Empowering Materials Science with Machine Learning: From Electron Density to Atomistic Simulations

A thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy

Taoyuze Lyu

School of Physics
Faculty of Science
The University of Sydney

2024

This research reported in this thesis was supported by the award of the Dean's International Postgraduate Research Scholarship to the PhD Candidate

Declaration of Originality

This is to certify that to the best of my knowledge, the content of this thesis is my own work. This thesis has not been submitted for any degree or other purposes. I certify that the intellectual content of this thesis is the product of my own work and that all the assistance received in preparing this thesis and sources have been acknowledged.

Taoyuze Lyu

10 June 2024

Abstract

The development of modern society requires new materials to handle resource and environmental crises. Traditionally, the search for new materials is based on large numbers of trials and errors. Meanwhile, plenty of data, especially from failures, are wasted rather than forms experiences that can be shared among researchers. Machine learning techniques can help deal with complex information and take part in the design and exploration loop of materials, accelerating the simulation and property prediction process. In this thesis, we mainly aim to implement machine learning techniques for microscopic understanding of materials.

In Chapter 1, the importance of materials and the obstacles to materials design is briefly introduced. Then the roles and recent studies of artificial intelligence for science are reviewed and discussed to address the ability of machine learning to concur the hard problems. Finally, the chapter outlines the main directions explored in this thesis: one is the use of machine learning to investigate iodide-related defect kinetics within the all-inorganic halide perovskite $CsPbI_3$, and the other is the application of neural networks to predict the electron density of material systems. The methodology used throughout the thesis has been explained in Chapter 2, including the fundamental frameworks of density functional theory, molecular dynamics simulation, the deep learning interatomic potential, and the concurrent learning strategy for dataset construction.

In Chapter 3, a study on rare events in iodide-related point defect hopping in $CsPbI_3$ was conducted. We found that the existence of iodide vacancies or interstitials can effectively reduce the creation barrier for anti-Frenkel disorders. These additional defects can assist the fast and long-range ion diffusion, as well as the annihilation of defects. We have also systematically studied the hopping barrier of iodide interstitials and the influence of temperature on their hopping process. The discovered multi-defect phenomenon can help explain the self-healing of halide defects in halide perovskite materials.

In Chapter 4, we have found and systematically investigated the planar confinement behaviour of iodide interstitials in $CsPbI_3$. According to the molecular dynamics and kinetic Monte Carlo simulations, this behaviour will influence the defect lifetime and diffusion of ions. Specifically, the interstitials will move faster in the ⟨110⟩ direction family, but ⟨100⟩

are the slowest directions. Moreover, we have also studied the influence of symmetry breaking in low-temperature phases. The interstitials obtain higher on-site energy in the equatorial planes of halide perovskite lattice, while lower on-site energy in the apical directions. This can further restrict the motion of interstitials, preventing them hop freely inside the equatorial directions. The studies of defect kinetics in halide perovskite not only reveal the potential diffusion patterns of halide ions but also provide strategies for tuning halide perovskite-based optoelectronic or resistive switching devices.

Besides the application of machine learning in material simulations, in Chapter 5, the Deep Charge model was developed and described. The Deep Charge model can predict electron density in space from the input structures with ab initio accuracy. The model has been tested in various materials systems including metals, semiconductors, and insulators. In addition, it is easily applicable to different structural systems, such as crystals and surfaces, and even more complex alloys and amorphous structures. The machine learning techniques provide a way to represent electron density with a linear time complexity, rather than the cubic complexity of traditional density functional theory calculations. This can facilitate fast first-principles calculation on larger systems and the physical property predictions.

Acknowledgments

Since I was a child, I have been curious about the world. I wondered why objects dropped in a moving car seemed to move with it; I wondered whether atoms could be split; I wondered whether electrons could be used to build houses. And when I learned that the sun had only five billion years left, I cried for the inevitable demise of such a magnificent entity. Such greatness, yet unable to last forever. When I first encountered physics, I was deeply attracted to it and decided to understand it.

I would firstly thank my lead Ph.D. supervisor Prof. Rongkun Zheng, who provided me with the opportunity to be involved in physics. He has given me the freedom and resources to explore the things that I am interested in. I appreciate his knowledge in many areas and his assistance in applying computational resources and writing reports and research articles. It was such a good fortune to be supervised by him. I would also like to thank my co-supervisor Dr. Feng Li, who is more like a peer, a friend of us. He has given me so much help during my research. Besides, I would like to express my gratitude to all the academic staff and teachers who helped me throughout my study. Prof. Xudong Yang has provided me with the research resources in Shanghai Jiao Tong University (SJTU) in the offshore period during Covid 19. I have learned the basic fabrication and characterization knowledge about specific materials from there. Prof. Zhicheng Zhong has had discussions with me and has provided me with plenty of valuable insights about computational physics. Prof. Cathy Stampfl, Prof. Kevin Varvell, and Prof. Holger R. Dullin have taught me the courses in the first two years of my Ph.D. candidature. Prof. Geraint Lewis and Dr. Benjamin Fulcher were on my progress evaluation meeting panel, and they were very responsible and sincerely helped me integrate into the physics family. Going back in time, I thank my undergraduate supervisor Prof. Bin Liu, who introduced me to the academic world and allowed me to experience computing research. I appreciate my middle school physics teacher Mr. Kun Wang for the courses he taught and for actively discussing my imaginative questions with me. Meeting them is truly a great fortune for me.

I would like to thank my fellow students. Dr. Ruhao Fang, Dr, Tiebin Yang, and Ms. Xingmo Zhang have provided me with the information of our university before and directly after my enrolment, which helped me a lot. Mr. Runkai Liu, Ms. Fang Zeng, and Mr. Josh

Maggiora have often discussed academic problems with me, which are invaluable for understanding physical research and I enjoy it very much. Discussion is one of the most important processes of research. Mr. Yuhang Liang has helped me revise my manuscripts. Mr. Xuejian Qin and Mr. Yifan Shan have helped me with coding in my research. Dr. Zijie (Jeff) Luo, who was in the same research group as me during my undergraduate studies, helped me a lot during my Ph.D. application period and when I first arrived in Sydney. Dr. Zhensheng Dai personally taught me how to conduct experiments and helped arrange tests at SJTU. I would like to thank all the partners for accompanying me through my Ph.D. time.

Last, and most importantly, I would like to thank my parents and grandparents, who have carefully accompanied me in my growth. Throughout my life, they have encouraged me so many times, respected my decisions, and given me the courage to explore the unknown. These are invaluable. Without them, I would not have become a determined and positive person and would be impossible to learn and understand physics. I would like to thank all the friends who have encouraged and helped me. When it comes to physics, I am still just an ignorant youth. I hope that one day I can pick up a beautiful shell from the sea of the unknown.

Authorship Attribution Statement

This thesis contains works that have been published or are under revision for publication. The materials contained in the works are included in Chapters 3, 4, and 5. I have played the primary role in these works. Details of my contribution and publication information are listed below:

- **Chapter 3** of this thesis is published as [1] as the front cover of the journal:

  Taoyuze Lv, Yuhang Liang, Fang Zeng, Feng Li, Xudong Yang, Jun Huang, and Rongkun Zheng*. *Kinetic Process with Anti-Frenkel Disorder in a $CsPbI_3$ Perovskite*. The Journal of Physical Chemistry Letters 2024 15 (10), 2929-2935.

  I proposed the study, constructed the dataset, analyzed the data, wrote the manuscript, and designed the front cover art for publication.

- **Chapter 4** of this thesis is under revision:

  Taoyuze Lv, Yuhang Liang, Fang Zeng, Feng Li, Xudong Yang, Jun Huang, and Rongkun Zheng*. *Anisotropic Motion of Iodide in $CsPbI_3$: The Role and Influence of Planar Confinement*. Chemical Science (under revision)

  I designed the simulation procedure, performed the simulations, analyzed, and drafted the manuscript.

- **Chapter 5** of this thesis is published as [2]:

  Taoyuze Lv, Zhicheng Zhong*, Yuhang Liang, Feng Li, Jun Huang, and Rongkun Zheng*. *Deep Charge: Deep learning model of electron density from a one-shot density functional theory calculation*. Physical Review B, 2023, 108(23): 235159.

  I designed the workflow, wrote the codes, collected the data, trained and tested the model, and wrote the manuscript.

Besides the above-listed works, I have also contributed to, or conducted many other studies that are not directly related to this thesis, as listed below:

- Taoyuze Lv, Yuhang Liang, Feng Li, Xudong Yang, Jun Huang, and Rongkun Zheng*. *Two-dimensional halide perovskites: A review on their orientations*. Science China Physics, Mechanics & Astronomy, 2023, 66(1): 217306.
- Zhensheng Dai, Taoyuze Lv, Julien Barbaud, Wentao Tang, Tao Wang, Liang Qiao, Han Chen, Rongkun Zheng, Xudong Yang*, and Liyuan Han. *Stable tin perovskite solar cells developed via additive engineering*. Science China Materials, 2021, 64(11): 2645-2654.
- Zhensheng Dai, Wentao Tang, Tao Wang, Taoyuze Lv, Xinhui Luo, Danyu Cui, Ruitian Sun, Liang Qiao, Han Chen, Rongkun Zheng, Xudong Yang*, and Liyuan Han. *Stable tin perovskite solar cells enabled by widening the time window for crystallization*. Science China Materials, 2021, 64(8): 1849-1857.

In addition to the statements above, in cases where I am not the corresponding author of a published item, permission to include the published material has been granted by the corresponding author.

Taoyuze Lyu

20 June 2024

As supervisor for the candidature upon which this thesis is based, I can confirm that the authorship attribution statements above are correct.

Prof. Rongkun Zheng

20 June 2024

Table of Contents

List of Figures

List of Tables

Table of Abbreviations

ϱ	ground-state electron density
T	temperature
P	pressure
V	volume
E	energy
$\mathbf{F}_i$	force on atom i
$\mathcal{D}$	descriptor
$\mathcal{F}$	fitting net
k_B	Boltzmann constant
E_a^D	self-diffusion activation energy
E_a^v	hopping activation energy
2D	two-dimensional
3D	three-dimensional
AI	artificial intelligence
AIMD	ab initio molecular dynamics
aF	anti-Frenkel disorder
CV	collective variable
DFT	density functional theory
DP	deep potential
DPMD	deep potential molecular dynamics
DP-GEN	deep potential generator
FA	formamidinium

GGA	generalized gradient approximation
HK	Hohenberg-Kohn
I_i	iodide interstitial
IP	in-plane
KS	Kohn-Sham
kMC	kinetic Monte Carlo
PCE	power conversion efficiency
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator
LDA	local density approximation
MA	methylammonium
MAE	mean absolute error
MSE	mean square error
MD	molecular dynamics
NVT	canonical ensemble
NPT	isothermal–isobaric ensemble
NVE	microcanonical ensemble
NEB	nudged elastic band
OP	out-of-plane
PBE	Perdew-Burke-Ernzerhof
RMSE	root mean square error
VdW	Van der Waals
V_I	iodide vacancy
VASP	Vienna Ab initio Simulation Package

Chapter 1.

Introduction

1.1. Resource crisis and why materials matter

Nowadays, with the development of human civilization, we are facing all kinds of crises due to the approaching of critical points in the complex social and environmental systems [3,4]. The consumption of energy and other resources in modern society is growing at an unprecedented rate. Take data and computation technology for instance, the training and inference processes of artificial intelligence (AI) are highly energy and water resource consumable, such as the generative large language models (LLM) ChatGPT from OpenAI [5] and Gemini from Google [6]. For example, the pictures in **Figure 1.1** below are generated by ChatGPT-4 using about nine dialogues, with a rough estimate of approximately 0.045 kWh of electricity consumed. In 2021, the total electricity consumption of Google was 18.3 TWh, with 10%~15% cost by AI [7]. If every standard search includes the LLM interaction, the energy consumption of Google will become as large as a country such as Ireland, about 29.3 TWh per annum [8]. AI servers are estimated to consume about 85.4~134.0 TWh of electricity in the year 2027 [8]. In addition, this consumption will only become more sizable as the complexity and number of parameters in AI models continue to increase, along with the growing demand for AI-generated content (AIGC) [9]. Besides supercomputing infrastructure, the energy needs of other manufacturing industries are also increasing to fulfil the urbanization. Overcoming bottlenecks in human development is a matter of urgency.

To address these obstacles, we need to allocate and manage resources more intelligently, while on the other hand, it is more important to explore new resources. Throughout the history of mankind, every great step forward has always been inseparable from the discovery and application of new materials, as shown in **Figure 1.1**. After recognizing copper and iron and smelting them, human productivity increased dramatically, leading to

a steady rise in population. The subsequent industrial revolutions advanced the widespread use of steel and polymers, laying the foundation for modern society. Today, the acquisition of monocrystalline silicon has brought us into the information age. In short, materials matter.

Figure 1.1 Schematics for the relation between human civilization and the discovery of new materials. The figures are generated by ChatGPT to show the rapid development of AI.

For the energy and environmental crises we are facing, advanced materials meeting various requirements are urgently awaited. In terms of energy consumption and efficiency for computing devices, semiconductor materials that are superior to monocrystalline silicon have the potential to exist [10–12]. Room-temperature superconductors could theoretically avoid heat loss in electronics to a great extent, even though their existence has not yet been experimentally established [13,14].

In addition, new materials can also provide breakthroughs to energy harvest and storage. Photovoltaic devices based on semiconductor materials endow us with the ability to harvest solar energy. For example, silicon is the traditional photoactive layer for solar cells and has been widely used in solar farms, aerospace, and everyday life. However, the production of monocrystalline silicon solar cells is highly energy intensive due to the high temperatures required for refining and melting the silicon from SiO_2. Meanwhile, halide perovskite materials such as methylammonium lead triiodide ($MAPbI_3$), formamidinium lead triiodide ($FAPbI_3$), and $CsPbI_3$, have been extensively studied since 2009 when their photovoltaic properties were first discovered by Miyasaka et al. [15,16]. The power conversion efficiency (PCE) of the single junction halide perovskites thin film solar cell has recently

surpassed 26% in the laboratory, which is comparable to traditional monocrystalline silicon solar cells [17–19]. Besides the efficiency, the fabrication of halide perovskite-based solar cells requires less energy with solution-based methods such as spin-coating, doctor blading, screen printing, etc. [20], which will reduce production costs to a considerable extent. Moreover, halide perovskite materials have their intrinsic advantages over silicon, they are direct band gap semiconductors, while silicon obtain an indirect band gap. This can lead to more efficient light harvesting and requires a thinner photo-active layer than the silicon case, which saves material and compresses the device sizes.

1.2. Current obstacles to materials research

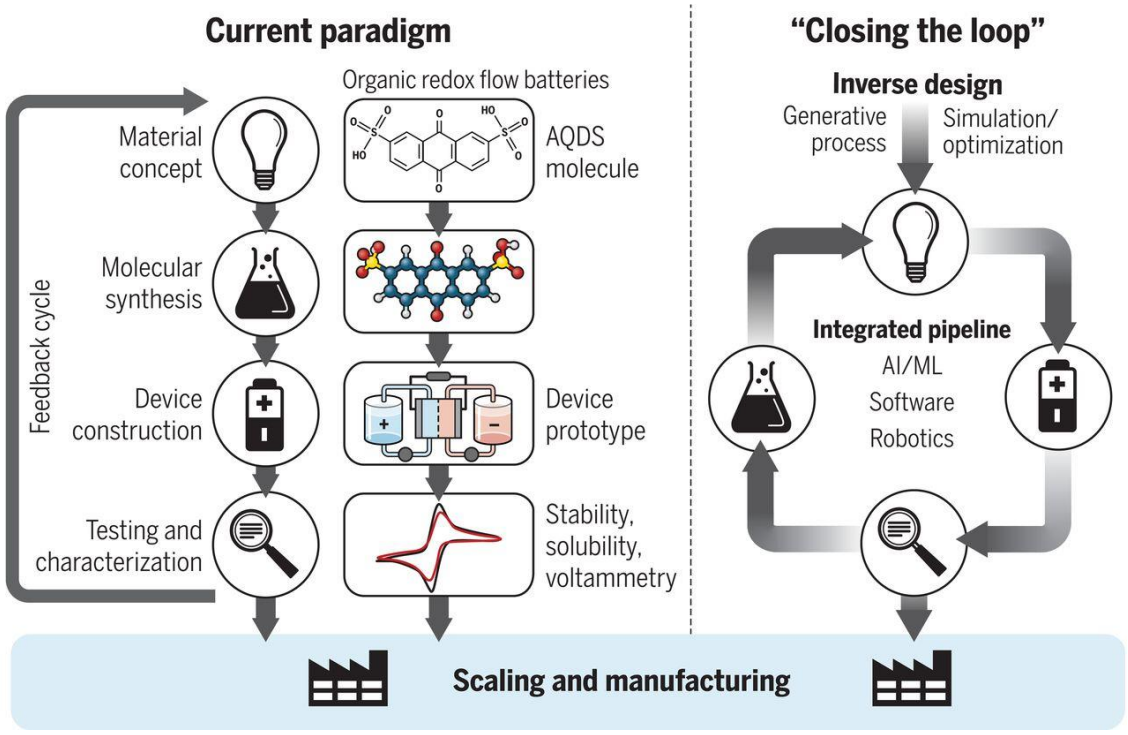

Figure 1.2 The comparison of the current science discovery loop and the inverse design loop. The left side is the current procedure for finding new materials, while the right side is the incorporation of the inverse design procedure. From ref. [21]. Reprinted with permission from AAAS.

However, although we understand that materials research is of vital importance to the future development of human society, there are two major obstacles that we are facing. The first

is the inverse design of materials [22], i.e. how could we find the ideal materials from a given target property? In theory, it is possible to design a room-temperature superconductor or a more "perfect" semiconductor material than silicon, but we can hardly know the specific composition and structure of that target material until we find one. Therefore, we are even less likely to know how to fabricate it. The current materials research is based on large amounts of trials and errors, as shown in the left side of **Figure 1.2**, though theoretical prediction and experiences can provide a little guidance. There is still a long way to achieve the ability of inverse design and implement it into the new materials discovery loop, as shown on the right side of **Figure 1.2**.

The second obstacle is the understanding of microscopic material properties. The above-mentioned advanced materials possess fascinating properties over traditional materials, but many of them have their problems. For example, halide perovskites are known for their instability, which can easily degrade to other non-photoactive matters [23]. To solve these problems and gain physical knowledge for designing new materials from scratch, a thorough understanding at the atomic level is necessary. Nevertheless, the characterization of materials at the atomic level remains challenging, primarily due to the intricate nature of atomic interactions and the limitations of current analytical techniques in achieving the resolution and sensitivity required to accurately observe and measure these complex phenomena.

Contemporarily, the most commonly used method for studying microscopic properties in materials and condensed matter physics is computation and simulation. Utilizing the quantum mechanics or empirical functions, the static and dynamical properties of materials could be derived with the assistance of supercomputing clusters. Two well-known computing methods are density functional theory (DFT) [24] and molecular dynamics (MD) [25]. DFT aims to solve the Kohn-Sham equation for electrons in a material, thereby providing a quantum mechanical description of the electronic structure [26]. This method is particularly powerful for predicting material properties such as electronic band structure [27], magnetic properties [28], and optical characteristics [29]. On the other hand, MD simulations offer insights into the behavior of atoms and molecules over time, allowing researchers to study the dynamic processes and thermal properties of materials at an atomic scale. MD can be divided into two main categories: ab initio MD (AIMD) [30] and classical MD [31]. AIMD borrows quantum mechanical methods for energy and force calculations

and uses these values to update the structure of systems. However, owing to the involvement of the degree of freedom of electrons, AIMD is highly computationally intensive. Meanwhile, classical MD relies on solving Newton's equations of motion for a collection of particles, using force fields derived from empirical data to model the interactions between atoms. However, although these computational methods could to some degree provide a more intuitive understanding of material properties on electronic and atomic scales, they still fall short of fully representing the materials at a comprehensive level.

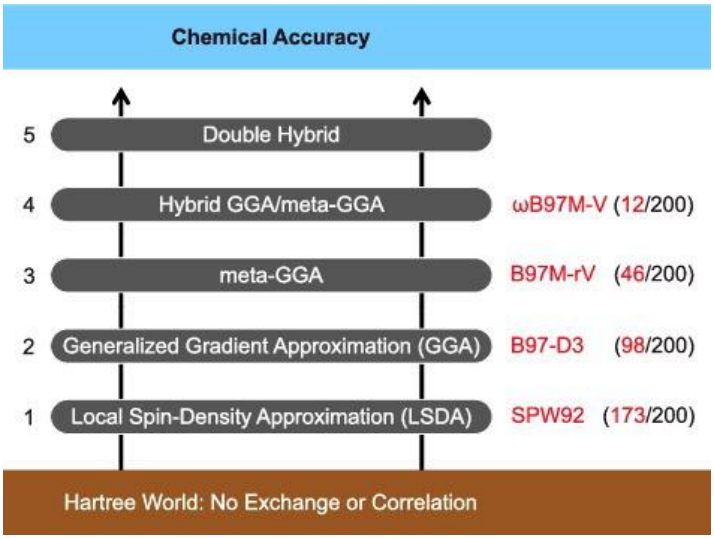

Figure 1.3 The Jacob's Ladder towards chemical accuracy. In each step, new physical content is added and the accuracy of the DFT model will be improved. Reproduced with permission from ref. [32], copyright 2017, Informa UK Limited, trading as Taylor & Taylor & Francis Group.

DFT is based on the non-interacting single-electron approximation of the full-electron Schrödinger equation, which takes the total electron density as the functional variable instead of considering all electron degrees of freedom. This formulation reduces the complexity of computing but introduces another problem: the exchange-correlation functional, which contains all the "hard-solving" parts of the energy functional. The general form of the exchange-correlation functional is unknown, and various approximations have been taken into account such as the local density approximation (LDA) [26] and

generalized gradient approximation (GGA) [33,34]. The errors in DFT can be generally attributed to the limitation of the exchange-correlation functionals. Currently, we are gradually approaching chemical accuracy by climbing the "Jacob's Ladder", as illustrated in **Figure 1.3.** However, as the precision continues to improve, the complexity of implementation rapidly increases, and computational costs also progressively rise. This leads to a reduction in the size of the systems that can be studied, limiting research on many complex systems and numerous properties. On the other hand, classical MD exhibits a notable rapidity, enabling its application to systems comprising millions of atoms. However, the accuracy of classical MD is contingent upon the force field employed, which is typically calibrated manually based on data derived from DFT or experimental results. The lack of details within the force field parameters could result in the generation of unrealistic microscopic kinetics, even though the macroscopic statistical properties might align with experimental values in some instances.

To solve the two major problems, more efficient ways are required to predict material properties rapidly with high precision and to simulate large systems over longer timescales. This indicates that we need to combine the accuracy of DFT with the speed of classical MD. The question then arises: how can such an integration be achieved?

1.3. The power of artificial intelligence for science

1.3.1. The prediction ability of AI

As mentioned above, the development of AI will lead to a significant increase in energy consumption in modern society. At the same time, however, AI is the first tool with intelligence invented by mankind, and its gradual rise in functionality and capability has the potential to provide new avenues for solving both of these challenges. A brand-new research field is rising: AI for science (AI4Science) [35].

AI4Science utilizes machine learning to accelerate and enhance scientific research across a broad spectrum of disciplines. This innovative approach harnesses the power of AI to analyze vast datasets, uncover latent patterns, and generate hypotheses at speeds and scales previously unattainable by human researchers alone. By integrating AI technologies such as machine learning, deep learning, and natural language processing, scientists can tackle

complex problems more efficiently, from understanding climate change dynamics to discovering new materials and drugs [36,37]. In 2021, researchers from DeepMind proposed the deep learning model AlphaFold for three-dimensional protein structure prediction [38]. This model outperformed other prediction methods significantly in the 14th Critical Assessment of Protein Structure Prediction (CASP14), with a median backbone accuracy of 0.96 Å r.m.s.d.$_{95}$ (Cα root-mean-square deviation at 95% residue coverage) compared to the next best performing method of 2.8 Å r.m.s.d.$_{95}$. The researchers have also launched the AlphaFold Protein Structure Database, containing the prediction results for unknown protein structures. In 2022, they uploaded around 200 million results, almost covering all the known proteins on earth [39], providing unprecedented information for the structural biology community.

Moreover, AI could also assist in solving fundamental physical problems from the observations. In 2019, Yosinski et al. proposed an AI model to learn the Hamiltonian $\mathcal{H}$ of various physical systems [40], and a year later Ho et al. constructed the Lagrangian Neural Networks that can parameterize arbitrary Lagrangians $\mathcal{L}$ from data in an arbitrary coordinate system [41]. $\mathcal{H}$ and $\mathcal{L}$ are all physical quantities of vital significance, especially for modern physics, which can capture the kinetics of a system with conservation laws. In 2020, Tegmark and Udrescu from the Massachusetts Institute of Technology proposed a recursive multidimensional symbolic regression model named "AI-Feynman" to discover the correct symbolic expression from the observed data [42]. This model correctly discovered 100 equations from the *Feynman Lectures on Physics*. Later in 2022, Liu and Tegmark presented another symbolic regression AI model for finding the hidden symmetries of a system [43], which is of fundamental interest to theoretical physics. In the same year, Battaglia et al. proposed a graph network with symbolic regression implementation that could successfully rediscover Newton's gravitation law from the orbital data of our solar system [44].

