *Stone-Wales-type transformations in carbon nanostructures driven by electron irradiation*. Physical Review B, 2011. **83**(24): p. 245420.
377. Kotakoski, J., C. Mangler, and J.C. Meyer, *Imaging atomic-level random walk of a point defect in graphene*. Nature Communications, 2014. **5**: p. 3991.
378. Kotakoski, J., et al., *From Point Defects in Graphene to Two-Dimensional Amorphous Carbon*. Physical Review Letters, 2011. **106**(10): p. 105505.

379. Susi, T., et al., *Atomistic Description of Electron Beam Damage in Nitrogen-Doped Graphene and Single-Walled Carbon Nanotubes*. ACS Nano, 2012. **6**(10): p. 8837-8846.
380. Lehtinen, O., et al., *Atomic scale study of the life cycle of a dislocation in graphene from birth to annihilation*. Nature Communications, 2013. **4**: p. 2098.
381. Robertson, A.W., et al., *Partial Dislocations in Graphene and Their Atomic Level Migration Dynamics*. Nano Letters, 2015. **15**(9): p. 5950-5955.
382. Warner, J.H., et al., *Dislocation-Driven Deformations in Graphene*. Science, 2012. **337**(6091): p. 209-212.
383. Zhao, J., et al., *Free-Standing Single-Atom-Thick Iron Membranes Suspended in Graphene Pores*. Science, 2014. **343**(6176): p. 1228-1232.
384. Quang, H.T., et al., *In Situ Observations of Free-Standing Graphene-like Mono- and Bilayer ZnO Membranes*. ACS Nano, 2015. **9**(11): p. 11408-11413.
385. Wang, H., et al., *Unraveling the Atomic Structure of Ultrafine Iron Clusters*. Scientific Reports, 2012. **2**: p. 995.
386. Ta, H.Q., et al., *Single Cr atom catalytic growth of graphene*. Nano Research, 2017.
387. Zhao, J., et al., *Direct in situ observations of single Fe atom catalytic processes and anomalous diffusion at graphene edges*. Proceedings of the National Academy of Sciences, 2014. **111**(44): p. 15641-15646.
388. Wang, W.L., et al., *Direct Observation of a Long-Lived Single-Atom Catalyst Chiseling Atomic Structures in Graphene*. Nano Letters, 2014. **14**(2): p. 450-455.
389. Ramasse, Q.M., et al., *Direct Experimental Evidence of Metal-Mediated Etching of Suspended Graphene*. ACS Nano, 2012. **6**(5): p. 4063-4071.
390. Zan, R., et al., *Metal-Graphene Interaction Studied via Atomic Resolution Scanning Transmission Electron Microscopy*. Nano Letters, 2011. **11**(3): p. 1087-1092.
391. Dorp, W.F.v., et al., *Nanometer-scale lithography on microscopically clean graphene*. Nanotechnology, 2011. **22**(50): p. 505303.
392. Zan, R., et al., *Graphene Reknits Its Holes*. Nano Letters, 2012. **12**(8): p. 3936-3940.
393. Lin, J., et al., *Flexible metallic nanowires with self-adaptive contacts to semiconducting transition-metal dichalcogenide monolayers*. Nature Nanotechnology, 2014. **9**: p. 436.
394. Jesse, S., et al., *Direct atomic fabrication and dopant positioning in Si using electron beams with active real time image-based feedback*. arXiv preprint arXiv:1711.05810, 2017.
395. Chuvilin, A., et al., *Direct transformation of graphene to fullerene*. Nature Chemistry, 2010. **2**: p. 450.
396. Girit, Ç.Ö., et al., *Graphene at the Edge: Stability and Dynamics*. Science, 2009. **323**(5922): p. 1705.
397. Kotakoski, J., D. Santos-Cottin, and A.V. Krashenninnikov, *Stability of Graphene Edges under Electron Beam: Equilibrium Energetics versus Dynamic Effects*. ACS Nano, 2012. **6**(1): p. 671-676.