1.3.2. The structure of AI

In brief, the well-trained AI can be regarded as mappings or functions from the input information to the target values. These functions generally have a large number of

parameters as the degrees of freedom to be tuned in the training process. An AI model M can be represented as $\boldsymbol{y} = M(\boldsymbol{x}; \boldsymbol{\theta})$, in which **y** is the target value, generally a vector with multiple elements; **x** is the input information, such as the data required for finding patterns; and **θ** is the tunable parameters of the model. Different from traditional function-solving problems such as giving **x** to find **y**, AI aims to find the functions themselves with the given **x** and **y**. There are numerous machine learning frameworks for finding the solution, but herein, only the broadly adopted artificial neural network approach is taken for demonstration [45].

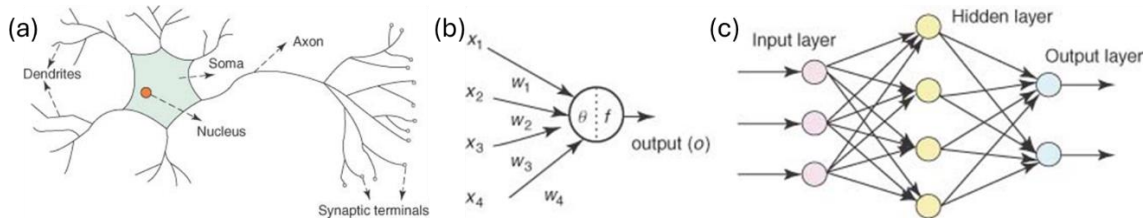

Figure 1.4 (a) the general structure of a mammalian neuron. (b) The architecture of an artificial neuron. (c) The structure of a multilayered neural network. Reproduced with permission from ref. [46], John Wiley & Sons, Inc.

Artificial neural networks solve practical problems by imitating biological nervous systems. As shown in **Figure 1.4(a)**, a neuron cell of mammals contains a cell body called soma, root-like dendrites, and a long wire-like axon. The mechanisms of information transmission and processing involve complex biochemical reactions. In brief, the dendrites are connected to other neurons and receive their transmitter substances, depending on which the cell lowers or raises its electrical potential. Once the potential is larger than a threshold, a pulse will be sent to the axon and then transmitted to other connected neurons.

A typical artificial neuron is shown in **Figure 1.4(b)**. The neuron takes several inputs (x_1, x_2, …), then manipulates the received information, and finally outputs the signal O. The output signal can be directly regarded as the final result or as the input signal to the subsequent neurons to form a multilayered neural network. To process the information, the neuron first assigns every input value (x_i) a weight (w_i), which determines its influence on

the output. These values are then summed together with a bias (b) term which determines the threshold. The obtained signal is finally input to the activation function (f) to determine the value of the output. The whole process can be written as:

$$O = f\left(\sum_i w_i x_i + b\right) = f(\boldsymbol{w} \cdot \boldsymbol{x} + b) \tag{1}$$

This is the information processing of a single neuron. To provide predicting abilities towards complex problems and mimic the biological neurons, the activation function f should be non-linear, such as the Rectified Linear Unit (ReLU), Sigmoid, and Tanh functions:

$$\mathrm{ReLU}(x) = \max(0, x) \tag{2}$$

$$\mathrm{Sigmoid}(x) = \frac{1}{1 + e^{-x}} \tag{3}$$

$$\mathrm{Tanh}(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{4}$$

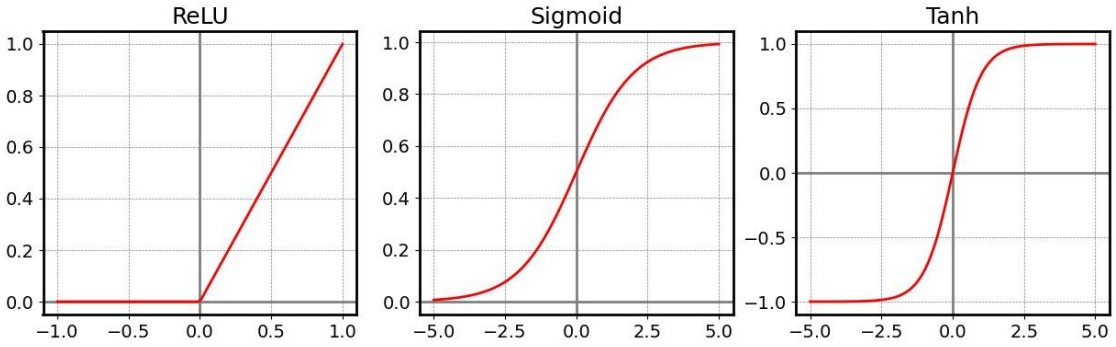

Figure 1.5 The three commonly used activation functions.

To form a neural network, we need a hidden layer containing multiple neurons, as shown in **Figure 1.4(c)**. The outputs of this layer can be directly input to the output layer, which gives out the results based on the information provided by the hidden layer, or to another

hidden layer. The layers could be connected forming a sequence, endowing the neural network with more non-linear capabilities. A neural network that contains a large number of hidden layers is called a deep neural network (DNN), which is a commonly used architecture for deep learning.

1.3.3. The training process

One of the powerful abilities of AI is learning, which is different from other tools created by mankind. As we believe, knowledge is contained in data. We need to first investigate and collect observations, then distil the fundamental rules behind them to understand the world. This is similar to the training process of AI. The neural networks can automatically adjust the above-mentioned parameters via algorithms to align the prediction to the provided data.

Generally, the training process requires two objects: the AI model and the datasets. Before the training, the dataset needs to go through a preprocessing pipeline [47,48], since the raw data are usually not directly readable for the model, or the data is unevenly distributed which could lead to bias while training [49]. The data should be split into two parts for training and validation, respectively. For a single training step, the preprocessed data are input into the model, and it will give a prediction based on the input. The predicted values are compared to the true results with a loss function, and finally, the algorithm adjusts the parameters of the model to minimize the loss of the model, which is called backpropagation [50,51]. The most well-known method for determining how to update the parameters of neural networks is gradient descent [52], which utilizes the chain rule of multivariant derivatives to backpropagate the loss onto every parameter. To obtain a model with enough predictability, thousands to millions of training steps are usually required, depending on the complexity of the data and the scale of the model.

After obtaining the optimized parameters, the model can be used for predicting unknown data without backpropagation.

1.4. Recent development of AI in materials science

As discussed above, there are oceans of complex problems in science that can be highly sophisticated for research. The fusion of AI and science could be the solution, which is experiencing a remarkable surge, with plenty of researchers promoting its development. Herein, I briefly described the machine learning studies focused on the field of materials science and condensed matter physics. In general, AI could assist scientists in three ways: property predictions, calculations and simulations, and experiment analysis.

1.4.1. AI in material properties prediction

Machine learning methods have been widely used in the property prediction of materials. In this task, the AI models are required to find out the mapping between the features from the composition of atomic species or the crystal structures to the target properties.

A number of AI models have mostly used the structures of materials together with the ab initio calculated properties for training and predicting, these data are relatively "clean", meaning no experimental errors. In 2017, Isayev et al. proposed the universal fragment descriptors model for inorganic materials [53]. This model can utilize the crystal structure to predict electronic properties such as band gap energy, mechanical properties such as bulk modulus and shear modulus, and thermodynamical properties such as heat capacity and thermal expansion coefficient. In the same year, Müller and Tkatchenko et al. proposed the SchNet based on deep tensor neural networks (DTNN) for the property prediction of molecules and solid materials [54,55]. SchNet obtains a high accuracy. For example, the mean absolute error (MAE) for the formation energy prediction is lower than 0.035 eV/atom for the Materials Project database. In 2018, Agrawal et al. proposed the ElemNet architecture for predicting material properties from only elemental composition information [56]. The model can effectively understand the interactions and similarities among different chemical species and can efficiently predict properties with high accuracy. In 2022, Ganguly et al. constructed an auto-encoder-based framework CrysXPP for property prediction in crystalline materials [57], which obtains a generally higher accuracy towards the state and elastic properties of materials than other compared models.

Several models have adopted the real experiment data for property prediction. In 2021, Gregoire et al. introduced the Hierarchical Correlation Learning for Multi-property Prediction (H-CLMP) framework for training and predicting multiple properties as a whole [58]. The model can predict spectral optical absorption in the complex metal oxide system with only the composition information. Compared to single-target models, the multi-target framework is more effective and efficient in understanding underlying chemical relationships. Since the compositional space is high dimensional, the correlation among different compositions and different properties is highly complex, providing more property information can assist the exploration of fundamental intrinsic correlations among properties and in efficiently finding out the relationship for accurate property prediction.

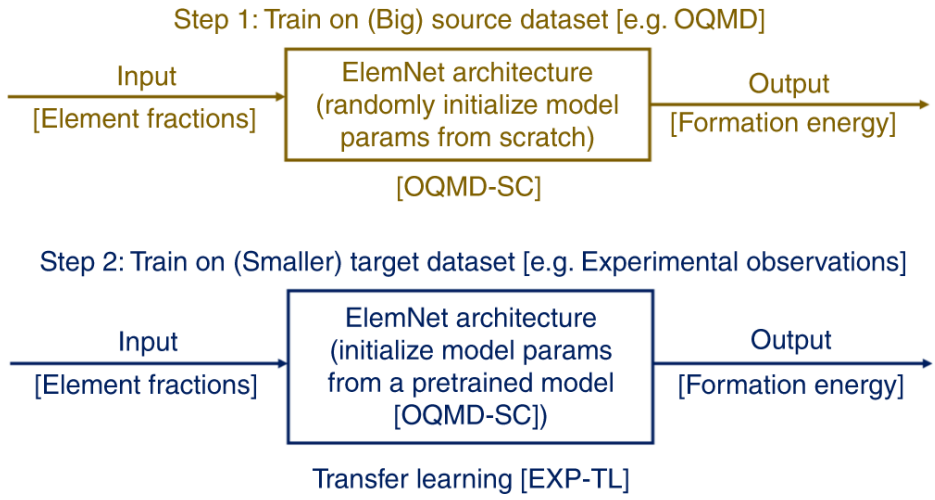

Figure 1.6 The procedure for transfer learning on finetuning the DNN model trained on DFT results towards experimental data. Reproduced with permission from ref. [59], 2019 Springer Nature Limited.

Some models combine the ab initio data with the experimental observations. In 2019, Agrawal et al. adopted the above-mentioned ElemNet architecture for predicting experimental properties based on a transfer learning strategy [59], as shown in **Figure 1.6**. They first trained the model with the Open Quantum Materials Database (OQMD) [60], a large database containing DFT results, and then fine-tuned the model based on the SGTE Solid SUBstance (SSUB) database containing experimental data. With this framework, the

model can perform better with even small experimental datasets. In 2022, they used the IRNet architecture [61] to predict the formation energy of materials [62]. They have trained the model based on the OQMD, and the experimental data are constructed with Matminer [63]. The addition of experimental data increases the accuracy of the model's prediction and outperforms that of pure DFT data, by about 20% better. This work shows that a small amount of ground truth data, i.e. the experimental observations, could increase the accuracy of the model.

1.4.2. AI in first-principles calculation

As previously mentioned, there are several obstacles in the field of computational materials science. AI can assist in dealing with problems via increasing the ab initio calculation speed or constructing machine learning potential energy surface.

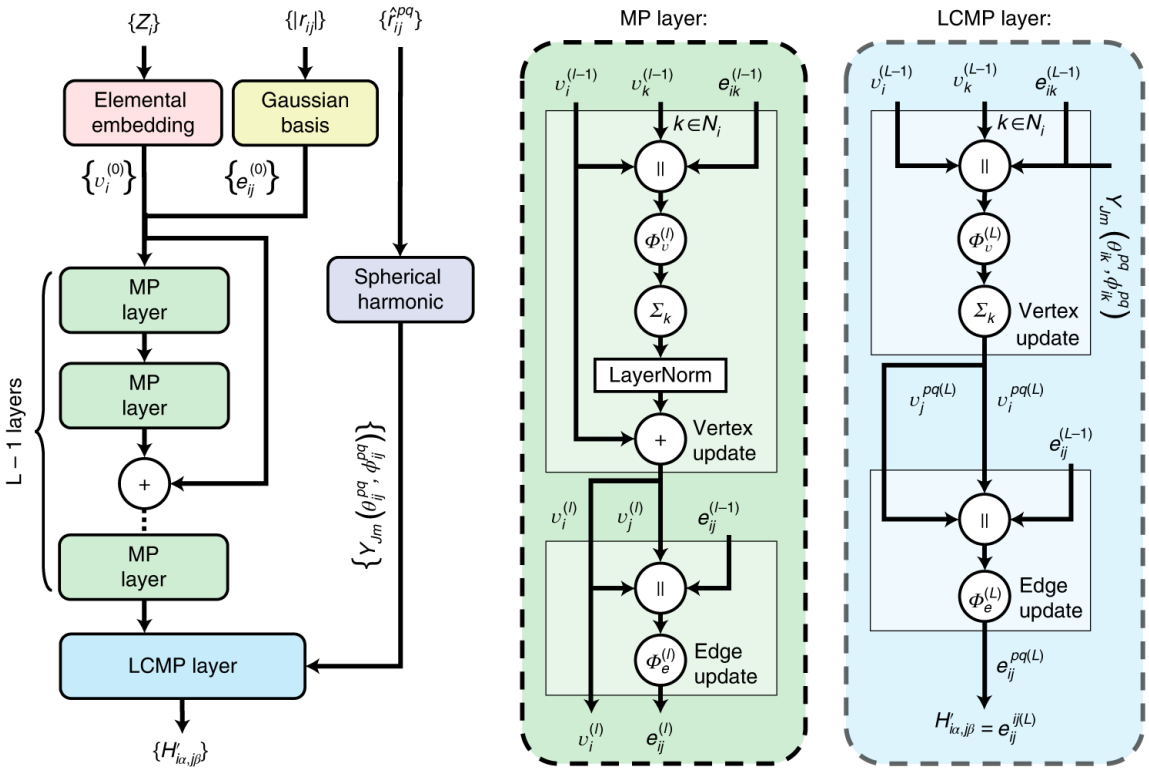

Figure 1.7 The model structure of DeepH. Reproduced with permission from ref. [64].

For the density functional theory calculations, one of the most time-consuming procedures is the self-consistent field (SCF) calculation. Nowadays, various machine learning methods assist in bypassing the SCF calculation. One way is to predict the Hamiltonian of a system, which encodes the electronic structure information. In 2017, Hegde and Bowen used bispectrum formalism to represent the local atomic environment and trained the machine learning model for Hamiltonian prediction with kernel ridge regression [65]. For interatomic energy, the RMSE of the model is 6 meV for copper and 18 meV for diamond. In 2022, Xu et al. developed the DeepH model by means of a message-passing neural network [64]. The model can predict the local Hamiltonian based on the environment and preserve the gauge covariance of the DFT Hamiltonian matrix. The atomic numbers and local coordinates are taken as the input features. As shown in **Figure 1.7**, firstly, the atomic numbers are embedded into initial vertex features and the local coordinates are expressed by Gaussian basis as the edge features. Then the features are successively updated via message-passing (MP) layers. Lastly, the features are used to represent the Hamiltonian and decoded via the local coordinate message-passing (LCMP) layer. The DeepH model was tested in the twisted vdW materials and shows good alignment to the DFT band structure and electric susceptibility.

On the other hand, the orbital-free DFT (OF-DFT) [66] framework assisted by machine learning can also increase the efficiency of ab initio calculations. The OF-DFT scheme said that the energy functional of a system is dependent purely on the electron density, which is the same as the original DFT idea, with a time complexity of O(N). Meanwhile, this approach is inaccurate since the kinetic energy functional (KEF) is hard to construct and can lead to large errors. Thus, the broadly used Kohn-Sham DFT was proposed, which requires the construction of the Kohn-Sham orbitals for energy functional calculations, with the time complexity of $O(N^3)$. If we could find a suitable KEF, the OF-DFT could be employed in material calculations. In 2022, Ryczko and Tamblyn et al. proposed the voxel DNNs method for computing KEF and its functional derivative [67]. The model predicts the KEF with sliced electron density data and shows agreement with the KS-DFT results in graphene lattice. In 2023, Suryanarayana et al. adopted the Δ-machine learning model for the OF-DFT scheme [68]. The model learned a correction term to the OF energies that captures the difference between OF and HK results. The tests showed that this model is efficiency and can be about 50% more accurate on energy prediction than the machine learning force field trained only on KS-DFT.

1.4.3. AI in materials simulation

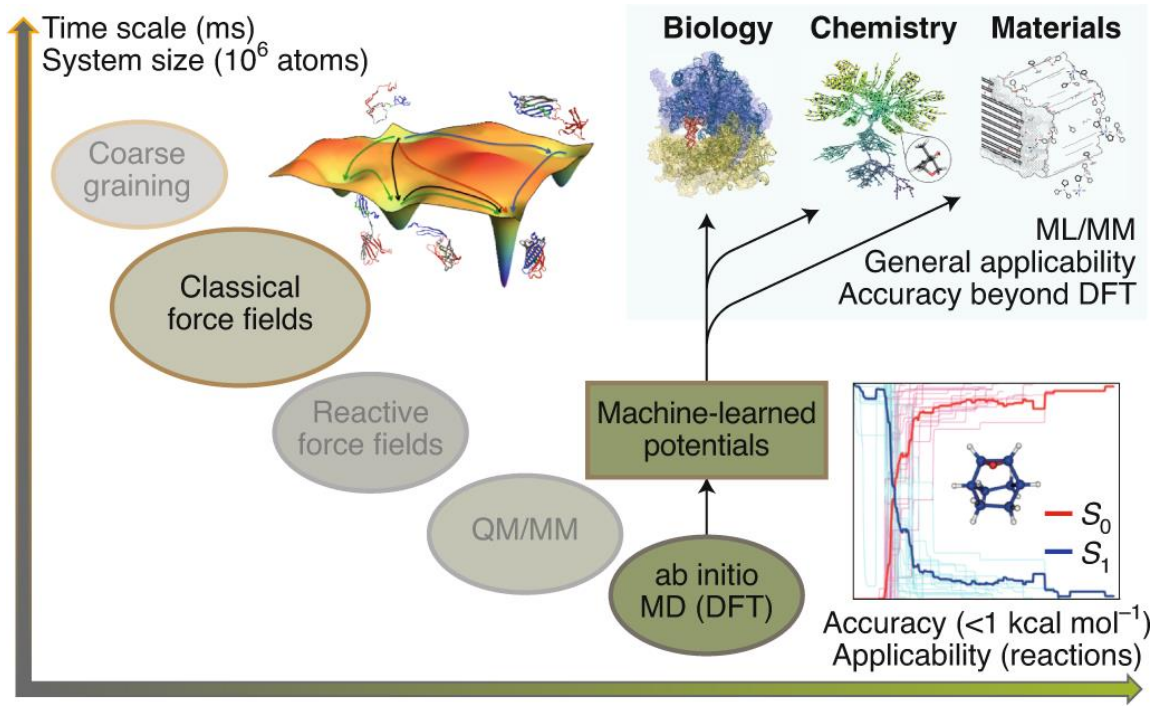

Figure 1.8 The advancement of machine learning potentials. The traditional methods are listed in the circles, in which the most accurate method normally obtains the lowest efficiency, while the fastest methods that can do large-scale simulations are of the lowest accuracy. The machine learning method could provide a way for large-scale simulation with ab initio accuracy. Reproduced with permission from ref. [69], Copyright © 2021, Springer Nature Limited.

At the same time, machine learning also empowers materials simulation, endowing them with accuracy comparable to ab initio calculations but easily performed in larger systems, as shown in **Figure 1.8**. The traditional interatomic potentials for MD simulations are manually fitted according to experimental or DFT results. However, this generally requires a large number of works for finetuning and loses details, making it hard to approach the ab initio accuracy. One possible solution is using numerical optimization to fit the potentials with a set of DFT results [70]. Early in 1995, Doren et al. adopted neural networks to represent the potential energy surface [71]. As discussed above, the neural network models obtain a large set of parameters such as biases and weights, which can be regarded as degrees of freedom for a complex function. Thus, they have the ability to approximate the potential energy surfaces.

Contemporarily, there are more advanced neural network models developed for material simulations. In 2010, Bartók et al. introduced the Gaussian approximation potential (GAP) method, which can capture the relation between atomic energy and the local environment from DFT results [72]. Later in 2021, based on the GAP method, Deringer et al. performed the GAP-MD simulation of systems with 100,000 silicon atoms for nanoseconds [73]. The simulation can provide atomistic insight into the phase transition in disordered silicon with DFT accuracy, which is impossible to obtain from ordinary DFT simulations. In 2017, Roitberg and Isayev et al. constructed the ANI-1 model based on the Behler and Parrinello symmetry functions from GAP, which obtains an RMSE of 1.3 kcal/mol for organic molecules [74].

In 2018, E et al. proposed the original deep potential molecular dynamics (DPMD) framework [75]. This model also predicts the energy and force on atoms based on knowledge learned from ab initio data. The neural network first extracts the environment information with an embedding network to construct the symmetry-conserved descriptors. Then the descriptors are further processed by the fitting net to obtain the energy and force on atoms. Owing to the outstanding efficiency and accuracy, DPMD achieved simulation at the scale of over 100 million atoms in 2021 [76]. Currently, DPMD has also implemented the attention mechanics for the embedding network [77], which provides the ability to train one single model over a wide range of atom species, providing the possibility to construct a large atomic model that contains every element on the periodic table.

Many of the machine learning potentials are based on the local atomic environment and ignore the long-range electrostatic interactions. Thus, several models have included the prediction of point charge, dipole and multipole moments' contribution [78–80]. The DPMD framework can also include long-range interactions via constructing the maximally localized Wannier centres [81].

1.4.4. AI in material experiments and analysis

Machine learning can also be applied to the process optimization of experiments, and even automatically conducting experiments. In 2016, Norquist et al. designed a support vector machine to construct procedures from failure experimental data for the reaction and

crystallization of templated vanadium selenites [82]. In 2019, Okubo and Chaikittisilp et al. used extreme gradient boosting (XGBoost) and random forest models for constructing the synthesis-structure relationship of zeolite materials [83]. As shown in **Figure 1.9**, first, the data were collected from published literature and datasets. Then the machine learning model will construct a descriptor, i.e. a high dimensional representation vector in latent space, for the synthesis procedure. The relationship between the synthesis procedure and the structure similarity is constructed, based on which the model can predict the synthesis condition for unknown zeolites. Finally, the prediction can be verified via experiments. Besides process optimization, there are also numerous models designed for characterization, such as microstructural analysis [84–86], spectroscopic analysis [87,88], atom probe tomography [89–91], etc.

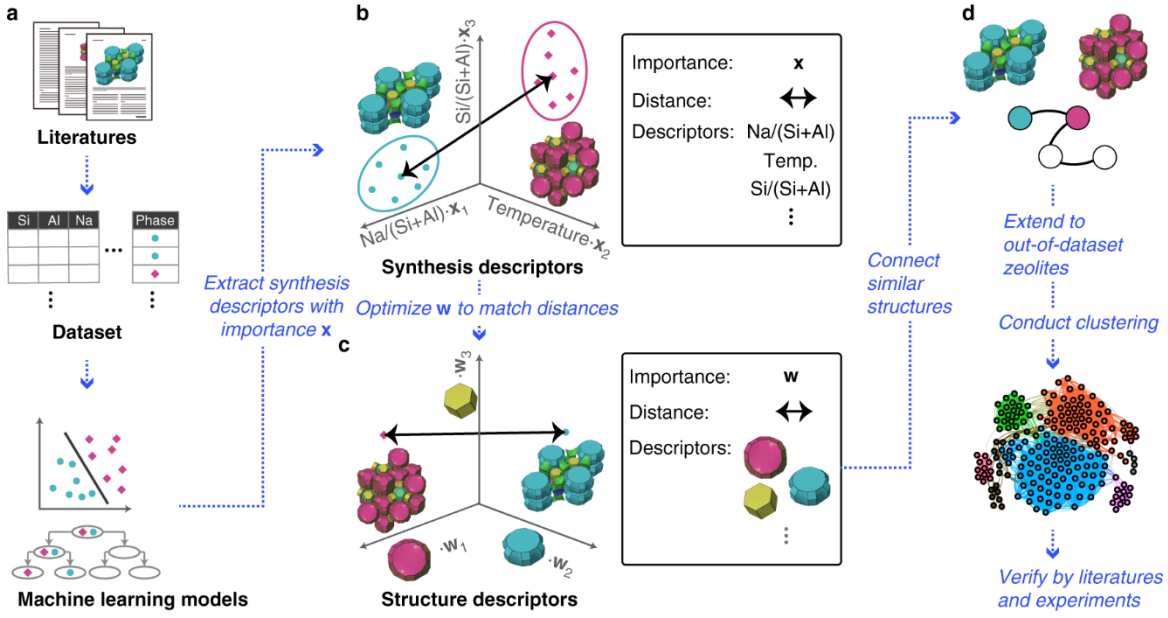

Figure 1.9 Workflow for finding the synthesis procedure for zeolites with different structures with machine learning. Reproduced with permission from ref. [83]. 2019 Springer Nature Limited.

More interestingly, AI and robotics have the ability to conduct experiments with less human interference. In 2020, a research group from the University of Liverpool built a robotic chemist, demonstrating the ability to find improved photocatalysts for hydrogen production

from water [92]. Using a batched Bayesian search algorithm, the system automatically performed 688 experiments in eight days and found optimal photocatalyst mixtures. In the same year, a group from the University of Glasgow constructed the “Chemputer” system, which can automatically explore the chemical space and discover new molecules and reactions [93]. In 2023, Gomes et al. from Carnegie Mellon University designed the *Coscientist* system based on multiple LLMs agents that can automatically plan, analyze, and perform experiments [94]. The AI system could search online for the synthesis process and design the experiment pathways, successfully synthesizing ibuprofen from raw materials. Moreover, it can solve real-world complex experiments optimization processes such as the Pd-catalyzed transformations. It can accurately predict the ratio of the reactants and obtain the desired product within 20 experiment iterations.

1.5. Overview of this thesis

1.5.1. The main challenges and general goals

In this thesis, machine learning models are used for both enhancing materials simulations and property predictions. The primary goal of this thesis is to address the challenges inherent in simulations by utilizing machine learning to uncover new phenomena in materials and develop predictive models for their properties.

Two specific topics have been chosen for study. For the sections in machine-learning-assisted simulation, a deep learning force field was constructed for the ion diffusion phenomenon in all-inorganic metal halide perovskite was investigated; for the section in materials property predictions, a deep learning model for electron density prediction was constructed and studied.

1.5.2. Deep-learning force field for $CsPbI_3$

1.5.2.1. Basic introduction to halide perovskite

Halide perovskite materials have been found and investigated since the past century [95–

97]. However, these materials did not gain widespread attention and large amounts of research interest until Miyasaka et al. discovered their unique photovoltaic properties in 2009 [15]. Over the following decade, researchers have optimized the halide perovskite-based photovoltaic devices quickly and achieved high power conversion efficiency [98]. The halide perovskites are generally described by the formula ABX_3, in which A and B are cations, and X is the halide anion. Typically, the perovskites for photovoltaic and optoelectronic devices usage take MA^+, FA^+, or Cs^+ as A-site cation; Pb^{2+} or Sn^{2+} as B-site cation; and I^- or Br^- as X-site anion. In this thesis, only the inorganic Cs^+ as A-site cations have been considered.

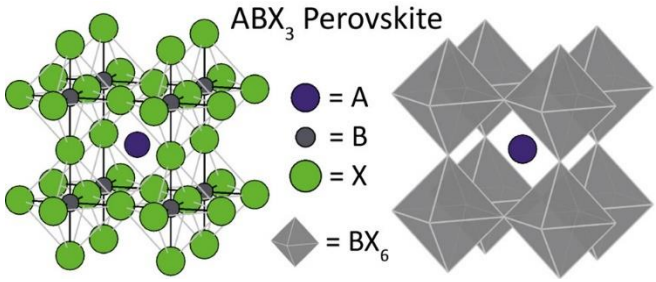

Figure 1.10 Schematics for the cubic halide perovskite structure. The left shows the atomic structure and the right shows the octahedra. Reproduced with permission from ref. [99]. Copyright © 2020 American Chemical Society.

As their name suggests, halide perovskites obtain the perovskite crystal structure, which obtains a framework of corner-sharing BX_6 octahedra with larger cation A inserted among them [100], as shown in **Figure 1.10**. Under room temperature, halide perovskite structures often exhibit lower symmetry compared to the cubic perovskite structure due to ion size mismatches. This results in spontaneous symmetry breaking, leading to tetragonal or orthorhombic lattice structures. The phase stability can be estimated by the Goldschmidt tolerance factor t [101,102] or the improved one τ for halide perovskite [103]:

$$t = \frac{r_A + r_X}{\sqrt{2}(r_B + r_X)} \tag{5}$$

$$\tau = \frac{r_X}{r_B} - n_A\left(n_A - \frac{r_A/r_B}{\ln(r_A/r_B)}\right) \quad (6)$$

Phase stability is critically important for the application of halide perovskites, as different phases exhibit varying efficiencies in light harvesting.

1.5.2.2. Ion diffusion in halide perovskite

Compared to other crystal materials, the halide perovskite lattices are relatively soft, so the ions inside them can move relatively more easily [104–106]. The active defect diffusion and migration can lead to numerous effects, such as phase transition or degradation, photocurrent hysteresis, non-radiative carrier recombination, etc. [107–109]. Thus, it is important to understand the ion diffusion phenomena in order to tune the device's performance.

So far, the computational method is necessary for studying the ion diffusion phenomena at a microscopic scale. Meanwhile, the simulations in defect structures normally require a large spatial scale and long timescale. Researchers need to choose between accuracy and efficiency.

The DFT calculations such as the nudged elastic band (NEB) [110] or AIMD are generally with first-principles accuracy but are limited to small supercells and a timescale for picoseconds. This can cause two problems. First, small supercells can lead to the interaction of defects themselves across the periodic boundary, leading to biased results. Second, a shorter timescale makes the final statistical results inaccurate and fails to detect rare events, which is important for transition state analysis and kinetic properties. On the other hand, MD simulations could give insight into thermodynamic properties such as activation energy for diffusion but lack of accuracy in interaction among atoms.

1.5.2.3. Contribution of this thesis

A dataset for $CsPbI_3$ all-inorganic halide perovskite was constructed from large numbers of

DFT results. Based on this, a deep-learning-based force field for MD simulation was trained, and we utilized the model to investigate iodide-related point defects in $CsPbI_3$. In Chapter 3, we performed the MD simulations and discovered a new kind of defect diffusion kinetics assisted by the anti-Frenkel disorder. The detailed pathways, energy barriers, and temperature-related properties were studied. In Chapter 4, we found out the planar confinement of iodide interstitials, which leads to different statistical properties and results in different diffusion patterns along crystal directions.

1.5.3. Electron density prediction

1.5.3.1. Basic introduction to electron density prediction

The electron density, as will detailly discussed in Chapter 2, is of fundamental importance for computational materials science and condensed matter physics. Based on the ground-state electron density, the other material's properties can be determined via density functional calculations. Therefore, the electron density was chosen to examine the predictability of machine learning rather than other properties.

Currently, there are several studies on developing machine learning models for electron density prediction, such as the Jacobi-Legendre charge density models [111] and ChargE3Net [112]. Several deep learning algorithms have been explored and show high accuracy towards specific material systems, such as the ordinary DNN with feed-forward structures and the graph neural networks with convolutional structures. In general, the prediction process requires the extraction of local environmental information. One approach is to construct a set of fingerprint functions for sampling the position and type of neighbouring atoms, while in the graph neural network approach, a set of atomic quantities are input and going through a sequence of equivariant functions. Both case requires a set of manually chosen functions. Since the resolution of the model for sampling local environmental information, these models could potentially reduce the accuracy in more complex systems. Besides, the predictability of the models is dependent on the functions chosen, which could make the generalization performance of the models worse. This makes one wonder if it is possible to design a DNN model with fewer manually chosen functions.

1.5.3.2. Contribution of this thesis

In this thesis, we have constructed a deep learning model for electron density prediction, named Deep Charge. The model takes use of an embedding net that can automatically extract the local environmental information, which is analogous to being able to find a set of sampling functions on its own. The model shows high accuracy towards DFT results with scalability to predict the electron density in larger supercells and is capable of representing electron density in various material systems. In Chapter 5, the detailed model construction and discussion of the ability of the model to keep the fundamental physical symmetries of the system. The Deep Charge model has been tested in various types of materials, including metal, semiconductor, and insulator, and also in various structures including crystalline bulk, surface, alloy, and amorphous structures. The model has shown a high accuracy towards all the systems and an obvious scalability in large-scaled systems.

Chapter 2.

Methodology

2.1. Density functional theory

2.1.1. The importance of density functional theory

Generally speaking, quantum mechanics provides us with the tools to compute the properties of materials from scratch. However, in real life, matters are collections of a macroscopically large number of particles, in the order of magnitude to the Avogadro constant N_A (6.02×10^{23}/mol). It is impossible to directly calculate their property from the fundamental atomic structures. On the other hand, even for calculations of the properties of a single atom, we may be faced with many-body interaction problems with tens or even hundreds of electrons. Calculating the properties obtains an exponential complexity increasing with the number of electrons that no one can truly solve, even for most of the atoms [113]. Density functional theory (DFT) is a well-known method for estimating the many-body problem in quantum mechanics calculations [114], making the ab initio calculations possible for us.

2.1.2. Harshness in Schrödinger equation

In a condensed matter system, plenty of nuclei and electrons. The stationary state of the N electrons and M atomic nuclei in the system can thus be represented by a many-body wavefunction $\Psi(\{\boldsymbol{r}_i\}_{i=1}^{N},\{\boldsymbol{R}_I\}_{I=1}^{M}) := \Psi(\boldsymbol{r}_1,\boldsymbol{r}_2,\dots,\boldsymbol{r}_N,\boldsymbol{R}_1,\boldsymbol{R}_2,\dots,\boldsymbol{R}_M)$ fulfilling the time-independent Schrödinger equation [115]:

$$\hat{H}|\Psi\rangle = \left(\hat{T}_e + \hat{T}_n + \hat{V}_{en} + \hat{V}_{ee} + \hat{V}_{nn}\right)|\Psi\rangle = E|\Psi\rangle, \qquad (7)$$

in which the electronic Hamiltonian $\hat{H}$ can be written as the sum of five parts: the kinetic energy operator for the electron and the nucleus $\hat{T}_e$ and $\hat{T}_n$, the electron-nucleus interaction $\hat{V}_{en}$, the electron-electron interaction $\hat{V}_{ee}$, and the nucleus-nucleus interaction $\hat{V}_{nn}$. Suppose there are N electrons and M nuclei, these operators can be written as:

$$\hat{T}_e = -\sum_{i=1}^{N} -\frac{\hbar^2}{2m_e}\boldsymbol{\nabla}_i^2 \tag{8}$$

$$\hat{T}_n = -\sum_{I}^{M} \frac{\hbar^2}{2M_I}\nabla_I^2 \tag{9}$$

$$\hat{V}_{en} = \sum_{i=1}^{N}\sum_{I=1}^{M} \frac{Z_I e^2}{|\boldsymbol{r}_i - \boldsymbol{R}_I|} \tag{10}$$

$$\hat{V}_{ee} = \sum_{i=1}^{N}\sum_{j=i+1}^{N} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} \tag{11}$$

$$\hat{V}_{nn} = \sum_{I}^{M}\sum_{J=I+1}^{M} \frac{Z_I Z_J e^2}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \tag{12}$$

where $\hbar$ is the reduced Planck constant, m_e is the mass of an electron, Z_I is the atomic number of the I's nucleus, e is the elementary charge, $\boldsymbol{r}_i$ is the coordinate of the i's electron, R_I is the coordinate of the I's nucleus, and M_I is the mass of the I's nucleus. To make the discussion simple, the spin degree of freedom is not included in the expression, while not affecting the results. The full differential equation includes the degree of freedom of all the electrons and nuclei, which is 3(M+N). In quantum mechanics, the superposition of single-particle states will provide an exponentially growing of parameters in describing the system [116], rather than a linear scaling in classical many-body systems. Thus, it is necessary to simplify the equation.

2.1.3. Born-Oppenheimer adiabatic approximation

The mass of a hydrogen nucleus is about 1836 times larger than that of an electron. Therefore, the movement of electrons is way faster than the nuclei. Based on this, we can assume that the nuclei are more or less "frozen" in space while the electrons are moving and relaxed fully to their favourable distributions, i.e. the time scales for the electronic and nuclear motion are significantly separated. This is the bedrock assumption of the Born-Oppenheimer (BO) approximation [117].

In the first step of BO approximation, we consider the kinetic energies of nuclei to be ignorable, and the interactions from frozen nuclei can be regarded as external potentials. The $\hat{T}_n$ term is subtracted from the total Hamiltonian. On the other hand, the Coulomb interaction among nuclei can be calculated externally and works as a constant that shifts the eigenvalues. Thus, the $\hat{V}_{nn}$ is also neglected from the original equation, leaving $\hat{H}_e = \hat{T}_e + \hat{V}_{en} + \hat{V}_{ee}$. Moreover, since the motion of electrons and nuclei can be calculated separately, the total wavefunction can be thus split into two parts:

$$\Psi(\{\boldsymbol{r}_i\}_{i=1}^{N}, \{\boldsymbol{R}_I\}_{I=1}^{M}) = \phi_e(\{\boldsymbol{r}_i\}_{i=1}^{N}; \{\boldsymbol{R}_I\}_{I=1}^{M})\phi_n(\{\boldsymbol{R}_I\}_{I=1}^{M}), \tag{13}$$

where ϕ_e and ϕ_n are the electronic and nucleus wave functions, respectively. Notice that although there are still nuclei coordinates in the electronic wave function, they now work as parameters rather than variables. The electronic Schrödinger equation becomes:

$$\hat{H}_e|\phi_e\rangle = E_e(\{\boldsymbol{R}_I\}_{I=1}^{M})|\phi_e\rangle \tag{14}$$

where E_e is the electron contribution to the total energy parametrized by the nuclei coordinates, also called the potential energy surface (PES).

After analyzing the motion of electrons, the motion of nuclei should be considered. Based on the separation of variables, chain rules, and several approximations [118], we can find the nucleus equation:

$$\left(\hat{T}_n + E_e + \hat{V}_{nn}\right)|\phi_n\rangle = E_{tot}|\phi_n\rangle \tag{15}$$

This means that once we obtain the potential energy surface of electrons, the motion of

nuclei can be updated accordingly.

2.1.4. Hohenberg-Kohn theorems

The BO approximation simplifies the many-body problem by reducing the degree of freedom from the separation of motions. We can only consider the 3N coordinates of electrons rather than 3(M+N). However, it is still a large number to deal with and cannot be calculated for larger atoms or systems.

The two Hohenberg-Kohn (HK) theorems take this direction a big step forward to further reduce the harshness [28,119,120], which utilized the ground-state electron density $\rho_0(\boldsymbol{r})$ with only three coordinates rather than all the electron positions:

$$\rho_0(\boldsymbol{r}) = \langle \phi_e | \hat{\rho}_0(\boldsymbol{r}) | \phi_e \rangle = N \int \left| \phi_e(\boldsymbol{r}, \{\boldsymbol{r}_i\}_{i=2}^N; \{\boldsymbol{R}_I\}_{I=1}^M) \right|^2 \prod_{i=2}^{N} d\boldsymbol{r}_i \qquad (16)$$

The first HK theorem states that the external potential $V_{ext}(\boldsymbol{r})$ of an interacting many-electron system is a unique functional of the $\rho(\boldsymbol{r})$ to within a constant, which means the $V_{ext}(\boldsymbol{r})$ can be written as $V_{ext}[\rho(\boldsymbol{r})]$. Since the external potential determines the Hamiltonian of the many-electron system, $\widehat{H}_e$ can thus be a unique functional of $\rho(\boldsymbol{r})$. The proof is described as follows [121].

Suppose there are two systems with only their external potentials that differ by more than a single constant, denoting as V_{ext} and V'_{ext}. We need to prove that the electron densities are different in the two systems. The electronic Hamiltonian of the two systems will be $\widehat{H}_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ext}$ and $\widehat{H}'_e = \hat{T}_e + \hat{V}_{ee} + \hat{V}'_{ext}$, which also differ only in the external potential term. The corresponding ground-state eigenfunctions ϕ_0 and ϕ'_0 for $\widehat{H}_e$ and $\widehat{H}'_e$ should also be different, and so is the ground-state energy E_0 and E'_0. According to the variational principle, the expectancy value for any trial wavefunction on the Hamiltonian should be larger than the ground-state energy, except for the specific ground-state eigenfunction. Thus, we could let the ϕ'_0 as a trial wavefunction on $\widehat{H}_e$ and obtain the relation:

$$E_0 < \langle \phi_0' | \hat{H}_e | \phi_0' \rangle = \langle \phi_0' | \hat{H}_e' | \phi_0' \rangle + \langle \phi_0' | \hat{H}_e - \hat{H}_e' | \phi_0' \rangle$$
$$= E_0' + \int \rho_0(\boldsymbol{r})[V_{ext}(\boldsymbol{r}) - V_{ext}'(\boldsymbol{r})]\, d\boldsymbol{r} \quad (17)$$

Meanwhile, we could also perform the same analysis with ϕ_0 and $\hat{H}_e'$:

$$E_0' < E_0 - \int \rho_0(\boldsymbol{r})[V_{ext}(\boldsymbol{r}) - V_{ext}'(\boldsymbol{r})]\, d\boldsymbol{r} \quad (18)$$

Adding these two equations will lead to a contradiction:

$$E_0 + E_0' < E_0' + E_0 \quad (19)$$

Therefore, different external potentials cannot lead to the same ground-state electron density, which proves the first HK theorem.

The second HK theorem states that the ground-state electron density minimizes the total energy of the system, which means the ground-state energy can be derived variationally. According to the first theorem, the Hamiltonian is determined by the electron density and thus determines the wave function. The wave function $\tilde{\phi}$ of a trial density $\tilde{\rho}(\boldsymbol{r})$ will lead to

$$\langle \tilde{\phi} | \hat{H} | \tilde{\phi} \rangle = E[\tilde{\rho}(\boldsymbol{r})] \geq E_0 = \langle \phi_0 | \hat{H} | \phi_0 \rangle \quad (20)$$

The equal sign only holds if the trial density is identical to the ground-state density.

The HK theorems reduce the 3N degrees of freedom to only three and provide the principle for finding the exact electron density, laying down the foundation for the density functional theory.

2.1.5. Kohn-Sham Equations

The dimensionality problem has been solved by the HK theorems, and now all the terms for system energy can in theory be expressed as functionals of electron density. However, there is still one complex problem: how to find the proper functional?

In 1965, Kohn and Sham proposed the famous Kohn-Sham (KS) equations to describe the

energy functional [26]. In their theory, the total energy functional of the electrons is expressed as:

$$E[\rho] = T_s[\rho] + \int V_{ext}(\boldsymbol{r})\rho(\boldsymbol{r})d\boldsymbol{r} + E_H[\rho] + E_{XC}[\rho], \quad (21)$$

where T_s is the KS kinetic energy, E_H is the Hartree (or Coulomb) energy, and E_{XC} is the exchange-correlation (XC) energy. Except for the E_{XC} term, the other functionals can be calculated via direct integration or summation over orbitals. E_{XC} contains all the hard-solving parts of the total functional, and the former problem further reduces to estimate the expression of E_{XC}.

The construction of an accurate exchange-correlation functional is still an active area in computational physics. One important class of approximations is the local-density approximations (LDA), which is included in the original paper on KS equations [26]. LDA estimates the XC energy at each point in space according to homogeneous electron gas with the same density on that point:

$$E_{XC}^{LDA}[\rho] = \int \varepsilon_{XC}(\rho)\,\rho(\boldsymbol{r})d\boldsymbol{r}, \quad (22)$$

where ε_{XC} is the XC energy for homogeneous electron gas. Based on the LDA approximations, more accurate generalized gradient approximations (GGA) can be constructed by considering the gradient terms of the density [122,123]. More accurate XC functionals are actively developing and some have already been used for DFT calculations, such as the meta-GGA [124–126].

2.2. Molecular Dynamics

2.2.1. Introduction to the classical molecular dynamics

MD is an old technique that has been used in a variety of scientific fields such as physics, materials science, chemistry, and biology. It has also been used to investigate properties from microscopic to mesoscopic scales. The goal of MD simulations is to analyze the time

evolution of the basic elements of a system, such as atoms, molecules, or nanoclusters, and give the dynamics and statistical properties of the system over a period of time [127].

The first MD simulation was proposed by Enrico Fermi et al. in 1953 to study the ergodic hypothesis in artificial systems with several different force laws [128]. Later in 1957 Alder and Wainwright simulated the elastic collisions in a hard sphere system [25]. The first known MD simulation on real systems was performed by Vineyard et al. in 1960, which investigated the radiation damage in solid copper with a Born-Mayer-type interaction [129]. Later in 1964, Rahman applied MD simulations to study the self-diffusion in liquid argon with a Lennard-Jones potential and showed good agreement with the experimental data [130]. Now, it has been used in much more complicated and massive systems, such as protein folding [131], phase transition of materials [132], and chemical reactions [133].

The classical MD simulation is based on numerically solving the Newtonian mechanics for the N-body problem. Suppose we know the interatomic potential between atoms, the force $\boldsymbol{F}_i$ on atom i can be derived via the gradient of the total potential V_{tot} "felt" by that atom:

$$\boldsymbol{F}_i = -\boldsymbol{\nabla}_{\mathbf{r}_i} V_{tot} \tag{23}$$

After obtaining the forces on atoms, the position and velocity for the next timestep of MD simulation can be calculated by solving the equation of motion.

2.2.2. Interatomic potentials

The core for MD simulations is the interatomic potentials (also called force fields in chemistry), based on which the system could evolve with time. Thus, an accurate interatomic potential is of vital importance for describing the target system. Many of the classical MD simulations were based on empirical potentials, which obtained parameters that were manually fitted to align with experimental or first-principles data. Formally, the potential term of a N-body system with atom coordinates $\{\boldsymbol{r}_i\}_{i=1}^{N}$ can be expanded into a summation of contributions [134]:

$$V(\{\boldsymbol{r}_i\}_{i=1}^N) = \sum_{i=1}^{N} V_1(\boldsymbol{r}_i) + \sum_{i<j}^{N} V_2(\boldsymbol{r}_i, \boldsymbol{r}_j) + \sum_{i<j<k}^{N} V_3(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) + \cdots, \quad (24)$$

where V_n represents the n-body term of the interaction. The one-body term can be regarded as the contribution of external potentials such as the electric field, and the two-body term is generally called the pair potential, which is the most used interaction. When there is no external influence on the system, the symmetries in the physical system will reduce the variables. In this case, the potential should only depend on the relative position between the atoms to fulfil the translational symmetry. Also, in many-body terms we could use the angle information between neighbors since the rotational symmetry exists:

$$V(\{\boldsymbol{r}_i\}_{i=1}^N) = \sum_{i<j}^{N} V_2(r_{ij}) + \sum_{i<j<k}^{N} V_3(r_{ij}, r_{jk}, \theta_{ijk}) + \cdots, \quad (25)$$

in which $r_{ij} = |\boldsymbol{r}_{ij}| = |\boldsymbol{r}_j - \boldsymbol{r}_i|$, and θ_{ijk} is the angle between $\boldsymbol{r}_{ij}$ and $\boldsymbol{r}_{jk}$. There have been plenty of interatomic potentials among various systems, and they can be classified into two groups: the parametric and non-parametric potentials.

The parametric potentials are developed with several parameters that can be manually tuned to fit the data. One of the simplest parametric potentials that is widely used is the 12-6 Lennard-Jones (LJ) potential [135]:

$$V_{LJ}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right], \quad (26)$$

where ε is the depth of the potential well, and σ is the distance for repulsion-attraction transition. These two parameters can be adjusted to tune the interatomic distance in the simulation. The first term with exponent 12 describes the repulsion force on atoms when they are in a short distance, while the second term with exponent 6 describes the London dispersion force [136]. LJ potential yields good results in simulations of noble gases.

Another example is the Coulomb-Buckingham potential used in ionic systems [137,138]:

$$V_{CB}(r) = Ae^{-r/\rho} - \frac{C}{r^6} + \frac{q_1 q_2}{4\pi\varepsilon_0 r}, \quad (27)$$

where A, ρ and C are the constants to be fitted, q_1 and q_2 are the charges of the ions in the pair. The first two terms together are the Buckingham potential describing the Pauli repulsion and van der Waals interaction between atoms. The last term describes the electrostatic potential.