398. Susi, T., et al., *Silicon-Carbon Bond Inversions Driven by 60-keV Electrons in Graphene*. Physical Review Letters, 2014. **113**(11): p. 115501.
399. Eigler, D.M. and E.K. Schweizer, *Positioning single atoms with a scanning tunnelling microscope*. Nature, 1990. **344**(6266): p. 524-526.
400. Kalinin, S.V., A. Borisevich, and S. Jesse, *Fire up the atom forge*. Nature, 2016. **539**(7630): p. 485-487.
401. Long, J., E. Shelhamer, and T. Darrell, *Fully convolutional networks for semantic segmentation*, in *Computer Vision and Pattern Recognition (CVPR)*. 2015, IEEE: Boston, MA, USA.
402. Wenzhi, L., et al., *Local crystallography analysis for atomically resolved scanning tunneling microscopy images*. Nanotechnology, 2013. **24**(41): p. 415707.
403. McLachlan, G. and D. Peel, *Finite Mixture Models*. 2000: John Wiley & Sons, Inc.
404. Azizi, A., et al., *Defect Coupling and Sub-Angstrom Structural Distortions in W_{1-x}MoxS₂ Monolayers*. Nano Letters, 2017. **17**(5): p. 2802-2808.
405. Ievlev, A.V., et al., *Quantitative Description of Crystal Nucleation and Growth from in Situ Liquid Scanning Transmission Electron Microscopy*. ACS Nano, 2015. **9**(12): p. 11784-11791.
406. Rushton, M.J.D. and A. Choneos, *Impact of uniaxial strain and doping on oxygen diffusion in CeO₂*. Scientific Reports, 2014. **4**: p. 6068.
407. Li, W. and J. Li, *Ferroelasticity and domain physics in two-dimensional transition metal dichalcogenide monolayers*. Nature Communications, 2016. **7**: p. 10843.
408. Arlia, D. and M. Coppola. *Experiments in Parallel Clustering with DBSCAN*. in *Euro-Par 2001 Parallel Processing*. 2001. Berlin, Heidelberg: Springer Berlin Heidelberg.
409. Vinyals, O., et al., *Starcraft ii: A new challenge for reinforcement learning*. arXiv preprint arXiv:1708.04782, 2017.
410. Bansal, T., et al., *Emergent complexity via multi-agent competition*. arXiv preprint arXiv:1710.03748, 2017.
411. Granter, S.R., A.H. Beck, and D.J. Papke Jr, *AlphaGo, deep learning, and the future of the human microscopist*. Archives of pathology & laboratory medicine, 2017. **141**(5): p. 619-621.
412. Dyck, O., et al., *Single atom force measurements: mapping potential energy landscapes via electron beam induced single atom dynamics*. arXiv preprint arXiv:1804.03729, 2018.
413. Garnelo, M., et al., *Neural Processes*. arXiv preprint arXiv:1807.01622, 2018.
414. Garnelo, M., et al., *Conditional Neural Processes*. arXiv preprint arXiv:1807.01613, 2018.
415. Battaglia, P.W., et al., *Relational inductive biases, deep learning, and graph networks*. arXiv preprint arXiv:1806.01261, 2018.
416. Duvenaud, D.K., et al. *Convolutional networks on graphs for learning molecular fingerprints*. in *Advances in neural information processing systems*. 2015.

417. Belsky, A., et al., *New developments in the Inorganic Crystal Structure Database (ICSD): accessibility in support of materials research and design*. Acta Crystallographica Section B: Structural Science, 2002. **58**(3): p. 364-369.

VITA

Artem Borisovich Maksov was born in Baku, USSR in 1988 to Boris Igorevich Maksov and Marina Vladimirovna Arzumanova and moved to Moscow at the age of 6 months. He attended Moscow State School 57. He graduated in 2005 and briefly attended State University Higher School of Economics before moving to USA. He earned associate's degree from Springfield Technical Community College in 2008 and transferred to University of Massachusetts Amherst, where he earned B.S. Chemistry and B.A. Economics degrees in 2011. During his college education, Artem worked as an undergraduate research student under Professor Michael D. Barnes, and that was when he got involved in work requiring image and signal processing in application to single molecule imaging and spectroscopy to uncover the connection between structure and optoelectronic properties. Through his advisor, Artem learned about the growing interest in such type of research at Oak Ridge National Laboratory, as well as about existence of the newly created Energy Science and Engineering Ph.D. program. Artem joined the University of Tennessee Knoxville and Oak Ridge National Laboratory Bredesen Center for Interdisciplinary Research and Graduate Education in 2013.