On the other hand, the non-parametric potentials, where there are no parameters to be found by hand. Generally, these potentials are constructed by machine learning [139]. During the training and predicting process, the model firstly constructs a set of representation $\{\boldsymbol{q}_i\}_{i=1}^{N}$ of the local environment of the N atoms from their coordinates and types, and then the model predicts the total potential energy via calculating the contribution ε on each atom in the system [140]:

$$V_{tot} = \sum_{i=1}^{N} \varepsilon(\boldsymbol{q}_i) \tag{28}$$

Generally, the machine learning models obtain a large number of degrees of freedom to be learned from data. Thus, they could provide more intrigue structures of the potential energy surface than the parametrized one and give more accurate results. However, the extrapolation ability of them should be tested carefully.

2.2.3. Ensembles

Similar to the real-life experiments, the MD simulation should be performed inside some specific environment, where the conditions could be controlled manually. These environmental factors could constrain the state of the system. The thermodynamic ensembles represent all the possible states of the system in statistical equilibrium in specific macroscopic constraints [141]. During the simulation, the thermostat and/or barostat are used to control the velocity of atoms to constrain the temperature or pressure of the system as desired. The thermostat and barostat can also keep the system stable throughout the numeric simulation, owing to the monitor and adjust of atom velocities. Therefore, it is of vital importance to choose an appropriate ensemble before the MD simulations.

2.2.3.1. Canonical ensemble (NVT)

In the canonical ensemble, the number of atoms N, the volume of the system V and the temperature T are conserved. During the simulation, the instantaneous temperature of the system is fluctuated. Thus a thermostat is used to detect and control the temperature change inside the system. Most thermostats are constructed by adjusting the equation of motion of the system by adding a "friction" term. For example, the adjusted equation of motion at time t for the i^{th} atom can be written as [142]:

$$\ddot{\boldsymbol{r}}_i(t) = m_i^{-1}\boldsymbol{F}_i(t) - \gamma(t)\dot{\boldsymbol{r}}_i(t), \tag{29}$$

where $\boldsymbol{F}_i(t)$ is the total force on that atom at time t, $\gamma(t)$ is similar to the "friction coefficient" in the Langevin equation but is not confined to positive values. $\gamma(t)$ determines the energy flow of the system. If $\gamma(t)$ is negative, the energy will be added into the system and vice versa. The $\gamma(t)$ is then constructed related to the temperature of the system, in order to control the velocity accordingly. For example, $\gamma(t)$ in the Hoover-Evans thermostat is derived as [142–144]:

$$\gamma(t) = \left[N_{dof}k_B T(t)\right]^{-1}\sum_{i=1}^{N}\dot{\boldsymbol{r}}_i(t)\cdot\boldsymbol{F}_i(t) \tag{30}$$

where N_{dof} is the degree of freedom of the system, k_B is the Boltzmann constant.

2.2.3.2. Isothermal–isobaric ensemble (NPT)

In the Isothermal–isobaric ensemble, the volume of the system does not remain constant as the canonical ensemble. Instead, the simulations adopt a barostat to control the pressure P around a constant value. There are several methods to control the pressure, such as the Andersen barostat [145], the Nose-Hoover style non-Hamiltonian equations of motion [146–148] and the Parrinello-Rahman method [149]. NPT ensemble is the most like the real-life experiment conditions that are coupled with the atmosphere pressure [150]. NPT ensemble is necessary for studying the crystallization of solid from liquid or the solid-state phase transition, which could require large deformation of the original simulation

cells [151,152]. Moreover, it is also important to investigate the mechanical properties of materials under external stress [153].

2.2.3.3. Microcanonical ensemble (NVE)

The Microcanonical ensemble conserves the number of atoms, the system volume, and the total energy. Thus, the system is isolated and does not exchange energy with the environment. Besides the energy, the momentum of the system is also conserved due to the lack of external forces [154]. The NVE ensemble is important if we want to study isolated systems or control the energy of the system [155].

2.3. Deep Learning Interatomic Potentials

As mentioned in **Chapter 1**, there are currently several machine-learning strategies for constructing interatomic potentials for MD simulations. In this thesis, the deep potential (DP) framework was adopted for the studies. Thus, in this section, I mainly describe the construction of interatomic potential following the DP architecture.

2.3.1. Deep potential framework

The deep potential framework is embedded in the DeePMD-kit package constructed with TensorFlow [156]. The core of the framework is the machine learning model. To train or apply the model in simulation, the trainer and inference sections are required. As shown in the workflow of **Figure 2.1**, the model can be used to predict the energy, force, etc. from the input information, which includes the sets of atom coordinates, atomic types, and boundaries. This is the inference section. In the training process, the predicted properties should be further directed to the trainer to compare with the reference values with the loss function. The parameters of the model are then tuned by the optimizer.

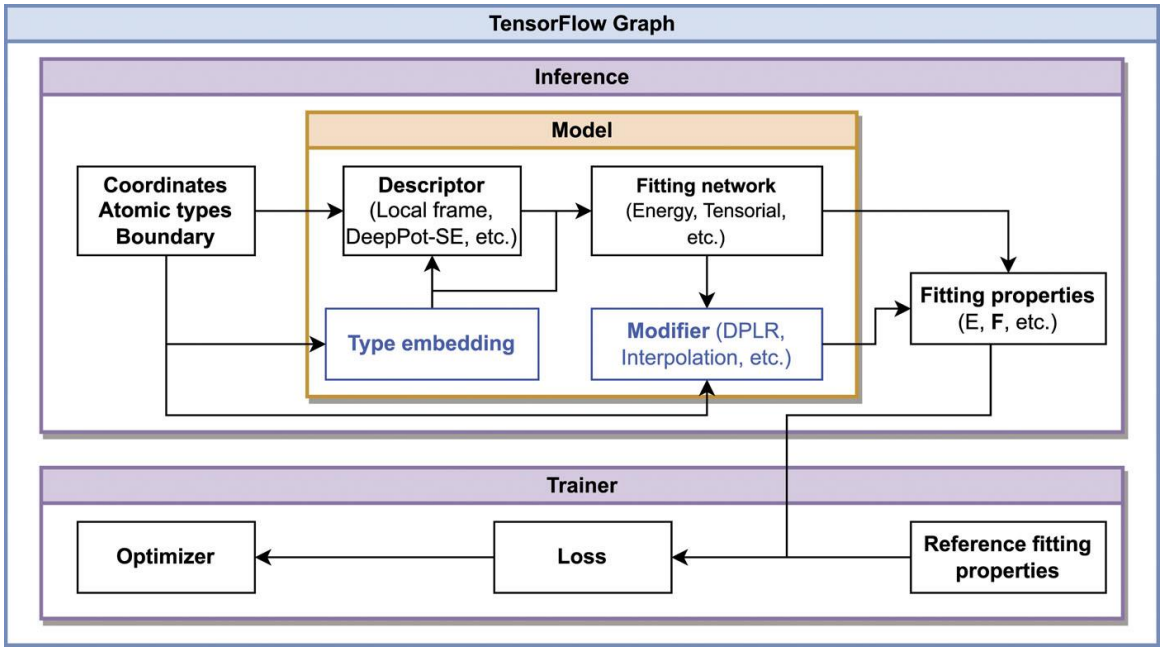

Figure 2.1 The general workflow of the deep potential framework. The blue boxes are optional for the model. Reproduced with permission from ref. [157]. © Copyright 2024 AIP Publishing.

2.3.2. The model architecture

The most important part of the deep potential framework is the machine learning model. In general, the prediction of an atomic property y on the atom i can be represented as:

$$\boldsymbol{y}_i = \mathcal{M}\left(\boldsymbol{x}_i, \{\boldsymbol{x}_j\}_{j\in n(i)}; \boldsymbol{\theta}\right) = \mathcal{F}\left(\mathcal{D}\left(\boldsymbol{x}_i, \{\boldsymbol{x}_j\}_{j\in n(i)}; \boldsymbol{\theta}_d\right); \boldsymbol{\theta}_f\right) \tag{31}$$

where $\mathcal{M}$ is the DP model, $\mathcal{F}$ is the fitting net, and $\mathcal{D}$ is the descriptor. $\boldsymbol{x}_i = (\boldsymbol{r}_i, a_i)$ is the 4D vector containing the Cartesian coordinate $\boldsymbol{r}_i$ and the atomic species a_i, denoting the degrees of freedom of the atom i. $\boldsymbol{\theta} = \{\boldsymbol{\theta}_d, \boldsymbol{\theta}_f\}$ is all the network parameters, in which $\boldsymbol{\theta}_d$ and $\boldsymbol{\theta}_f$ are the parameters for the descriptor and the fitting net, respectively. The global properties such as the system energy can be computed via summation over the corresponding atomic property:

$$\boldsymbol{y} = \sum_{i=1}^{N} \boldsymbol{y}_i$$

The DP model is constructed via multi-layered neural networks, each layer may have a residual network (ResNet) structure [158] or an ordinary neural network. For the descriptor part, several different kinds of architecture can be implemented, such as the local frame descriptor [159], two-body embedding smooth edition *se_e2_a* [160], three-body embedding smooth edition *se_e3* [161], etc. In the studies of this thesis, the descriptor *se_e2_a* with full environmental information is generally adopted, which is a many-body representation of the local environment that takes the distance between neighbouring atoms as the input information. The descriptor is constructed to obey several fundamental symmetries:

1. **Translational Symmetry**. The translation of the whole system should not influence the prediction of physical quantities.
2. **Rotational Symmetry**. The predicted scalar quantities should be invariant of the rotation of the whole system. Moreover, the vector quantities should be equivariant with the rotation.
3. **Permutational Symmetry**. The predicted values should be independent of the order of atoms in the neighbour list.

The detailed form of the descriptor *se_e2_a* will be discussed later in **Chapter 5**. After obtaining the representation of the local environment, the information is further input to the fitting net, which will predict the potential energy of the system by summation over the contribution of atomic energy. The fitting net is a fully connected feed-forward network that outputs a zero-order tensor $\mathcal{F}_0$. The energy is described as:

$$E = \sum_{i=1}^{N} E_i = \sum_{i=1}^{N} \mathcal{F}_0(\mathcal{D}^i) \tag{32}$$

Based on the energy, we could also obtain the atomic force by differentiating:

$$F_{i,\alpha} = -\frac{\partial E}{\partial r_{i,\alpha}} \tag{33}$$

where the α denotes the α-th component of the vector. Moreover, if the simulation obtains a periodic boundary condition, the virial tensor can be derived as:

$$\Xi_{\alpha\beta} = \sum_{\gamma} \frac{\partial E}{\partial h_{\gamma\alpha}} h_{\gamma\beta} \tag{34}$$

where $h_{\alpha\beta}$ is the β-th component of the α-th basis vector of the lattice.

2.3.3. Loss function

In brief, the loss function used in the DeePMD-kit is constructed by the mean square error (MSE) among the predicted values. For example, the MSE for energy and forces can be written as:

$$L_E(\{\boldsymbol{x}_i\}_{i=1}^N; \boldsymbol{\theta}) = \frac{1}{N}(E - E^*)^2, \tag{35}$$

$$L_F(\{\boldsymbol{x}_i\}_{i=1}^N; \boldsymbol{\theta}) = \frac{1}{3N}\sum_{i=1}^{N}\sum_{\alpha=1}^{3}\left(F_{i,\alpha} - F_{i,\alpha}^*\right)^2, \tag{36}$$

where the symbols with a superscript asterisk denote the labels, or real values, of the corresponding predicted values. The general loss function for one frame of data can be written as:

$$L(\{\boldsymbol{x}_i\}_{i=1}^N; \boldsymbol{\theta}, \tau) = \sum_{k} p_k(\tau) L_k(\{\boldsymbol{x}_i\}_{i=1}^N; \boldsymbol{\theta}) \tag{37}$$

where $p_k(\tau)$ is a prefactor to determine the weight of the specific type of loss. It can be related to the training step τ, meaning that the weight of different parts of the loss can gradually change throughout the training process.

2.4. Datasets Construction

A high-quality dataset is essential to train an interatomic potential that effectively represents particular matter systems. Meanwhile, since the construction of the dataset

requires a large amount of DFT calculations, there are two issues when collecting data:

1. How to reduce the size of datasets via discarding redundant and similar data.
2. How to determine that the datasets are sufficient for representing the systems.

In this section, I briefly describe the workflow of the Deep Potential Generator (DP-GEN) for dataset construction [162]. In general, the DP-GEN uses a concurrent learning framework for automatically determining the convergence of the dataset, based on which the code will decide whether or not to add new data to the dataset.

2.4.1. Concurrent learning framework

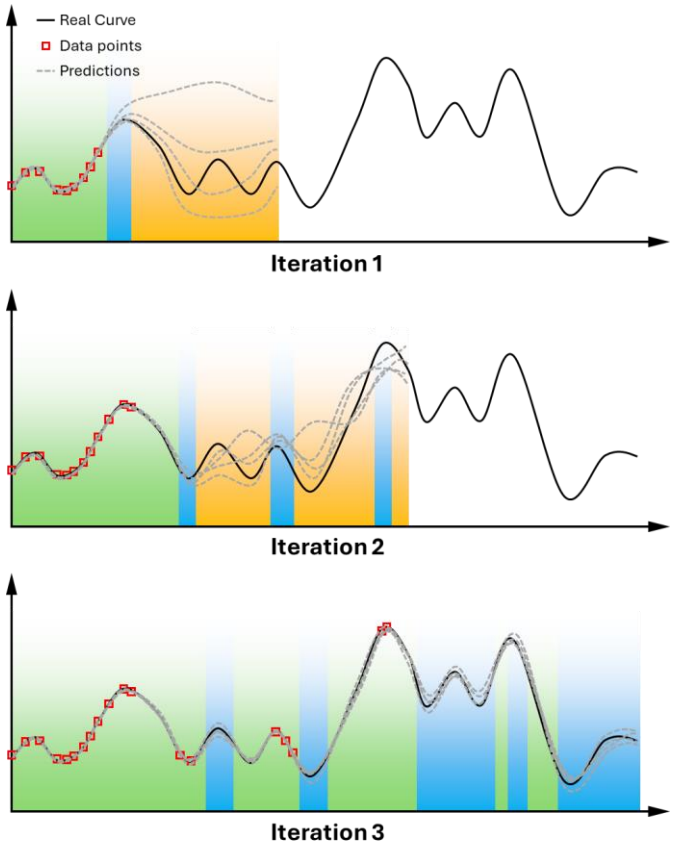

Figure 2.2 Schematics of concurrent learning for fitting a function. The areas marked green, blue, and yellow denote the accurate, candidate, and failed areas for fitting, respectively.

Usually, the datasets for deep learning interatomic potential are constructed by several AIMD trajectories of the target systems. Meanwhile, this approach has intrinsic flaws. Firstly, in order to collect enough data from the configuration space of the systems, a large number of long-period AIMD simulations should be performed. Secondly, In the AIMD simulations, the configurations of nearby timesteps are similar, which is possible to be redundant for the training process. To circumvent the problems, the concurrent learning procedure was proposed [162].

Figure 2.2 demonstrates the basic idea of concurrent learning. The datasets are created through an iterative way. Suppose we would like to fit a complicated function (the black line), the traditional way is to randomly sample datapoints along the line, generating large datasets and reducing the training efficiency. This is also hard to perform when we have no information about the required range of the function. In the concurrent learning process, we first collect a few datapoints from a small range of the function, as shown in the topmost plot in **Figure 2.2**. Then, multiple DP models (four in this case) are simultaneously trained and initialised with different neural network parameters. Since the dataset in this iteration cannot fully cover the range we are interested in, the prediction of models in the extrapolate region is highly possible to be different from the real values (the grey dashed lines). Moreover, the four DP models will provide different predictions for the extrapolate region, due to varying initial parameters. We now need to construct an indicator to quantify the accuracy among predictions of models, denoted as ϵ. If ϵ is smaller than a threshold σ_{lo}, the predicted value is regarded as accurate (the green areas in **Figure 2.2**); if it is larger than another threshold σ_{hi}, then regarded as "totally wrong" (the yellow areas); and if it is in between the two thresholds, the datapoint corresponding to that prediction is regarded as a candidate for collection (the blue areas). Next, the datapoints labelled as candidates are collected for the next iteration of concurrent learning. The collection procedure is considered finished when all the datapoints are within the accurate threshold.

This concurrent learning framework will dramatically reduce the dataset size by ignoring redundant data marked as accurate, and automatically determine whether the dataset is sufficient for a system.

2.4.2. Workflow of DP-GEN

DP-GEN has implemented an integrated workflow for exploring, labelling, and training DP models, as shown in **Figure 2.3**. The schedular determines when jobs are submitted and confirms job completion. Based on an initial AIMD dataset, the schedular firstly trains several initial DP models in parallel. The initial DP models are then used in DPMD simulations in the interested systems. The trajectories are analyzed, and the schedular determines whether a frame requires to be collected based on the ϵ defined as:

$$\epsilon = \max_i \sqrt{\frac{1}{N_m}\sum_{\alpha=1}^{N_m}\left(F_{\alpha,i}(\mathcal{R}) - \frac{1}{N_m}\sum_{\beta=1}^{N_m}F_{\beta,i}(\mathcal{R})\right)^2}, \tag{38}$$

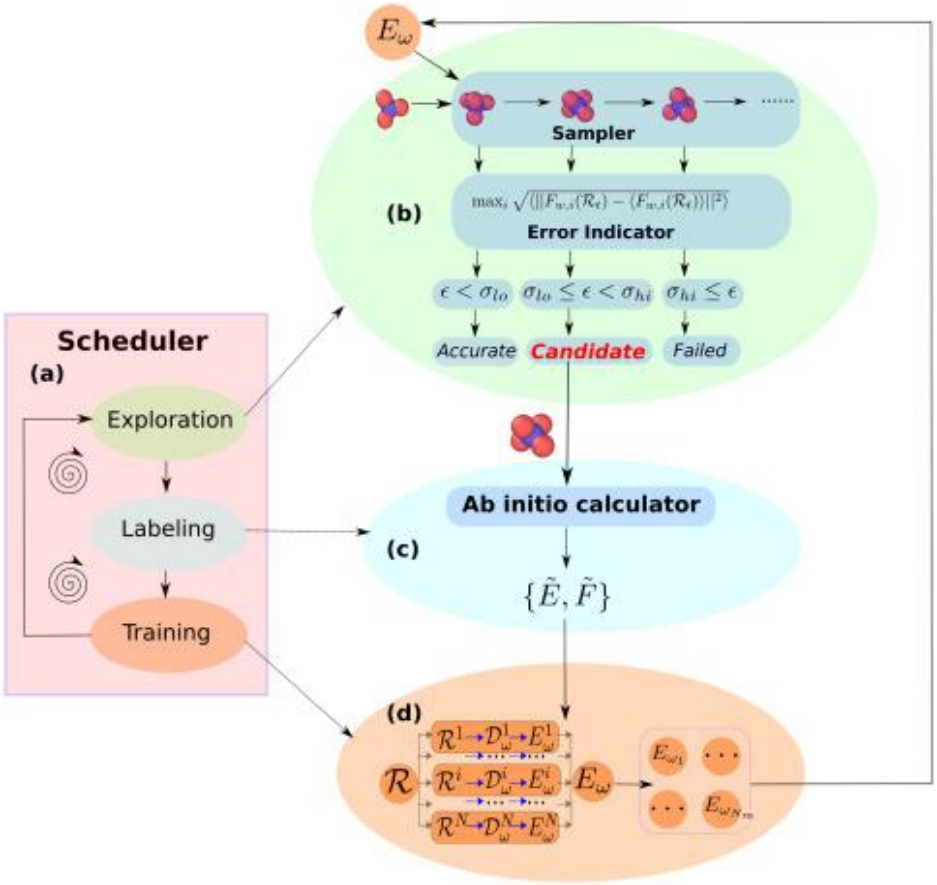

Figure 2.3 Schematic for the workflow of DP-GEN. Reproduced with permission from ref. [162]. Copyright © 2020 Elsevier B.V.

where N_m is the number of DP models, $\mathcal{R}$ is the configuration of the specific frame, and $F_{\alpha,i}(\mathcal{R})$ is the i-th component of the predicted force on configuration $\mathcal{R}$ using the α-th model. In brief, the code uses the maximum force error between the models to determine the accuracy for each frame. If the ϵ of a frame is inside the candidate region, this frame is collected for the labelling process.

After the exploration, the collected structures are then automatically transformed into the input format for Ab initio calculators. The single-point energy calculations with the set parameters are performed. Then the scheduler collects the energy and forces (and virials, if required) from the calculation results, and transforms them into datapoints. The new dataset containing collected data is then used for training DP models for the next iteration.

When all the frames in the simulation range are marked accurate, the DP-GEN procedure will stop. The final dataset can be constructed with all the initial datasets and the collected datapoints. Based on this dataset, the DP models with more complex structures and parameters can be trained for further accurate simulations.

Chapter 3.

Kinetic Process with Anti-Frenkel Disorder in $CsPbI_3$ Perovskite

3.1. Abstract

Halide perovskites are rich in ionic diffusion phenomena due to their low activation energy. The soft lead-iodide lattice can in theory endow the system with more complex defect collaborative motions. In this work, we systematically investigated the hopping mechanics of iodide interstitials among various defect behaviours such as anti-Frenkel disorder creation and annihilation. We found that the existence of iodide vacancies and interstitials can effectively lower the creation barrier of additional anti-Frenkel disorder in halide perovskite. The free energy barriers for generating additional Frenkel defect pairs vary range from 0.25 to 0.43 eV in proximity to original iodide defects at 300 K. This finding suggests that the spontaneous creation of a specific level of anti-Frenkel disorder facilitates long-range annihilation and defect hopping processes.

3.2. Introduction

Point defects play important roles in the performance and stability of halide perovskite materials. The soft lattice and solution-based fabrication methods commonly employed in these materials result in the presence of numerous defects within the halide perovskite lattice, with a density ranging from 10^{10} to 10^{18} cm^{-3} [163–165]. These defects will further assist the ion diffusion process, causing the hysteresis of current-voltage curves and thus resulting in efficiency and performance decrease in perovskite-based optoelectronic devices and field-effect transistors [166–169]. Furthermore, mobile defects could

accumulate into clusters, thereby promoting the degradation of halide perovskites [170]. Meanwhile, fast ion diffusion can assist the development of ion conductors and resistive switching devices [171,172]. Previous studies have highlighted that defects related to halide anions such as iodide dominate the fast diffusion process due to their low activation energy [173–175]. This diffusion is closely related to two defects: iodide interstitial (I_i) and vacancy (V_I). Among them, iodide vacancies have been extensively researched [176,177], in that they have been regarded as the dominant halogen-related defects contributing to ion diffusion, due to the low formation enthalpy (0.08 ~ 0.14 eV) via the "full" or "partial" Schottky disorder reaction, as shown in Eqs. 1 and 2 with Kröger–Vink notation [174,178]:

$$\emptyset \rightleftharpoons V_A' + V_M'' + 3V_X^{\cdot} + AMX_3 \quad (39)$$

$$\emptyset \rightleftharpoons V_A' + V_X^{\cdot} + AX \quad (40)$$

where $\emptyset$ represents the ideal perovskite lattice, A is the large cations (Cs^+, methylammonium (MA^+), etc.), M is the metal cations (Pb^{2+}, Sn^{2+}, etc.), and X is the halide anions (I^-, Br^-, etc.). This phenomenon is also applicable to perovskite oxides, including $SrTiO_3$ [179,180].

On the contrary, the creation of interstitial is mainly through anti-Frenkel disorder (*aF*) (Eq. 3), with a formation enthalpy of 0.82 eV [181].

$$\emptyset \rightleftharpoons V_X^{\cdot} + X_i' \quad (41)$$

This is much higher than the Schottky disorder and has thus received less attention in previous studies. However, research has indicated that *aF* becomes predominant in $MAPbI_3$ at temperatures higher than 290 K [182,183]. The additional *aF* can probably decompose into I_i and V_I, leading to a fast ion diffusion [175,184] and providing additional charge carrier traps [185]. Nevertheless, it is still questionable about the detailed pathways of *aF* creation and I_i diffusion, especially in the all-inorganic $CsPbI_3$ perovskite. Therefore, the finite-temperature behaviour of halide in perovskite still requires systematic research, which is vital for understanding ion diffusivity in halide perovskites.

Herein, we implement deep potential molecular dynamics (DPMD) [159] together with metadynamics [186] to study the microscopic motion pathways of iodide interstitials and

vacancies in $CsPbI_3$. Our analysis delves into the intricate hopping pathways of iodide-related defects. We discover that the existence of I_i or V_I reduces the creation barrier of *aF* by 0.5 eV relative to that in ideal crystals under 0 K. Furthermore, employing enhanced sampling techniques, we determine that the free energy barriers for the *aF*-assisted OP hopping of I_i decrease even further to the range of 0.25-0.43 eV within the temperature range of 300-600 K. This finding suggests the feasibility of high-temperature pathways for *aF* formation around existing defects. Notably, the formation of *aF* not only facilitates the long-range annihilation and self-diffusion of I_i and V_i but also disrupts the planar motion behaviour of I_i. The insights enhance our understanding of self-diffusion phenomena in halide perovskite materials, beneficial to the development of stable perovskite-based optoelectronic devices and ion conductors.

3.3. Results and Discussion

To efficiently capture the finite temperature hopping events of iodide-related defects, we first ran several isothermal–isobaric (NpT) ensemble DPMD simulations above 500 K. Under these temperatures, $CsPbI_3$ adopts a cubic structure in our simulations, as shown in **Figure 3.10** of **Section 3.5**. Thus, we can ignore the anisotropy of defects across different lattice directions. In general, the hopping of I_i occurs more frequently than that of V_I, ascribing to a lower activation energy.

The sampled hopping pathways for I_i and V_I are systematically presented in **Figure 3.1**. Notably, I_i prefers to align parallel within the Pb-I planes, indicating the ability to adopt two distinct orientations at the same site. I_i obtains two primary kinds of pathways: in-plane (IP) and out-of-plane (OP). The IP hopping occurs most frequently by rotating centred on an adjacent Pb atom, as shown in the first column of **Figure 3.1**. This process results in short-term planar motion of I_i. Conversely, OP hopping changes the orientation of I_i and can be achieved through two distinct pathways. The first is on-site OP hopping via rotation around the crystal axis where I_i is located, leaving the position of I_i unchanged (second column in **Figure 3.1**). The second pathway, off-site OP hopping, involves simultaneous changes in position and orientation (third column in **Figure 3.1**). Meanwhile, V_I exhibits a simpler structure, leading to only one type of hopping pathway.

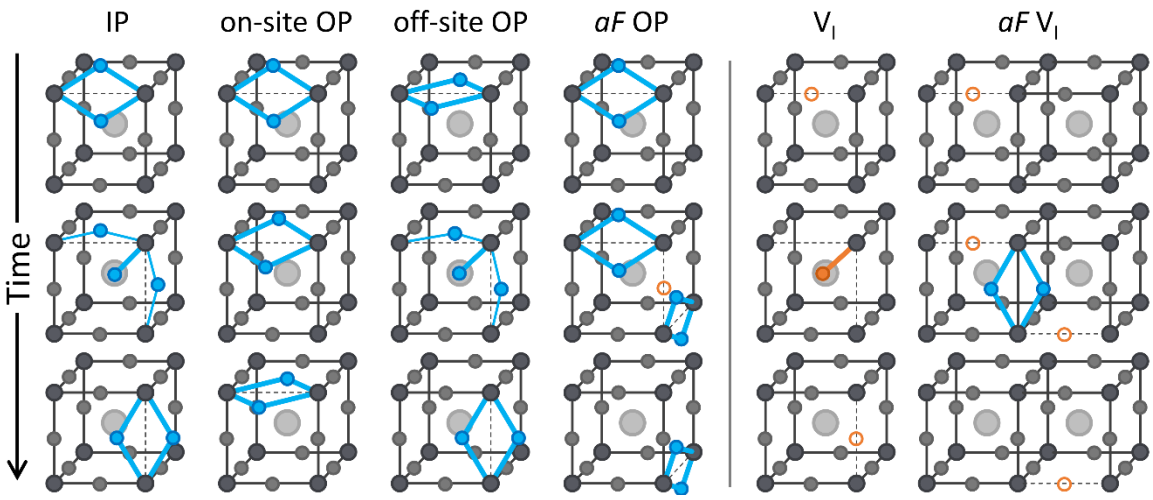

Figure 3.1 Schematics for different I_i and V_I hopping pathways sampled from MD simulations. The left four and right two columns are I_i and V_I hopping events, respectively. Time increases from top to bottom. The blue and orange circles demonstrate the iodides taking part in the I_i or V_I hopping, respectively. The orange hollow circles denote the approximate position of V_I.

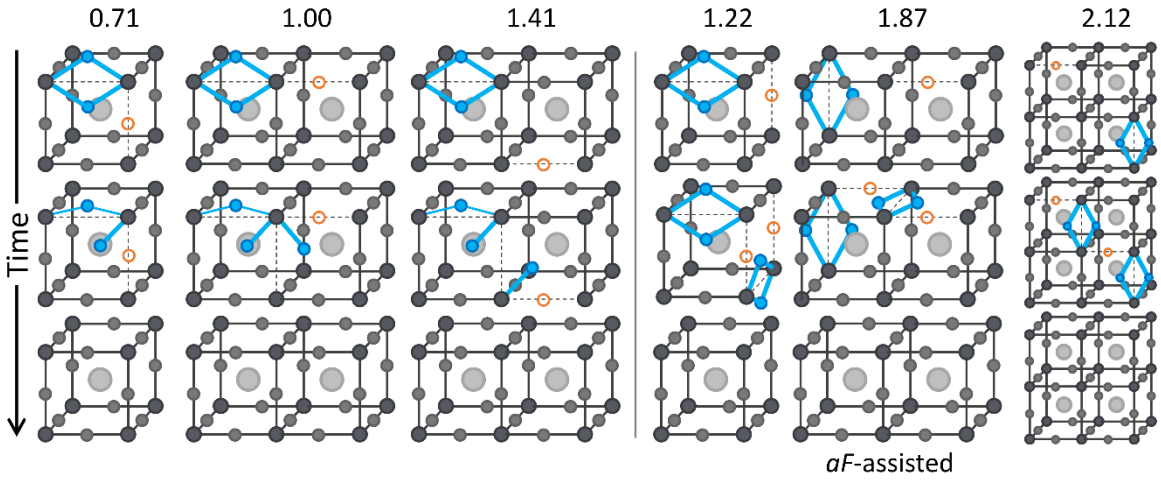

Figure 3.2 The different pathways for I_i - V_I annihilation. The left and right three columns are annihilations without and with the assistance of aF, respectively. The number on the top of each column denotes the approximate relative distance between I_i and V_I in the unit of cubic cell lattice.

Besides the conventional pathways, our study reveals that *aF* could further enhance the kinetics of the defect, as shown in the fourth and final columns of **Figure 3.1**. Specifically,

the off-site OP of I_i and V_I hopping can be accompanied by creating an I_i-V_I pair in their vicinity. Then the original point defects annihilate the newly formed counterparts, remaining a defect on the other side and finishing a long-range hopping event.

We further studied the annihilation of *aF*, i.e., the I_i-V_I pair. The annihilation events can be categorized by the distance between I_i and V_I. The shortest annihilation path occurs inside one unit cell (first column of **Figure 3.2**). The initial conformation is highly unstable. Once I_i and V_I encounter each other, the annihilation will occur within a few picoseconds. Almost all the *aF* created by thermal fluctuation disappears in this form. Long-range annihilation typically occurs across multiple unit cells and involves the cooperative motion of two or more iodides. The first type shown in the second column of **Figure 3.2** has been previously studied [187,188]. It can occur rapidly by rotating around Pb atoms, displaying relatively higher stability compared to short-range configurations. Besides, another long-range annihilation pathway emerges by involving the simultaneous occurrence of fast IP hopping of I_i and short-range annihilation. Iodides are like the teeth of two interlocked gears, rotating around the Pb axis. The exact mode of annihilation depends on the conformation in which I_i and V_I meet.

Moreover, similar to the defect hopping events, the *aF* can also assist the long-range annihilation process. The processes can be classified into three types based on the distance between I_i and V_I, as shown in the last three columns of **Figure 3.2**. Type 1 involves one single cubic unit cell, type 2 involves two unit cells, and type 3 is in three unit cells. An additional *aF* can be created between the existing I_i-V_I pair, with the newly generated I_i and V_I residing in proximity to the initial V_I and I_i, respectively. With the help of *aF*, annihilation could happen over long distances. Moreover, the I_i and V_I in types 1 and 2 are not in the same lattice plane, indicating the potential interactions between the V_I and OP I_i

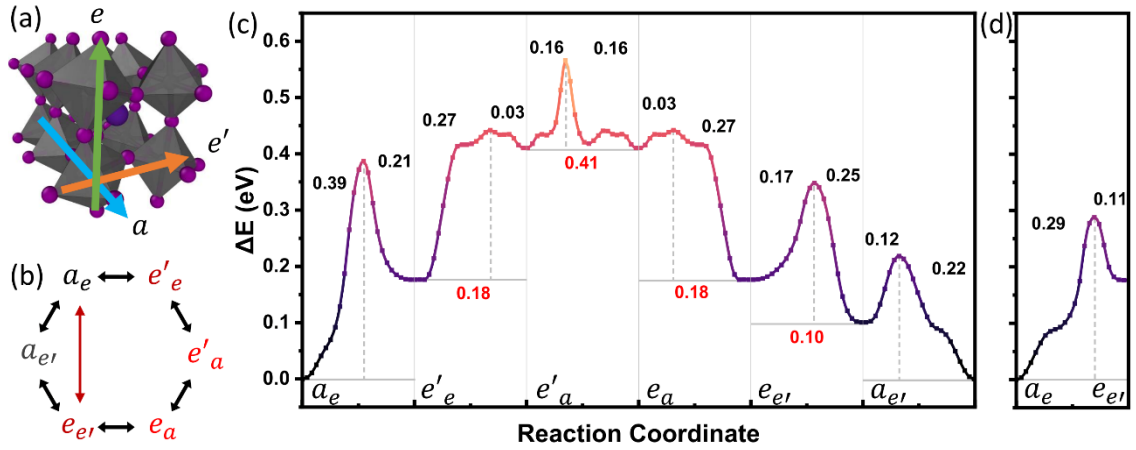

Figure 3.3 The NEB energy barrier of I_i IP and OP hopping in different conformations. (a) Schematic for the apical (a) and two equatorial (e and e') directions in the orthorhombic perovskite structure. We denote the shorter equatorial direction as e'. (b) Schematic for possible one-step pathways between defect states. The colour represents the relative energy of different defect states, in which e_a and e'_a obtain the highest energy, and ae the lowest energy. The red arrow denotes the one-step hopping outside the original axis and plane. (c) The energy barriers for the defect hopping process along different axes, in which $a_e \leftrightarrow e'_e$, $e'_a \leftrightarrow e_a$ and $e_{e'} \leftrightarrow a_{e'}$ are IP transitions; $e'_e \leftrightarrow e'_a$, $e_a \leftrightarrow e_{e'}$ and $a_{e'} \leftrightarrow a_e$ are OP transitions. The red font indicates the relative energies of the states and the black font indicates the energy barriers. (d) the energy barrier for $a_e \leftrightarrow e_{e'}$ one-step hopping. The numbers beside the peaks represent the energy barriers of hopping. The numbers below the curves represent the relative energy to a_e, which obtains the lowest energy.

We then conducted an in-depth analysis of I_i hopping events with nudged elastic band (NEB) calculations [110,189] to systematically study the sampled pathways without the assistance of *aF*. The $CsPbI_3$ crystal presents three perovskite structures within the temperature range of 200-600 K: cubic, tetragonal, and orthorhombic [190]. As shown in **Figure 3.3(a)**, there are three iodide sites in the orthorhombic structure: one is apical (*a*), which lies along the longest lattice vector; the other two are equatorial (*e* and *e′*) which lie in the equatorial plane. Moreover, it is insufficient to represent the states of I_i with the sites *a*, *e*, and *e′*, as I_i adopts two orientations at each site, resulting in six states. To facilitate clear discussions, we denote each state with a letter and a subscript. The letter indicates the specific iodine site, while the subscript denotes the normal direction of the orientation. For example, $a_{e'}$ represents I_i in the apical site oriented in the *e′* direction. Therefore, the

hopping between states with the same subscript is IP hopping, whereas those with the same letter are OP hopping. Similarly, for V_I, states can be denoted using subscripts corresponding to the site: V_a, V_e, and $V_{e'}$. Experimentally, $CsPbI_3$ adopts a cubic structure above 550 K, in which the three Pb-I axes are symmetrical to each other, leading to identical on-site energy of the six defect states and identical diffusion barriers in the three directions. At about 450 ~ 550 K, the cubic phase spontaneously breaks into the tetragonal phase via Pb-I octahedral tilting, resulting in the energy difference of apical direction and equatorial plane. Thus, the diffusion inside the equatorial plane is different from that between apical and equatorial sites. At lower temperatures, twisting of the axial Pb-I chain occurs. The two apical states are no longer degenerate, which makes the hopping barriers along all three Pb-I axes different.

We systematically calculated the energy barriers associated with IP and OP hopping between the identified sites, and the results are shown in **Figure 3.3(b)-(d)**. Due to the inherent symmetries, the e'_e and $e_{e'}$ states are identical, as well as e'_a and e_a. Conversely, the apical positions obtain two distinguishable states, in which a_e is the most stable state for I_i, and $a_{e'}$ exhibits a slightly higher energy, approximately 0.1 eV above a_e. The transition $e'_a \leftrightarrow e_a$ obtains a low IP hopping barrier of 0.16 eV. However, both states located in the equatorial planes possess the highest on-site energies. On the other hand, the $a_e \leftrightarrow a_{e'}$ is the most favorable reversible OP pathway. Generally, the equatorial plane is suitable for IP hopping owing to the lower and symmetric IP barrier of 0.16 eV, while the apical positions are suitable for OP hopping.

The energy differences among states provide valuable insights into the tendencies of defect movements. I_i is energetically favourable to stay at the apical sites but avoids lying in the equatorial planes. Once the defect transitions out of e'_a or e_a states, returning to these states becomes relatively challenging. In the meantime, the $a_e \leftrightarrow e'_e$ and $e_{e'} \leftrightarrow a_{e'}$ IP hopping pathways all include components moving along the apical direction. Consequently, the self-diffusion pattern of I_i stretches along the apical direction and contracts within the equatorial plane, indicating a propensity for easier movement along the apical direction. We have also verified whether there are other one-step transition pathways among the six states. As shown in **Figure 3.3(d)**, the only pathway we have found is the $a_e \leftrightarrow e_{e'}$ transition, which possesses an energy barrier 0.06 eV lower than the two-step pathway $a_e \leftrightarrow a_{e'} \leftrightarrow e_{e'}$. All the other transitions require a two or three-step pathway.

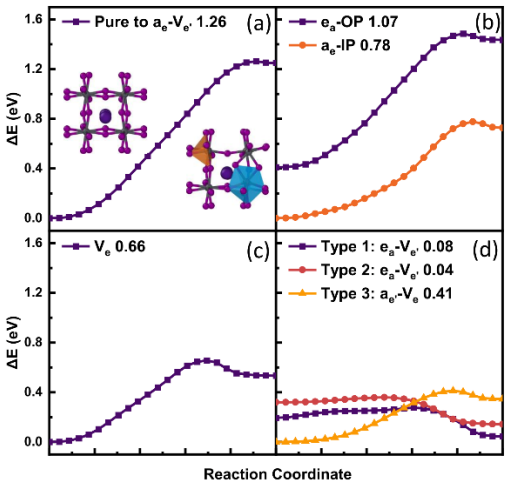

Figure 3.4 The energy barriers of aF creation (a) in an ideal crystal; (b) near an existing I_i; (c) near an existing V_I, and (d) near an existing I_i-V_I pair with the type 1, 2, and 3 conformations. The numbers following the labels are the energy barriers of creating aF.

We then focus on the analysis of *aF* formation. As mentioned above, the spontaneous formation of *aF* near existing defects can thereby enhance the diffusion and annihilation processes. Previous studies have reported a formation energy of 0.84 eV for *aF* in halide perovskites [191], a value lower than that typically observed in other materials (For example, 7.6 eV for O in SiO_2 and ~2 eV for O in TiO_2) [192,193]. However, this formation energy is still much higher than 10 k_BT (200–600 K) and cannot explain the observed formation of *aF* within a reasonable timescale in MD simulations. To address this, we performed NEB calculations to demonstrate that the presence of individual defects or defect pairs can facilitate the formation of an additional *aF*. The pathways with minimum energy barriers are shown in **Figure 3.4**. In addition, due to the existence of multiple cases of the initial defect state and the relative position between the defects, there could be multiple different paths for the same type of creation, and therefore we show these paths in **Figure 3.13**.

We first studied the energy landscapes for *aF* creation in the ideal crystal for reference, as shown in **Figure 3.4(a) and Figure 3.13(a)**. The creation of *aF* necessitates an iodide to rotate towards a nearby I site, which is highly unstable and may not obtain a stable final state. We identified three distinct pathways with different conformations, with the lowest barrier of 1.26 eV. However, the final energy states of $a_e - V_{e\prime}$ and $e'_e - V_a$ are very shallow,

indicating that a minor perturbation could lead to the annihilation of *aF*. Thus, external stimuli such as light illumination that can reformulate the energy landscape and provide enough energy for iodide to escape the unstable states are necessary [194]. The $e_a - V_{e\prime}$ conformation, although relatively stable, presents a higher barrier. The results indicate that it is difficult to form *aF* in an ideal $CsPbI_3$ crystal without external stimuli, despite a relatively small energy barrier. Even if it can be created, the displaced iodide is highly possible to revert and annihilate the defect pair.

Meanwhile, the presence of I_i in the lattice reduces the energy barriers for the creation of metastable *aF* through two pathways: OP or IP relative to the I_i lattice plane, as shown in **Figure 3.4(b)**. A typical *aF*-assisted OP transition follows the process:

$$e_a \rightarrow e_a + V_{e\prime} + a_e \tag{42}$$

This process normally requires 3 steps of ordinary IP and OP hopping as shown in **Figure 3.3**. Among the six different defect conformations, e_a exhibits the energy for the creation of *aF* OP most, while a_e obtains the hardest formation of *aF*. This difference is induced by the relative position of the original and created I_i. The final states in both cases are degenerated, which are all $a_e + V_{e\prime} + e_a$. The e_a itself possesses the highest on-site energy among all I_i states, requiring only the creation of the a_e state with the lowest on-site energy to complete the process. On the other hand, the opposite holds for a_e, which requires to generate the high energy e_a. Moreover, as shown in **Figure 3.13(a)**, the creation barrier of a_e from the ideal lattice is much smaller than that of e_a. This also increases the gap of their *aF* creation to some extent, although it is not in an ideal crystal in this case. In the V_I cases, the creation of additional *aF* exclusively occurs through the IP pathway. As shown in **Figure 3.4(c)** and **Figure 3.13(c)**, the associated barriers are generally smaller than those observed for I_i.

For each type of *aF*-assisted I_i-V_I annihilation, we systematically examined the energy barriers associated with creating additional *aF*, considering various relative positions of I_i and V_I. These conformations were labelled based on their respective sites. In these conformations, we have found six pathways for type 1, five pathways for type 2, and three pathways for type 3 to create relatively stable *aF* structures, as shown in **Figure 3.4(d) and Figure 3.13(d)-(f)**. This indicates that the I_i and V_I in shorter distances increase the number

of stable intermediate *aF* states. While the energy barriers vary largely for different structures, they all remain lower than the creation of *aF* in the ideal crystal. The relative energy for the initial I_i-V_I state primarily depends on the state of I_i, while the position of V_I only causes slight energy shifts. For the $a_e - V_{e\prime}$, we can see that the energy barrier of type 2 structure is lower than types 1 and 3, and type 1 exhibits the highest energy barrier among the three types.

To understand the observed pattern, we analyzed the transition structures, as illustrated in **Figure 3.14-Figure 3.16** of **Section 3.5**. One trivial result is that the largest displacement occurs on the iodides in between the I_i-V_I pair, forming the additional I_i in the final state. The iodide movement is most extensive in the first type of *aF* creation events, whereas it is minimal in the second type. Additionally, the transition requires distortion in the Pb-I octahedral structure. For the first type of *aF* creation, the bending of iodides opposite to the I_i-V_I pair significantly distorts the Pb-I octahedral structure, inducing a substantial rise in the overall energy of the structure. In contrast, the deformation of the octahedral structure is less severe for the second and third types, resulting in a lower energy penalty.

In these energy landscapes, we identified the smallest barriers in the $e_a - V_{e\prime}$ structure, ranging from 0.04 to 0.08 eV, which is remarkably low and comparable to the OP barrier (0.03 eV) of e_a and e'_a. It is also notable that the final states with an additional *aF* show higher stability compared to the initial state. This observation, coupled with the fact that the e_a or e'_a state obtains the highest energy among the I_i, is evident that e_a is quite unstable. In the $e_a - V_{e\prime}$ conformation, the additional I_i is created on the a_e state obtaining the lowest energy. This process can be represented as:

$$e_a + V_{e\prime} \rightarrow e_a + a_e + 2V_{e\prime} \qquad (43)$$

The e_a and newly created a_e can thus further annihilate the nearest $V_{e'}$ of them, leading to rapid and long-range defect annihilation.

In general, the energy barrier for creating an *aF* can be significantly reduced in the presence of surrounding defects. Without additional defects, *aF* is relatively harder to be stabilized in pure perovskite crystals. Notably, the lowest creation barrier is observed in the vicinity of I_i-V_I pairs. The reduced energy barrier can be attributed to the lattice strain induced by the presence of I_i and V_I. I_i results in the distortion of the surrounding lattice and provides

a negative charge distribution. As shown in **Figure 3.17**, the negatively charged Pb-I rhombus distorts the lattice and pushes nearby iodides away from their equilibrium positions. Conversely, the presence of positively charged V_I pulls iodides towards it, as shown in **Figure 3.18**. Under the influence of thermal fluctuations, it becomes possible for an iodide to temporarily escape its equilibrium position, becoming an interstitial and leaving a vacancy in its original position. This process occurs with a low energy barrier. Moreover, we have also found that the presence of V_I can decrease the hopping and transition of I_i states, as discussed in detail in **Section 3.5** and **Figure 3.19**. These phenomena further elucidate the self-healing behaviour observed in halide perovskite materials [195]. The *aF*-assisted process not only increases the finite temperature diffusion length of iodide but also extends the annihilation radius of iodide-related defects, especially at higher temperatures. Firstly, the I_i and V_I can diffuse fast in the lattice, which increases the possibility that they meet each other per unit time. Once the free I_i has moved to a position within three unit cells of a V_I, the *aF*-assisted annihilation is possible to occur to complete one "self-healing" event.

In the context of high-temperature behaviours, the NEB calculations discussed above provide valuable insights but are limited in capturing the pathways entirely, especially across different perovskite phases. Temperature can influence the free energy of the defect configuration, altering the number of intermediate states and pathways in the defect hopping process due to varying entropy changes among states. To explore high-temperature *aF*-assisted I_i hopping phenomena, we employed well-tempered metadynamics MD simulations. We used three-dimensional collective variables (CVs) constructed with environmental similarity [196] of different I_i states. The positive or negative value of each CV denotes two orientation states on one type of site (*a*, *e*, or *e'*). While these CVs may not perfectly encapsulate all transition states, especially in cases where the three on-site OP rotations overlap near the origin point of the CV space, they are sufficient for studying *aF*-assisted processes.

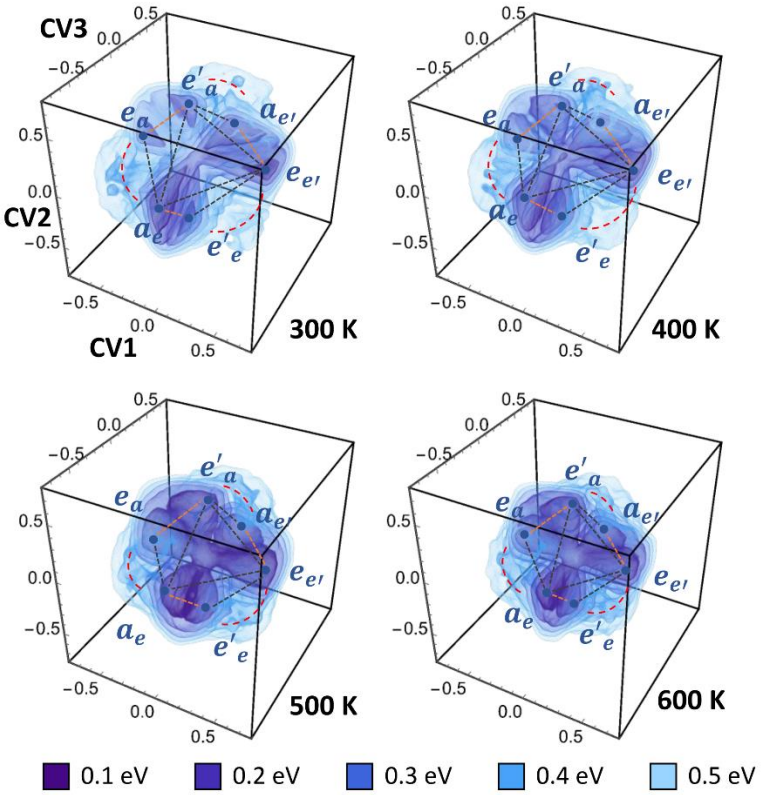

Figure 3.5 The free energy landscape in a three-dimensional CV space of I_i hopping under 300, 400, 500, and 600 K. The points and symbols denote the equilibrium positions of defect states. These points in the space form an octahedron in CV space, as denoted with grey and orange dashed lines, which become symmetrical under higher temperatures. The orange dashed lines denote the IP hopping process between states, and the red dashed lines denote the aF-assisted process between states. The isosurfaces in the figure are with the levels of 0.1, 0.2, 0.3, 0.4, and 0.5 eV, from dark blue to light blue, respectively.

The free energy landscapes of 300-600 K are drawn in **Figure 3.5**. The shapes of the whole landscapes are similar to three-bladed propellers. Within each blade, the transitions signify IP hopping on specific planes, while ordinary OP rotations of I_i occur near the origin of the CV space. Notably, the three blades show clear differences at 300 K and 400 K, where one blade appears smaller than the other two. This blade corresponds to the transition lying on the equatorial planes, with the free energies of e_a and e'_a being the highest among the six states—a trend consistent with the NEB results shown in Figure 3(c). This difference can be attributed to the gradual transition from pseudo-cubic phases of perovskite to the cubic phase as the temperature rises, resulting in a symmetric lattice.

The *aF*-assisted OP processes are marked with red dashed lines in **Figure 3.5**. These transitions occur between two adjacent blades, which means only three pathways exist: $e_a \leftrightarrow a_e$, $e'_e \leftrightarrow e_{e'}$ and $a_{e'} \leftrightarrow e'_a$. No other transitions are sampled within a suitable energy scale due to the structure of I_i. We summarized the *aF* creation barriers for each type of defect in **Table 3.2** of **Section 3.5**. The three pathways for *aF* creation have different free energy barriers in 300 and 400 K. The most probable pathways are the $e_a \rightarrow a_e$ and $e'_a \rightarrow a_{e'}$ with a barrier of about 0.25 eV. The transition between $e_{e'}$ and e'_e obtains the highest free energy of about 0.4 eV. On the other hand, similar to the IP transitions, higher temperatures pull the barriers of all the transitions into alignment and decrease these barriers to about 0.30 eV. As a result, these barriers come closer to the value of 10 k_BT, particularly at higher temperatures, signifying that a substantial population could potentially exist within crystal structures.

The above discussions are all about the all-inorganic halide perovskite $CsPbI_3$, which obtains a simple structure. Meanwhile, we could infer similar *aF* phenomena in organometal halide perovskite such as $MAPbI_3$ and $FAPbI_3$. Owing to the existence of asymmetric organic cations, the energy landscape could be more complex and different from the case in $CsPbI_3$. During the creation of *aF*, the motion of iodide will lead to the lattice relaxation assisted by cation rotation. The hydrogen interaction between organic cations and halide anions differs in different relative orientations [197], which may lead to elevated or reduced energy barriers for some *aF* creation processes.

Moreover, *aF*-assisted phenomena require that there be originally existing point defects in the lattice. Usually, the fabrication process, thermal fluctuation and external stimuli such as light irradiation can indeed stimulate the creation of aF [198–200]. However, the creation barrier is still not considered low, and in the meantime, after the aF formation, I_i and V_I can easily annihilate directly with each other rather than separate away. The existence of grain boundaries and surfaces, nevertheless, allows defects to arise more commonly. Previous research has shown that the surface and grain boundary of halide perovskite could reduce the Frenkel formation energy of I_i-V_I pair, especially in the case of PbI-terminated surface can reduce that to 0.03 eV [188,201]. In addition, the grain boundaries have attractive forces towards I_i and V_I [202]. I_i and V_I will tend to migrate towards PbI- and MAI-terminated boundaries, respectively, which assist the separation and accumulation of defects in the system [203]. These findings also suggest that the *aF*-assisted mechanism

could have more important roles at places such as grain boundaries, and then cause degradation of the bulk materials [204]. Further studies are required to obtain detailed conclusions.

3.4. Conclusion

In summary, we have thoroughly investigated the influence of anti-Frenkel disorders in the microscopic diffusion kinetics of iodide defects. We observed that the negatively charged iodide interstitial is preferred to be on apical sites in the perovskite and enables stretching of self-diffusion pattern along the apical direction. Moreover, we demonstrated that the presence of point defects can facilitate the creation of additional anti-Frenkel disorders, elongating the interaction range of defect pairs and the defect diffusion pathways. Under higher temperatures, the free energy barrier of the hopping pathway assisted by the anti-Frenkel disorder can be further decreased due to the entropy term, which effectively enhances the self-healing mechanisms of $CsPbI_3$. Our work sheds light on the mechanics of defect diffusion and shows that halide perovskite is an ideal material for studying the defect creation and annihilation phenomenon in materials science.

3.5. Supplementary Section

3.5.1. Theoretical methods

3.5.1.1. Density functional theory calculations

DFT calculations of $CsPbI_3$ were performed with Vienna Ab-initio Simulation Package (VASP) [205] with projector-augmented wave (PAW) pseudopotentials [206] and Perdew-Burker-Ernzerhof parametrization [122]. A plane-wave energy cutoff of 500 eV was used. The DFT-D3 method with the Becke-Johnson damping function [207] was used for describing van der Waals dispersion interaction. Phase transitions and lattice constants can be reproduced more accurately by using this setup [208]. To be noticed, the HSE06 functional [209] with spin-orbit coupling (SOC) was not used for geometry optimization

and energy calculations, which is considered to be the most accurate ab initio functional for electronic structure calculation. This is due to the extreme computational consumption and the large dataset we need to construct. Moreover, it is previously reported that SOC has little effect on force calculation [200] which is important in our study. Monkhorst-Pack scheme k-points mesh with a k-spacing of 0.2 was used. *Ab initio* molecular dynamics simulations were performed with the same setups as the single point calculations and a 2 fs timestep was chosen. The canonical (NVT) ensemble was used and controlled with a Nosé-Hoover thermostat [210,211]. $CsPbI_3$ instead of the more popular organometal halide perovskites is chosen for this study to exclude the influence of asymmetry from organic cations on halide defect motion [212].

3.5.1.2. Dataset construction

Table 3.1 The correspondence of system labels and conformations of the datasets.

System	**Type**	**Supercell**	**Defects**
000	α	$2 \times 2 \times 2$	
001	α	$2 \times 2 \times 2$	V_{Cs}-V_I
002	γ	$\sqrt{2} \times \sqrt{2} \times 1$	
003	δ	$2 \times 1 \times 1$	
004	α	$3 \times 3 \times 3$	I_i-Cs_i
005	δ	$4 \times 2 \times 1$	I_i-V_I
007	PbI_2	$2 \times 2 \times 1$	
009	α	$4 \times 4 \times 4$	
010	γ	$2 \times 2 \times 2$	
011	δ	$4 \times 2 \times 1$	
012	γ	$2 \times 2 \times 2$	V_{Pb}-Pb_i V_{Cs}-Cs_i
015	γ	$\sqrt{2} \times \sqrt{2} \times 1$	I_i-V_I
016	α	$3 \times 3 \times 3$	I_i-V_I
017	α	$2 \times 2 \times 2$	V_{Pb}-Pb_i
018	α	$3 \times 3 \times 2$	I_i-V_I
019	α	$2 \times 2 \times 2$	V_{Cs}-Cs_i

All the datasets used here were constructed with DP-GEN [162]. The initial datasets containing 15000 data points were extracted from several 1-ps-long AIMD trajectories at 300 K. The 10 starting structures were obtained by randomly perturbing the box size (in

±0.02) and atomic positions (in ±0.02 Å) of α- and δ-$CsPbI_3$. The remained datasets were constructed using the concurrent learning strategy provided by DP-GEN. We have explored the conformations of perovskite phases and non-perovskite phases, and the lattices with defects, as shown in **Table 3.1**. The conformations with a max model deviation of force between 0.03 and 0.2 were chosen to be the candidates for DFT single-point energy calculations. During the exploration of DP-GEN, model deviation MD runs of the first 7 iterations were under NVT thermostat within the temperature range of 200–650 K, and all the other iterations were under isothermal-isobaric (NPT) thermostat with constant pressure at 1 standard atmosphere.

3.5.1.3. Deep learning potential

The deep learning potential we used was trained with DeepMD-kit [159], which extracts the local environments of atoms and then maps them to their energies and forces. The utilization of deep learning potential enables us to extend our simulations beyond the limitations of supercell size and timescale inherent in density functional theory (DFT) calculations [205]. The DFT-calculated datasets from DP-GEN can be directly used for training DP models. Our model is equipped with a three-hidden-layer embedding net with (25, 50, 100) nodes and a fitting net with (240, 240, 240) nodes. The *se_e2_a* descriptor was used with both angular and radial information of atomic configurations considered. A cutoff radius of 8.5 Å smoothing start at 1.0 Å was used. The model was trained for 2 000 000 steps with learning rate decays from 10^{-3} to 10^{-8} and the energy, force, and virial prefactors in the loss function varied from 0.02 to 1, from 1000 to 1, and from 0.01 to 0.1 during the training, respectively. The trained model then went through a compressed training procedure to accelerate computation [213].

3.5.1.4. Nudged elastic band calculations

The NEB calculations were performed with LAMMPS and DeepMD-kit. 8×8×8 supercells were used to avoid the periodic boundary effects [214]. A total of 24 images were performed in one calculation. We first optimize the structure and volume of the ideal $CsPbI_3$

structure with the conjugate gradient algorithm, with energy and force stopping tolerance of 10^{-40} and 10^{-40} eV/ Å. Then the defects were set inside the systems and another minimization process with the cell shape fixed. The NEB was performed using the "quickmin" style with a force-stopping tolerance of 0.001 eV/ Å.

3.5.1.5. Molecular dynamics simulations

The molecular dynamics simulations of this study were carried out using LAMMPS code patched with DeepMD interface and Plumed [215]. All the simulations were done with periodic boundary conditions and the NPT ensemble. The migrations of iodide-related point defects were studied using 8×8×8 cubic supercells (around 2560 atoms), for which the defect densities are approximately 7.5×10^{18} cm^{-3}. The Nose-Hoover thermostat [210] and Parrinello-Rahman barostat [148] were adopted to control the temperature and pressure of systems, respectively. The simulations were done at 200–600 K in 1 standard atmosphere, which covers the temperature range for fabrication and application of $CsPbI_3$ perovskite materials. The timestep of the simulation was set to 2 fs. The data for single point defect jumping events were collected from simulations in 4 ns and the first 20 ps in each trajectory were dropped to guarantee an equilibrium state.

The well-tempered metadynamics MD simulations were performed with CVs constructed with environmental similarity multicolvar implemented in Plumed 2.8. The broadening parameters of these CVs were set to 0.5. The maximum values of environmental similarity of each defect state were extracted with a beta of 0.05. We then constructed the final CVs by subtracting the values of the two states on each site. We have chosen 0.04 for the width of Gaussian hills for all three CVs. Since the free energy landscapes change with the temperature, we used different Gaussian heights and bias factors for different temperatures. The Gaussian heights for 300 and 400 K were set as 0.041 eV, and those for 500 and 600 K were 0.031 eV. The bias factors for 300, 400, and 500 K were 3, and that for 600 K was 2. The Gaussian hill addition frequency is set to 500 timesteps.

3.5.2. Accuracy of the model

Various tests were performed to verify the accuracy of DPMD, as shown in **Figure 3.6**-**Figure 3.12**. In **Figure 3.12**, we have calculated 2500 frames from an MD trajectory containing the defect motion and annihilation events in a 160-atom system. The absolute errors of DPMD prediction of system energy are shown in **Figure 3.12**(a), with hopping and annihilation events marked with red points. As shown in **Figure 3.12**(b), most of the errors are below 0.045 eV which is high and can accurately represent the DFT results. The errors in transition states are generally lower than 0.1 eV. The accuracy of a defect-free system may be slightly higher compared to the system with defects.

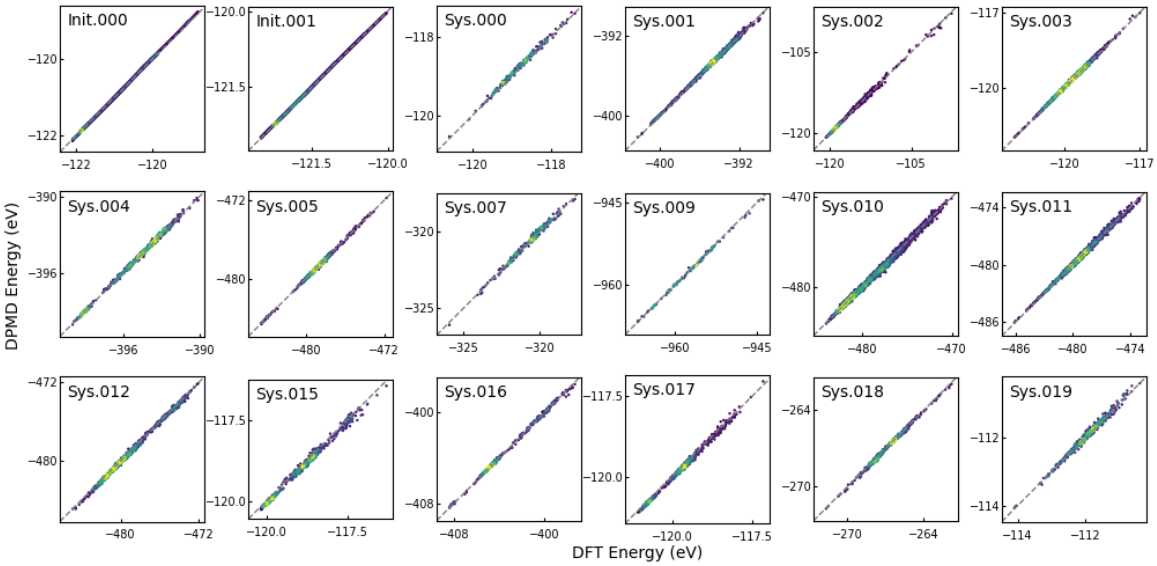

Figure 3.6 Comparison of energy values in different systems. The horizontal and vertical axes denote the value obtained from ab initio calculations and DPMD simulations, respectively.

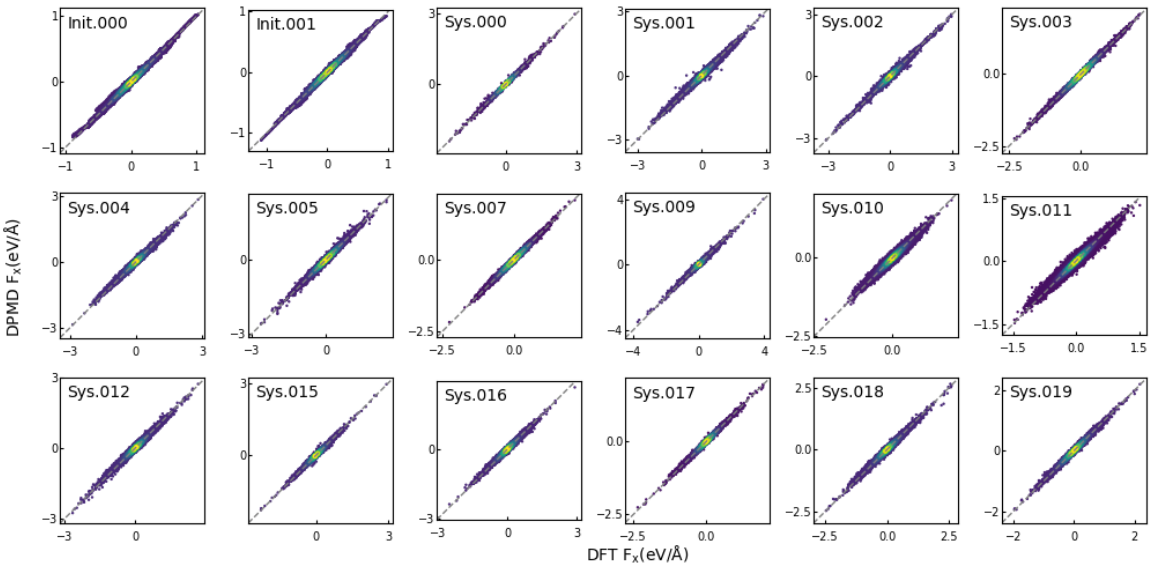

Figure 3.7 Comparison of the x-axis component of forces in different systems.

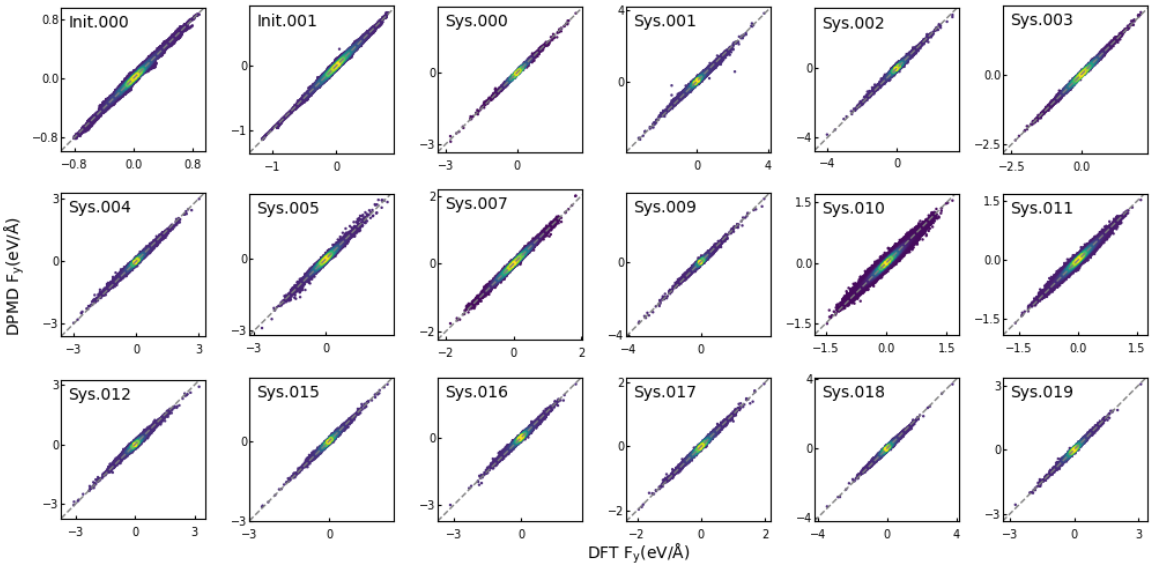

Figure 3.8 Comparison of the y-axis component of forces in different systems.

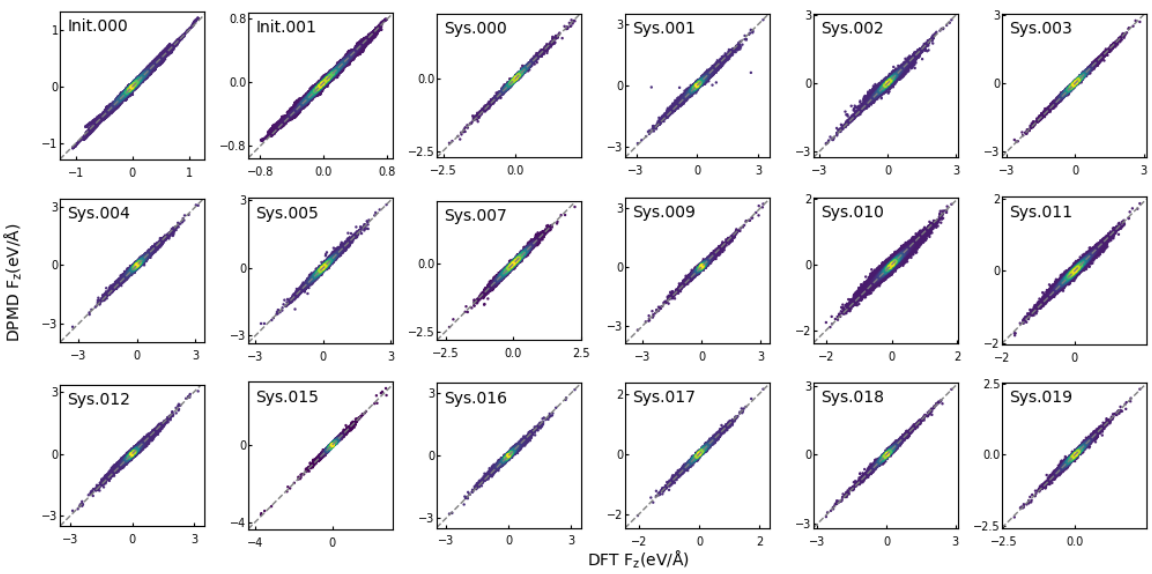

Figure 3.9 Comparison of z-axis components of forces in different systems.

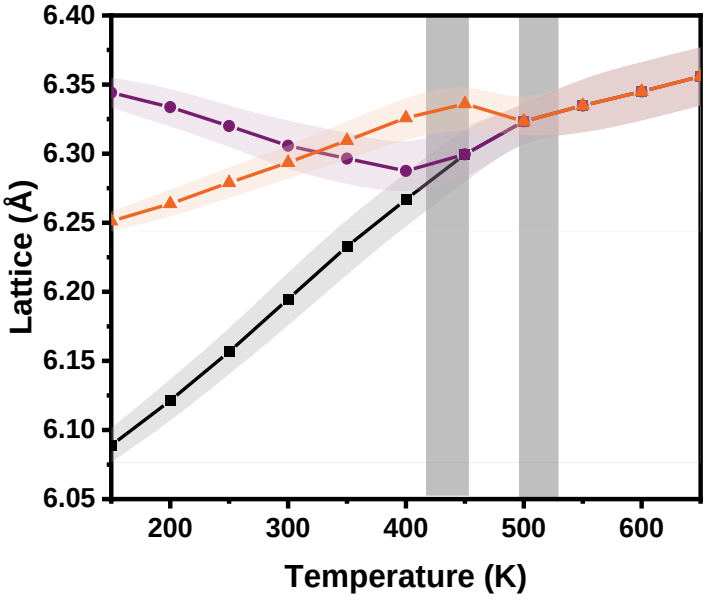

Figure 3.10 The pseudo-cubic lattice parameters versus temperatures extracted from DPMD simulations. The shadowed areas near curves denote the fluctuations of the lattice parameters. The grey bands denote the range of phase transition.

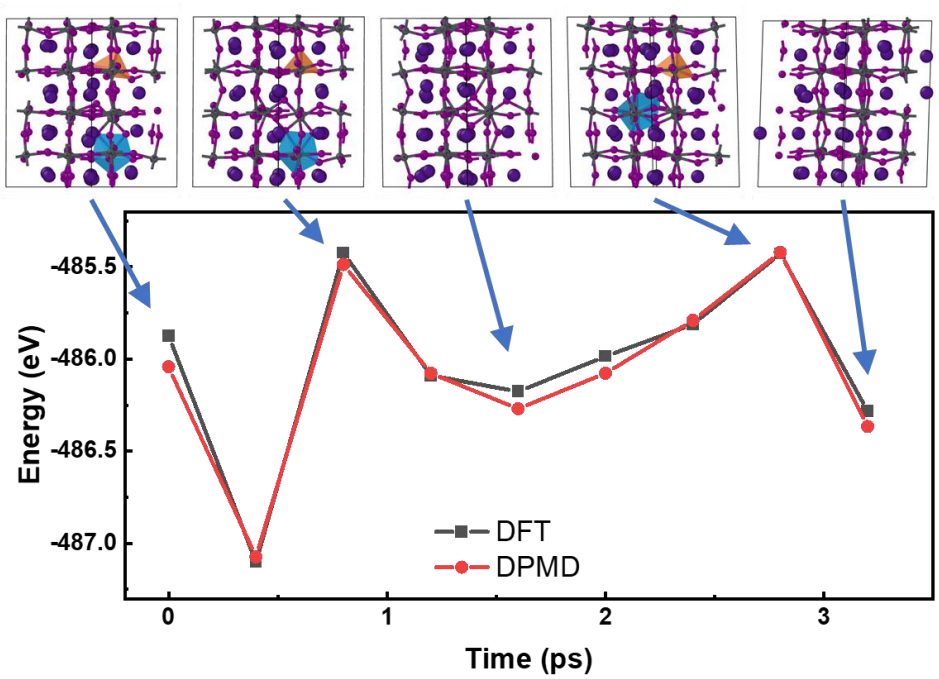

Figure 3.11 Calculated DFT and DPMD energies for an annihilation event sampled from MD simulations. The top pictures denote key structures during annihilation, as pointed out with arrows.

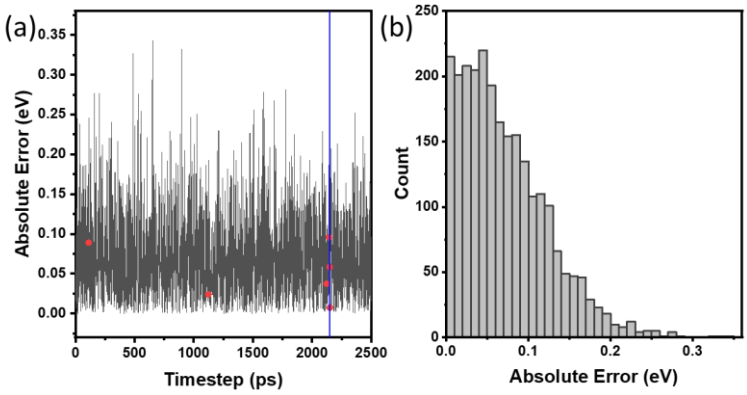

Figure 3.12 Examine the accuracy of the DP model predicted energy. (a) Absolute error curve of DPMD energies for a trajectory contains defects hopping and annihilation events, marked with red dots. The blue line separates the frames with defects (left) and without defects (right). (b) Histogram for the absolute error of all the frames.

3.5.3. Other calculation results

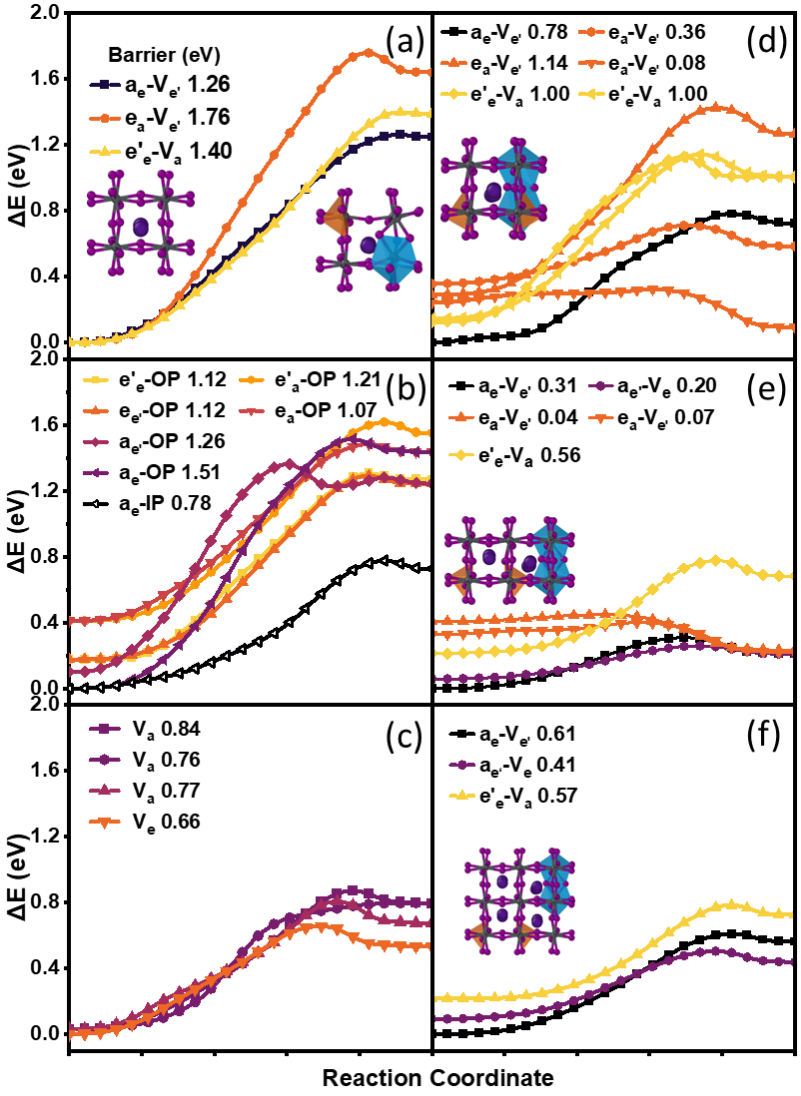

Figure 3.13 Additional trajectories for each type of *aF* creation, which also includes the lowest energy landscape shown in Figure 4. The energy barriers of *aF* creation (a) in an ideal crystal; (b) near an existing I_i; (c) near an existing V_i. (d), (e) and (f) are the creation barriers of *aF* near an existing I_i-V_I pair with the type 1, 2, and 3 conformations, respectively. The numbers following the labels are the energy barriers of creating *aF*.

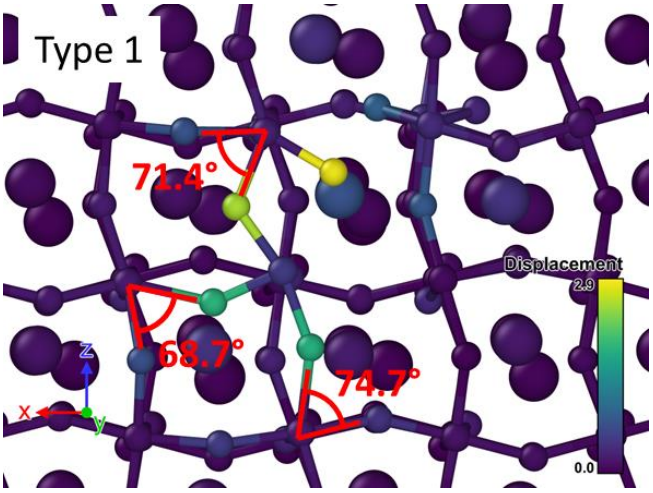

Figure 3.14 The displacement magnitude and typical I-Pb-I angles for the transition state of the first type *aF* creation.

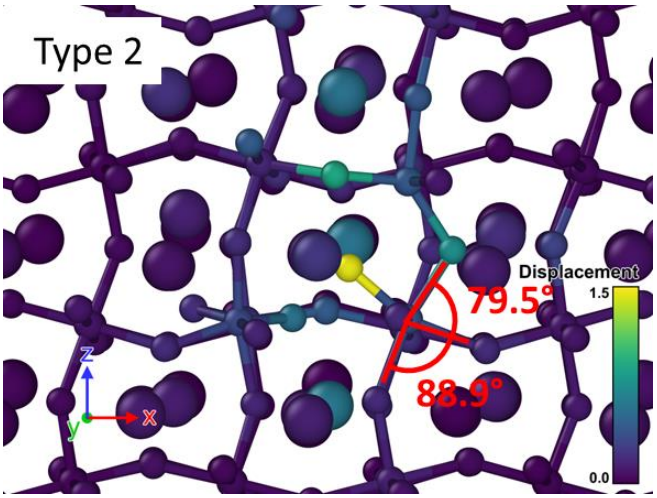

Figure 3.15 The displacement magnitude and typical I-Pb-I angles for the transition state of the second type *aF* creation.

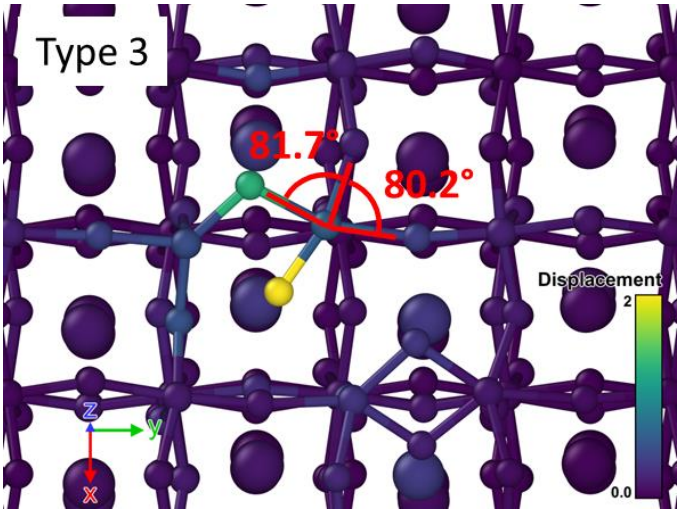

Figure 3.16 The displacement magnitude and typical I-Pb-I angles for the transition state of the third type *aF* creation.

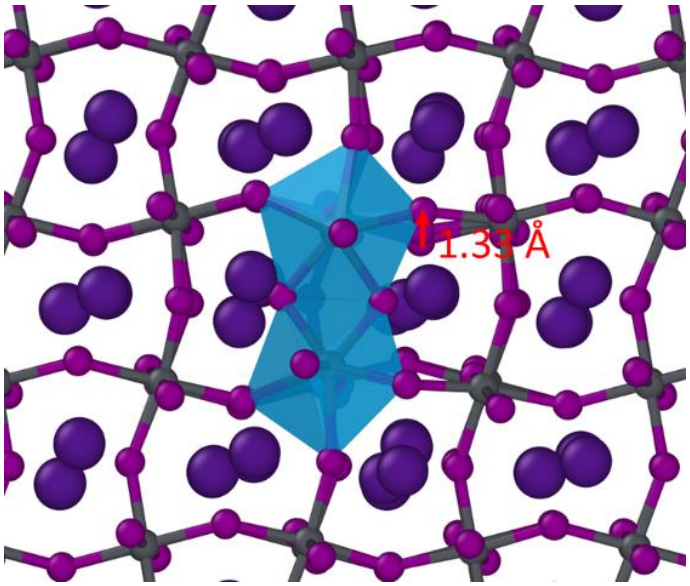

Figure 3.17 The displacement of the iodide resulted from an adjacent I_i.

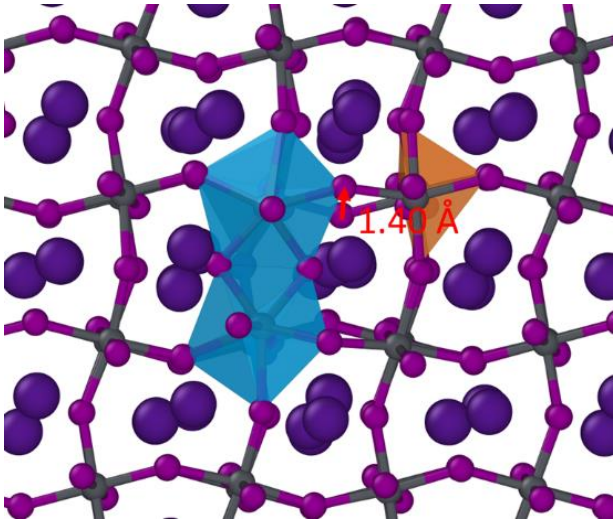

Figure 3.18 The displacement of the iodide resulted from an adjacent I_i and a V_I.

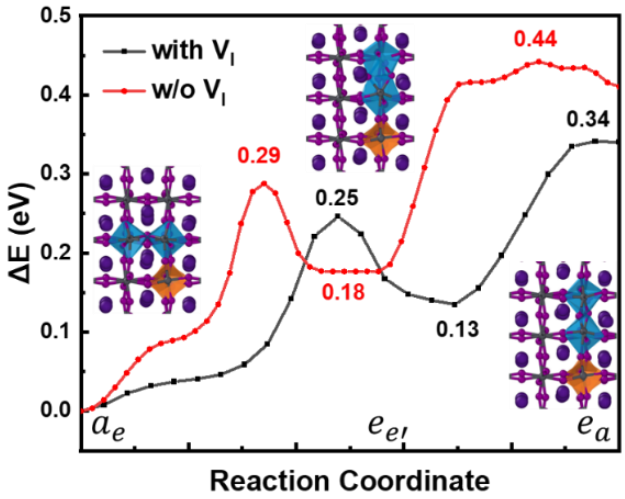

Figure 3.19 The NEB energy landscape for a_e-e_a transition with and without the influence of $V_{e'}$. The insets are the structures of the a_e, $e_{e'}$ and e_a states, respectively.

The influence of V_I on I_i hopping. We chose $a_e - V_{e\prime}$ state as the initial state to be investigated since it obtains the lowest energy as shown in **Figure 3.13**(d). In addition, the final e_a state will be aligned in the same plane as the $V_{e\prime}$, which increases the possibility of fast defect annihilation with only the IP hopping process. The transition obtains two substeps. First, a_e jumps into the $e_{e\prime}$ state, then to the e_a state. As shown in **Figure 3.19**,

the transition barriers of the two hopping processes decrease by about 0.04 and 0.10 eV compared to those of I_i alone, respectively. In addition, the relative energy of $e_{e'}$ and e_a states also decrease by 0.04 and 0.07 eV, respectively. The result shows that the presence of V_I could decrease the transition barriers of I_i hopping, thus in turn promoting the motion of I_i.

Table 3.2 The *aF* creation barrier of different I_i states under different temperatures.

***aF* barrier (eV)**	300 K	400 K	500 K	600 K
$\mathbf{a_e}$	0.39	0.40	0.35	0.30
$\mathbf{e'_e}$	0.40	0.39	0.39	0.31
$\mathbf{e_{e'}}$	0.43	0.41	0.35	0.33
$\mathbf{a_{e'}}$	0.43	0.39	0.36	0.27
$\mathbf{e'_a}$	0.26	0.29	0.37	0.28
$\mathbf{e_a}$	0.25	0.31	0.32	0.32

Chapter 4.

Anisotropic Motion of Iodide in $CsPbI_3$: The Role and Influence of Planar Confinement

4.1. Abstract

The unique soft lattice and octahedral structure of halide perovskites facilitate distinct defect diffusion behaviours that are crucial for their performance in halide perovskite materials and devices. In this study, we employ deep-learning-enhanced molecular dynamics to investigate iodide self-diffusion in $CsPbI_3$. Our results demonstrate that the planar confinement on interstitial iodides will lead to different diffusion kinetics from that of vacancies. At lower temperatures, the motion of interstitials is further restricted, preventing them from lying inside the equatorial plane. Findings derived from kinetic Monte Carlo simulations elucidate that while planar confinement acts to expedite the annihilation process between proximate vacancies and interstitials, it paradoxically serves to prolong the overall lifetime of defects over an extended period. Moreover, the results reveal that the ⟨100⟩ family is the slowest movement direction of iodide interstitials in the long run, while ⟨110⟩ representing the fastest. This planar-confined behaviour not only influences defect dynamics but also offers new avenues for enhancing the performance and longevity of halide perovskite-based electronics by guiding defect management strategies.

4.2. Introduction

In the last decade, halide perovskites have attracted significant research interest due to their unique optoelectronic properties, making them promising for photovoltaic and optoelectronic devices [216–218]. Notably, organometal halide perovskites, such as

$MAPbI_3$ (where MA = methylammonium) and $FAPbI_3$ (where FA = formamidinium), have achieved power conversion efficiencies exceeding 25% when incorporated as the photo-active layer in solar cells [219,220]. However, the degradation of organic compounds in these materials poses a significant challenge, leading to reduced efficiency and stability [221,222]. To overcome these issues, researchers have widely studied all-inorganic halide perovskites such as $CsPbI_3$ that intrinsically exhibit enhanced thermal stability and resistance to degradation [223,224]. However, the defects in perovskites tend to undergo ion migration, especially under heat and illumination, causing unwanted phase transitions and phase segregation under certain conditions, as well as hysteresis issues, which remain a critical hurdle [187,225–228]. Therefore, it is vital to get a deep understanding of the mechanistic origins of ion migration within these materials.

To date, research has been focused on the diffusion of halide vacancies and interstitials, particularly iodide vacancies, which are identified as the primary halogen-related defects due to their low formation enthalpy in $CsPbI_3$ [167,176,178,229]. Halide interstitials are also important for studying since they could also contribute to the performance instability [230,231], and could be related to super-fast diffusion induced by planar confinement [175]. While previous nudged elastic band (NEB) [110] calculations have provided insights into the transition states and energy barriers of ion diffusion under zero Kelvin [174,232–234], this DFT-based approach often fails to capture the complexity of all potential pathways and cannot provide a comprehensive understanding of defect diffusion at finite temperature, where thermal fluctuation introduces additional diffusion pathways. This results in a diverse collection of different defect states and different diffusion pathways, particularly for the interstitial iodides which obtain a more complex configuration. Moreover, current classic molecular dynamics (MD) simulations have predominantly focused on higher temperature ranges [184,208,229], and neglected the anisotropic diffusion pathway that emerges at room temperature, thereby highlighting the need for quantitative analyses of defect diffusion at finite temperatures.

In this study, we implemented deep learning molecular dynamics (DPMD) and kinetic Monte Carlo (kMC) [235] simulations to investigate the self-diffusion behaviours of iodide interstitials and vacancies in $CsPbI_3$. Our results reveal that iodide interstitials (I_i) in $CsPbI_3$ exhibit a quasi-two-dimensional (2D) diffusion behaviour, which facilitates annihilation of the near-range I_i and iodide vacancy (V_I) pairs but increases the pair lifetime by about 20%

for distant defect pairs, particularly under lower temperatures. Further, we identified the speed of self-diffusion of defects along the common crystal direction and the energy preference of defects on different lattice sites. Such a preference towards directions and sites can significantly restrict the movement of I_i and therefore lead to performance differences among different orientations. This study provides a comprehensive understanding of the anisotropic motion of iodide interstitials in $CsPbI_3$ and offers new perspectives for the design of more stable and efficient perovskite-based devices.

4.3. Results and Discussion

4.3.1. Trajectory analysis

We first studied the influence of temperature on the self-diffusion of point defects. **Figure 4.1** and **Figure 4.8** illustrate the trajectories of I_i and V_I. In these figures, the 3D trajectories are projected to the [100] lattice plane according to the cubic structure. Specifically, the z-axis is aligning with the apical direction in low temperature phases. In **Figure 4.1**, there exist a number of vertically and horizontally distributed trajectories, which represents the movement inside the xz and yz planes, respectively. Except for these trajectories, the areas covered by trajectories with same colours represent the defect movement inside the same xz planes, such as the trajectory in the x range -20 ~ -80 Å in 300 K, etc. Therefore, it shows that the movement of Ii is always inside a specific xy, xz, or yz plane, demonstrating predominantly planar-confined motion within a specific cubic lattice plane in the {100} family, with occasional transitions to other planes, similar to observations in $MAPbI_3$ [175]. This contrasts with the trajectories of V_I shown in **Figure 4.8**, which exhibit more freedom of movement in all directions along the Pb-I octahedral structure. This indicates that I_i has two primary motion pathways: in-plane (IP) hopping within a specific crystal plane, and out-of-plane (OP) hopping to another plane. As shown in **Figure 4.1**, the OP hopping is more prevalent at high temperatures: the I_i trajectories transform to other planes more often at 500-600 K than at 300-400 K, suggesting that lower temperatures favour planar-confined motion. Notably, trajectories from 300 and 400 K seem to suffer more confinement, in which defects tend to move within the planes that include the apical axis (the XZ and YZ planes in **Figure 4.1**) but rarely within the equatorial planes (the XY plane). This

phenomenon is attributed to the symmetry breaking at low temperatures, which is discussed later in **Section 2.5**.

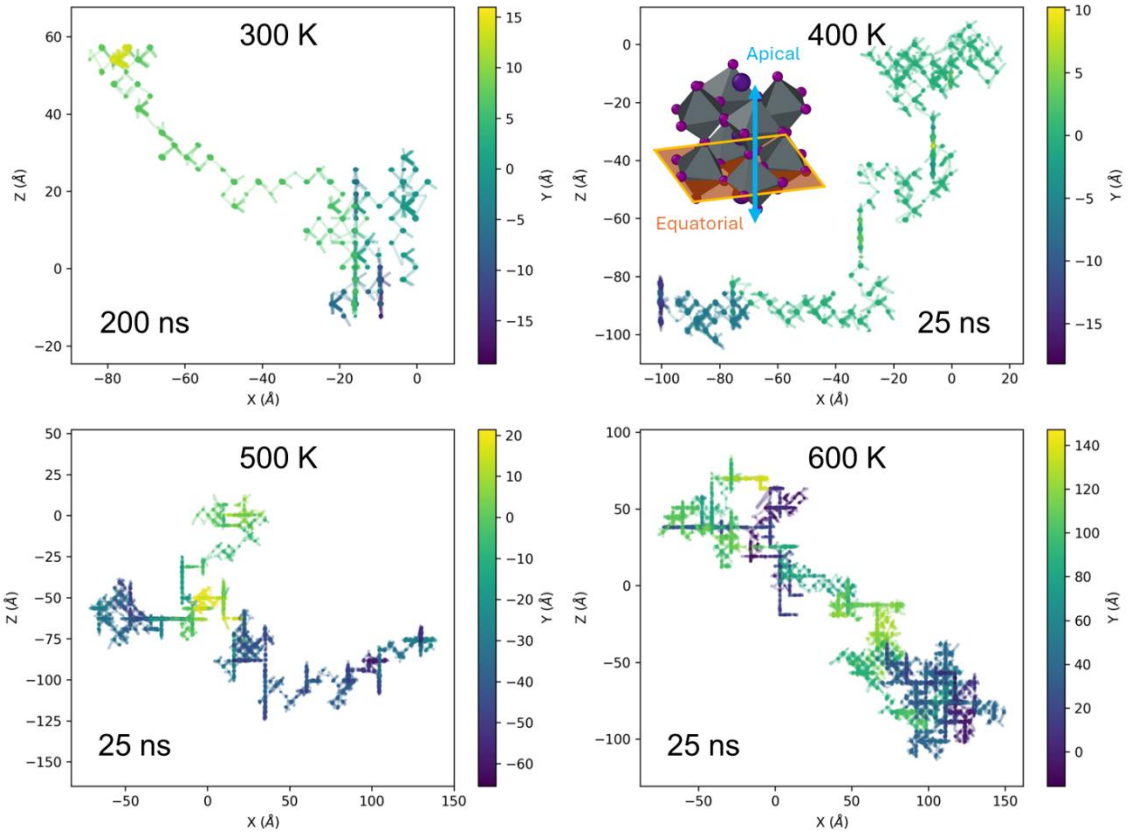

Figure 4.1 The unwrapped projected trajectories of I_i at 300, 400, 500, and 600 K. The simulation timescale is marked in each figure. The colour map represents the depth of points in trajectories inside the figure. In the figure, the z-axis represents the apical direction, while x and y represent the two equatorial directions. The schematic inserted shows the position of the apical direction and equatorial plane with the blue line and orange plane, respectively.

To compare our results with previous studies [236], we have estimated the self-diffusion activation energy E_a^D of I_i and V_I using the Arrhenius plot, as listed in **Table 4.1** and drawn in **Figure 4.2**. Different from previous MD studies that determined the 3D diffusion coefficients at higher temperatures (in cubic phase, denoted as α), our analysis further includes orthorhombic phase (β) and tetragonal phase (γ) to assess phase influences on defect diffusion. We have also listed the results from other studies in **Table 4.1**. Generally,

the activation energy at the α phase varied among studies, and our results are roughly in the middle. **Figure 4.2**(a) reveals that I_i diffuses approximately 1.5 orders of magnitude faster than V_I, especially at lower temperatures, as also evident from the difference in varying area sizes covered by the trajectories in **Figure 4.1** and **Figure 4.8**. Previous experimental analysis of halide diffusion within $MAPbI_3$ shows that the diffusion coefficient of halide species is about 10^{-13} to 10^{-12} cm^2/s at 75~125℃ [177], which is similar to the current results. This suggests that I_i possibly dominates the diffusion under such a temperature range.

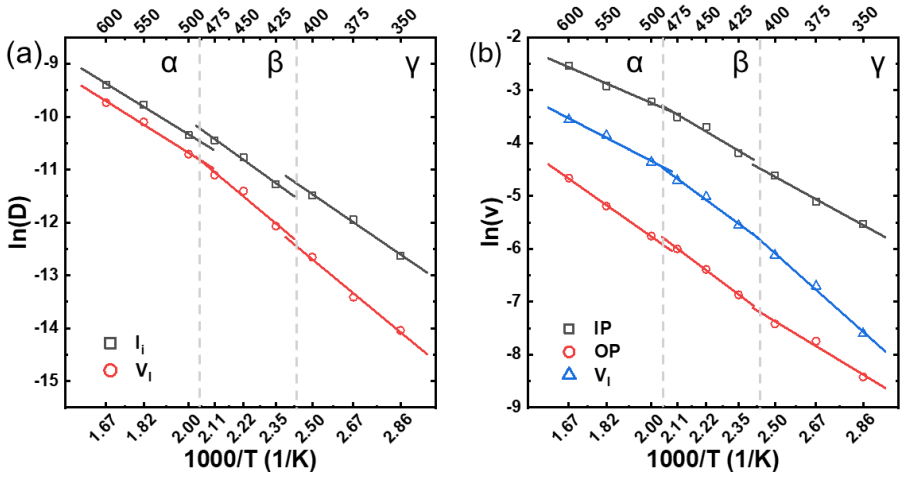

Figure 4.2 (a) The Arrhenius plots for I_i and V_I defect diffusion. The black squares and red circles are the diffusion coefficients derived from the mean square displacement of the iodide species in the systems containing a single I_i and V_I, respectively. The lines show the fitting results in each phase. The temperature zones where different phases are located are roughly divided by dashed lines. (b) The Arrhenius plots for I_i IP, I_i OP, and V_I hopping events in each phase.

Overall, the E_a^D for both I_i and V_I is around 0.25 eV, aligning with a recent MD study [184] and also similar to the experimental and computational studies for the organometal halide perovskites (in the range 0.1-0.3 eV) [171,175,237–239]. However, the values are slightly lower than several NEB calculation results for the organometal ones (in the range of 0.3-0.5 eV) [233,234,240]. This discrepancy may result from the influence of lower symmetry phases. The E_a^D values for α phases are the lowest among the three perovskite phases, with β and γ phases showing similar E_a^D values. Lower temperatures lead to lattice symmetry

breaking, causing the halide sites to obtain various on-site energy. Therefore, the defects could pin to low-energy sites with a higher energy barrier, thereby increasing the E_a^D. Moreover, the organic cations in $MAPbI_3$ and $FAPbI_3$ are polarized and can rotate in space under high temperatures, while they could be largely frozen when temperature decreases. The cations near defects could relax to a lower energy state via rotation, further increasing the energy barriers.

Table 4.1 Self-diffusion activation energy E_a^D and pre-exponential factor D_0 fit from self-diffusion coefficients.

	Phase	D_0 (10^{-6} cm^2/s)	E_a^D (eV)	Ref. E_a^D (eV)
I_i	α	1.03 ± 0.34	0.25 ± 0.02	0.28 (MD) [208], 0.20 (MD) [184]
	β	3.56 ± 2.69	0.29 ± 0.03	
	γ	3.08 ± 2.00	0.28 ± 0.02	0.16~0.39 (NEB) [1]
V_I	α	0.80 ± 0.40	0.25 ± 0.02	0.19 (MD) [208], 0.34 (MD) [184]
	β	6.04 ± 9.72	0.34 ± 0.06	
	γ	4.80 ± 4.67	0.33 ± 0.03	0.58 (NEB) [241]

Table 4.2 Hopping activation energy E_a^ν and attempt frequency ν_0 for I_i hopping.

	Phase	ν_0 (THz)	E_a^ν (eV)
I_i IP	α	2.23 ± 1.05	0.17 ± 0.02
	β	10.50 ± 15.03	0.24 ± 0.06
	γ	5.89 ± 3.16	0.22 ± 0.02
I_i OP	α	2.20 ± 0.44	0.28 ± 0.01
	β	4.06 ± 0.88	0.30 ± 0.01
	γ	0.74 ± 0.91	0.24 ± 0.04
V_I	α	1.64 ± 0.70	0.21 ± 0.02
	β	12.73 ± 12.01	0.30 ± 0.04
	γ	75.4 ± 69.97	0.36 ± 0.03

While E_a^D offers a general overview of the diffusion barrier, it is unable to distinguish between different types of hopping events. Therefore, we calculated the activation energies E_a^ν from the hopping frequency ν (ps^{-1}) from MD trajectories, i.e. the hopping events per unit time, for a better understanding of IP and OP hopping of I_i, as detailed in **Figure 4.2(b)** and **Table 4.2**. In general, E_a^D and E_a^ν exhibit similar trends, despite some differences in values. This discrepancy arises because E_a^D is derived from the mean square displacement (MSD) of all iodine ions, whereas E_a^ν is based on the frequency of the diffusive defect. The E_a^ν for I_i IP and V_I hopping have the same trend with E_a^D, with higher barriers under lower temperatures. Interestingly, one anomalous point is that the OP hopping of I_i obtains a lower barrier in the low-temperature phase, which means the γ phase facilitates more OP transitions compared to the α phase. This observation is further explored in **Section 4.3.5**.

4.3.2. Random walk properties

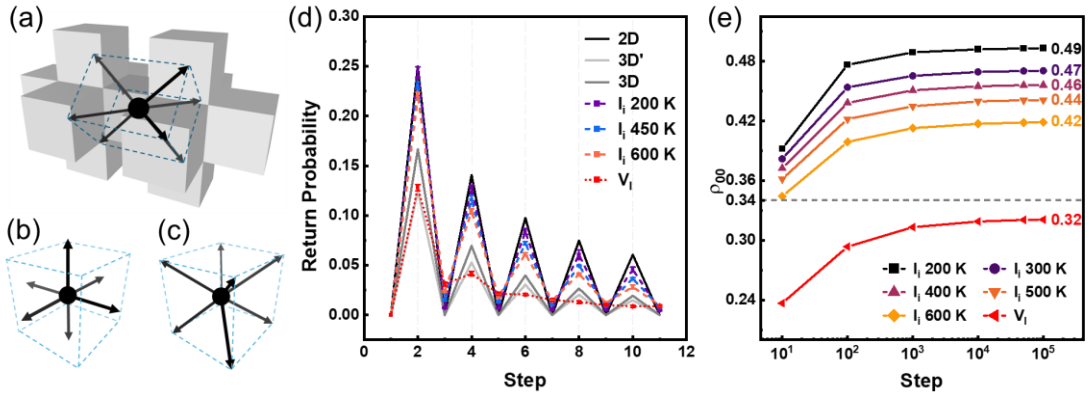

Figure 4.3 (a) The random walk grid for iodide-related defects. (b) and (c) are the random walk directions for 3D and 3D' lattices, respectively. In the 3D lattice, the walker can jump along three orthogonal axes and six directions in total, which is identical to moving through the six faces of a grid. Meanwhile, in the 3D' lattice, eight identical directions can be chosen, which is regarded as jumping through vertices. (d) The return probability for defect self-diffusion in 11 steps of hopping. (e) The accumulated return probability for defects in various temperatures. The grey dashed line represents the accumulated return probability for the random walk on the ideal 3D cubic lattice.

To further study the dynamics of iodide-related defect hopping and verify the planar-constrained motion behaviour, we performed kMC simulations with the E_a^v derived from our MD results. A key characteristic of random walks across different dimensions is the probability of returning to the original position after N steps [242]. For 2D and 3D random walks on square and cubic lattices, respectively, returning to the origin requires an even number of steps. The return probabilities for 2D [243] and 3D random walks after 2N steps can be derived as (see **Section 4.6** for derivation):

$$P_2(2N) = \left[\frac{1}{2^{2N}}\binom{2N}{N}\right]^2 = \left[\frac{(2N)!}{2^{2N}(N!)^2}\right]^2 \quad (44)$$

$$P_3(2N) = \frac{1}{6^{2N}}\binom{2N}{N}\sum_{i=0}^{N}\binom{N}{i}^2\binom{2i}{i} \quad (45)$$

Meanwhile, the behaviour of 3D random walks in a cubic lattice differs slightly from the iodide hopping process in the cubic perovskite structure. Each iodide has 8 nearest sites with D_{4h} symmetry [**Figure 4.3(a)**], while the 3D random walk on a cubic lattice has 6 identical nearest sites with O_h symmetry [**Figure 4.3(b)**]. For comparison, we also showed the return probability for 3D random walks on 8 sites (denoted as 3D'), see **Figure 4.3(c)**.

The return probability values for I_i and V_I, along with the analytical values for 2D and 3D random walks are shown in **Figure 4.3(d)**. The results indicate that the return probability of I_i self-diffusion at temperatures at 200 K is highly approaching that of the 2D random walk within 11 steps of hopping. Even at 600 K, the return probability still exhibits clear 2D behaviour. In contrast, the hopping of V_I is like the 3D' random walk. One difference between 3D' random walk and V_I hopping is that V_I has a probability to return to its origin after an odd number of jumps, which is attributed to the jumps along the three edges of one face of the Pb-I octahedral. These results demonstrate that the diffusion of I_i exhibits a quasi-2D random walk behaviour, while the hopping of V_I is more like 3D' random walk behaviour. This difference is due to the anisotropic hopping mechanisms of I_i, as mentioned above.

The quasi-2D random walk behaviour of I_i has implications for the dynamic equilibrium of defect density of halide perovskites, i.e. the defect can be easily created and annihilated [244]. Due to the relatively soft lattice [177], Frenkel defect pairs can be easily

created in halide perovskites under light irradiation [181,194]. For the short-range annihilation, we can analyze the accumulated returning probability for I_i and V_I of the Frenkel pairs to evaluate the possibility of annihilation. ρ_{00} is defined as the probability of returning to the original site within an infinite time. If ρ_{00} is equal to 1, the random walk is "recurrent", and will inevitably pass through every point in space within infinite time. If ρ_{00} is smaller than 1, on the other hand, the random walk is called "transient", then walkers may never reach certain points. According to Pólya's recurrence theorem [245], random walks in 1D and 2D lattices are recurrent, while a 3D random walk on a cubic lattice has a ρ_{00} of approximately 0.34, indicating a transient nature and the inherent difficulty of returning to the origin over an infinite timeline.

Figure 4.3(e) shows that the ρ_{00} value gradually asymptotes to values less than one for all types of defects hopping in 10^5 timesteps, indicating that the I_i self-diffusion is transient in the long run. As the temperature increases, the probability of returning decreases. Additionally, the returning probability is about 0.42 even at very high temperatures like 600 K, which is larger than 0.34 for simple 3D random walks. At temperatures below 400 K, the I_i motion exhibits approximately equal probability of returning to its origin and escaping. On the other hand, the ρ_{00} value of V_I is 0.32, slightly smaller than that in the 3D cubic lattice. The data suggests that the quasi-2D random walk behaviour would facilitate the annihilation of defects, subsequently decreasing defect concentration.

4.3.3. Defect lifetimes

However, in reality, perovskites exhibit a considerable number of defects. Intuitively, if I_i and V_I are situated on two planes that are at a distance from each other, this planar confinement should, in fact, hinder the annihilation of defects, and the ρ_{00} is mainly contributed by the short-range domain directly after the creation. Therefore, we further investigated the impact of this restriction on the lifetime of defects. In our previous MD analysis (**Section 2.1**), we focused on the short-range properties of defects, observed within a relatively limited spatial and temporal scope, where the estimated defect density reached magnitudes of 10^{19} cm^{-3}. In contrast, real halide perovskite crystals typically exhibit lower defect densities, as indicated by various measurements [164,165]. Furthermore, the effects

of planar-confined motions in short ranges cannot represent their impact on the whole. To address this discrepancy and assess the broader implications of planar-confined motion, we conducted extensive kMC simulations involving both mobile I_i and V_I. Although kMC is less accurate than other simulation methods in our case, it can provide simulations with larger supercell sizes for different defect concentrations under longer time scales. Also, it is computationally faster so that we can simultaneously perform thousands of simulations at a time to derive the statistical information of a system.

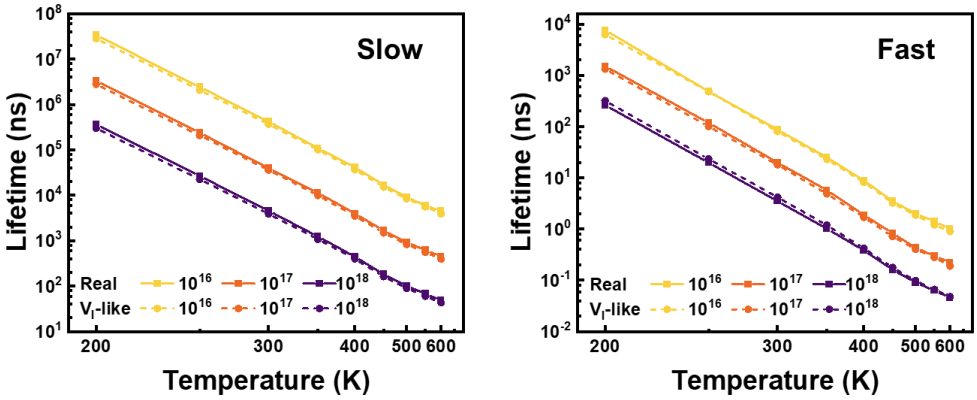

Figure 4.4 The slow and fast decay mode of Frenkel defect. The lifetime with respect to temperatures in different defect densities (in units cm^{-3}) is plotted. The dashed lines represent the corresponding curves when no planar motion of I_i. The data were fitted with tri-exponential decay of the ratio of annihilated pairs.

In our simulations, I_i and V_I were randomly placed on the halogen sites within the perovskite structure, with annihilation considered to occur when their separation fell below the distance of one unit cell. It is worth noting that the actual annihilation range might extend further, facilitated by the anti-Frenkel disorder process [1]. The defect lifetimes were determined by fitting the defect annihilation rates over time with exponential functions across multiple kMC trajectories in each temperature and density. We discovered that defect annihilation occurs through multiple modes, resulting in several distinct lifetimes. Therefore, we adopted tri-exponential fits in our analysis. In the three modes, the slowest and medium mode of annihilation obtains a similar trend. Thus, we only show the results in **Figure 4.4** for the slowest and fastest mode.

Defect lifetimes decreased as temperatures rose, owing to the accelerated defect self-diffusion in higher temperatures that increases the encountering rate. Besides, higher defect densities lead to shorter defect lifetimes. To show the influence of the planar-constrained motion of I_i on lifetime, we have also drawn the results when I_i moves with the same pattern as V_I for comparison, in which no planar constraints exist, denoted as "V_I-like". This comparison revealed that planar confinement notably extends the slow-mode lifetime, especially at temperatures below 350 K. However, as temperatures increase, leading to more frequent OP hopping by I_i, the differences in lifetimes between the actual and V_I-like scenarios begin to diminish across varying defect densities. Specifically, at 200 K, the slow-mode lifetimes for defect densities of 10^{16}, 10^{17}, and 10^{18} cm^{-3} increased by 19.65%, 19.80%, and 21.26%, respectively, compared to the V_I-like cases, highlighting a pronounced effect at higher defect densities. Meanwhile, as the defect density decreases, the discrepancy between the real defect lifetime and the V_I-like case decreases. Conversely, the lifetimes of the fast mode exhibited a diverging pattern. Compared to the V_I-like case, the fast-mode lifetime increased by 22.33%, 13.88%, and decreased by 17.73% at concentrations of 10^{16}, 10^{17}, and 10^{18} cm^{-3} under 200 K, respectively. This indicates that at a high-concentration defect environment, the hopping pattern of I_i actually increases the defect lifetime.

The planar-confined motion can also be explained by the 2D random walk properties. We attribute the slow mode to the annihilation of distant defects through V_I and OP hopping, while the fast mode to the annihilation dominated by IP hopping for nearby defects. For the slow annihilation mode, at lower defect concentration, the I_i and V_I are less likely to be in close proximity or within the same plane. Thus, the I_i approaching V_I requires slower OP hopping. Additionally, the quasi-2D random walk behaviour of I_i becomes more significant at lower temperatures. This results in a reduced annihilation rate of defects. Conversely, for the fast mode at high defect concentrations, the probability of defects being in the same plane increases, making IP hopping a more critical factor in facilitating rapid annihilation. During this time, the V_I-like case, due to its more distinct "transient" nature, actually becomes less favourable for defect annihilation.

In real circumstances, the newly created Frenkel defects are close to each other, leading to rapid annihilation before effective. Based on the conclusions above, the capability for rapidly annihilating closer defects partially explains the self-healing properties of

perovskites [195,246].

4.3.4. Self-diffusion in different directions

Understanding the preferred crystallographic directions for ion diffusion within halide perovskite systems is crucial for op timizing the performance of optoelectronic and ion conductor devices. The structure of halide perovskites and their characteristic planar-confined motion suggest differential impacts on diffusion along the common crystallographic directions: ⟨100⟩, ⟨110⟩, and ⟨111⟩. To quantify diffusion capabilities, MSD serves as a valuable metric, allowing us to project defect displacement in specific directions to derive the projected MSD.

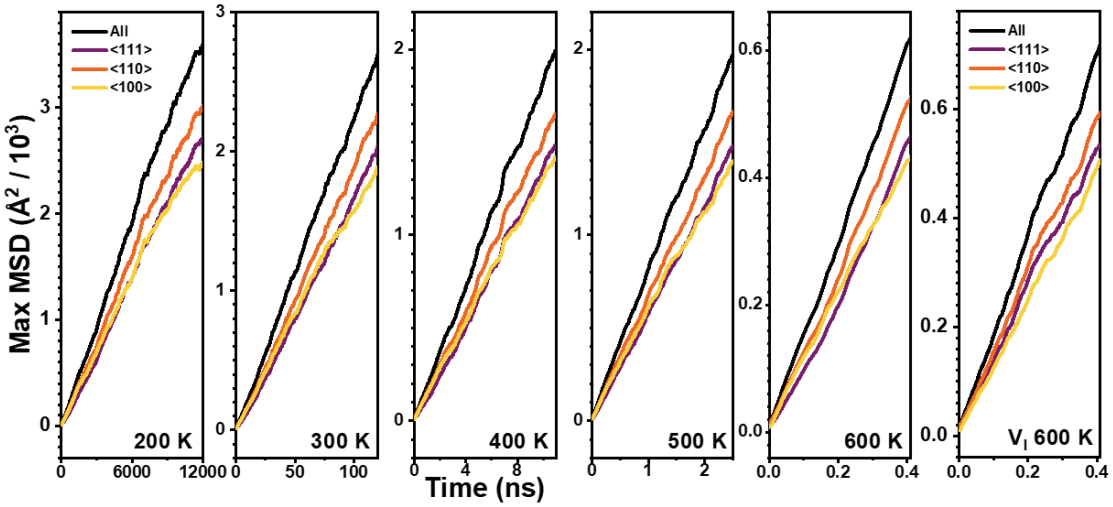

Figure 4.5 The MSD extracted from kMC simulations in different temperatures. The black lines represent the total MSD of defects. The other lines represent the MSD in each specific lattice direction. The rightmost figure is for V_I, while the others are for I_i.

Figure 4.5 presents the average MSD across 400 kMC trajectories for I_i and V_I diffusions under different temperatures. Notably, the ⟨110⟩ directions consistently exhibit the most rapid increase in MSD, surpassing the ⟨100⟩ and ⟨111⟩ directions under all temperature conditions. Intuitively, this is understandable, since I_i and V_I mostly move along the edge

of the Pb-I octahedral, the displacement vector for every single hopping event is along one ⟨110⟩ direction. Any movement in other directions are the combination of movement in ⟨110⟩ directions. Meanwhile, the identification of the slowest diffusion direction varies with the duration or distance of defect diffusion. Initially, the ⟨111⟩ directions show the slowest MSD. This is because of the planar confinement of I_i, since the ⟨111⟩ displacement requires the assistance of OP hopping, which happens much fewer than IP hopping. But to our surprise, over time, the ⟨100⟩ directions gradually become the slowest, as shown in the rightmost figure of **Figure 4.5**. This suggests that although individual I_i diffusion behaviours remain constant, the collective quasi-2D diffusion properties of I_i populations tend to average out, aligning more closely with V_I migration patterns over extended periods.

An interesting aspect for further exploration is how this directional shift in the slowest diffusion speeds might influence material properties. Considering the case of I_i diffusion under 300 K, where the transition from ⟨111⟩ to ⟨100⟩ as the slowest diffusion direction occurs at around 85 ns, corresponding to an MSD of about 1400 $Å^2$ and an estimated defect displacement of about 37 Å – a distance spanning only a few lattice lengths. Thus, the diffusion of the ⟨111⟩ direction does not need to be considered in practice, except in specific scenarios such as near grain boundaries or where defects are close. These observations suggest that the movement of I_i along the ⟨100⟩ directions is inherently slower, making the {100} orientation a potential barrier to iodine diffusion in devices and potentially inhibiting halide diffusion. Therefore, for optoelectronic devices that benefit from slower diffusion rates, the {100} orientation is preferable. Conversely, devices designed to maximize ion diffusion speed, such as ion capacitors or resistive switching devices [247,248], would benefit from a {110} orientation. This anisotropic behaviour of defects partially explains the performance enhancement through crystal-direction engineering reported in previous experimental studies [249,250].

4.3.5. Anisotropy from the lattice

In this section, the anisotropy effects of orthorhombic perovskite lattice were investigated, particularly the observed affinity of I_i for apical sites. As shown in **Figure 4.6**, To validate this observation and remove the influence of the initial defect site on our simulation

outcomes, we performed MD simulations with I_i placed in varying initial positions demonstrated in **Figure 4.9**. Within these figures, we use "*e*" and "*e'*" to denote the two equivalent equatorial directions, while "*a*" denotes the apical direction. The plotted lines trace the orientation of defects over time in these specified directions, with values between zero and one signifying the degree of parallelism to a crystallographic direction.

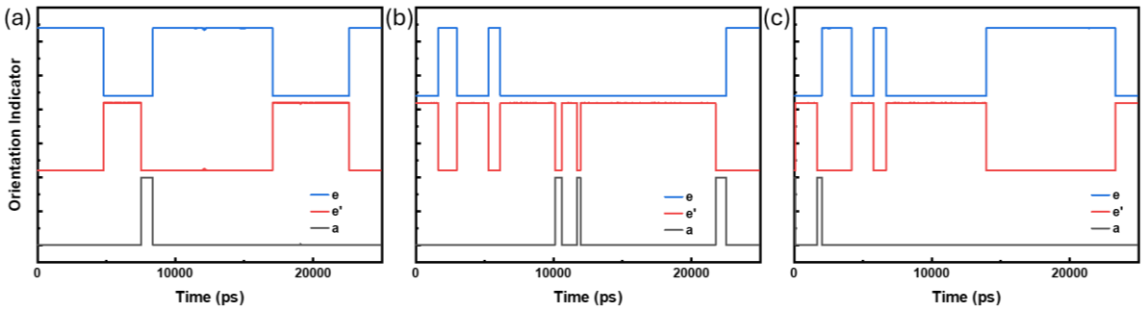

Figure 4.6 The orientation of I_i over time under 350 K. The initial state of each figure is set up in the (a) apical site oriented in e direction, (b) apical site oriented in e' direction, and (c) equatorial site oriented in a direction.

From **Figure 4.6(a) and (b)** we can see that the defect starting at the two apical sites with different orientations will be stable and occasionally transfer from one to the other. Meanwhile, I_i could only hop and lie in the equatorial site in a few cases, and will quickly return to other sites. It is also clear in **Figure 4.6(c)** that the initial I_i inside the equatorial plane hopped rapidly into other planes at the very beginning of the simulation. This represents the defect states lying in equatorial planes possess energetically unfavourable states.

The phenomenon means that the I_i diffusion behaviour varied in temperature, which is also observed in **Section 4.3.1**. In α-phase the three directions are identical, no preference on defect orientation exists. Thus, the diffusion is statistically the same on every lattice plane in the {100} family. While in β- and γ-phases, the preference of lying in apical sites could lead to enhanced movement along the apical directions and inhibit the diffusion inside the equatorial planes.

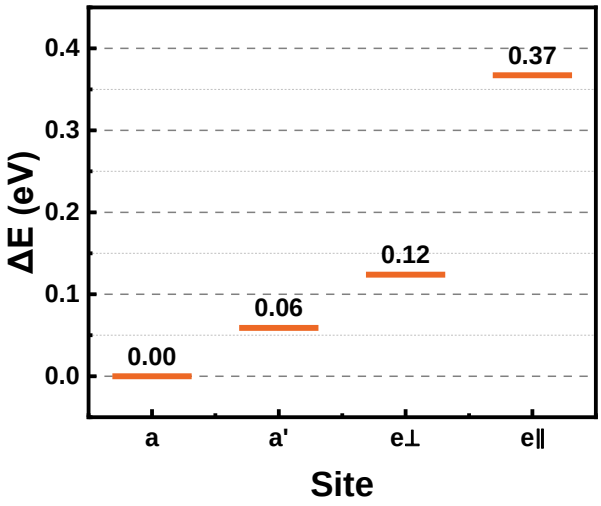

Figure 4.7 The relative energy between defect states. The symbols a and a' denote the two apical defect sites with different orientations, while symbols $e_{\perp}$ and $e_{/\!/}$ denote the equatorial defect site with the I_i configuration perpendicular and parallel to the equatorial plane, respectively. The energy of a is set at 0 eV for reference.

To energetically examine the origin of this behaviour, we have performed DFT calculations to analyse I_i states across various sites and orientations (**Figure 4.7**). A symmetry analysis with Pymatgen's [251] point group analyser reveals the equivalence of iodides in equatorial positions. Hence, it is sufficient to select and compute configurations for I_i at apical and a single equatorial position. We found that the apical sites obtain the lowest energy among the others. The two defect configurations on apical sites are different in energy. On the other hand, the defects in the equatorial plane with an orientation toward the apical direction (denoted as $e_{/\!/}$) are energetically unfavourable. The energy differentials between these configurations create an "energy well", effectively confining the defect within the apical Pb-I chain. Knowing the energy of each state, the probability for a certain defect configuration can be estimated following the Boltzmann distribution:

$$p_i = \frac{1}{Z} e^{-\frac{\varepsilon_i}{k_B T}} \tag{46}$$

where p_i is the probability of a defect in state i, Z is the partition function, ε_i is the energy of state i, k_B is the Boltzmann constant, and T is the temperature. The results show that the possibilities of finding I_i in the two apical configurations are approximately 85% and 12%,

respectively. This suggests that I_i primarily seats at the apical sites, and that inside the equatorial plane is noticeably less. Thus, the two apical states may interchange through OP rotation which provides extra OP events but hardly contributes to the total diffusion. This result also indicates that for the application in optoelectronic devices, it is crucial to adjust the material structure to align the apical direction with a horizontal orientation to inhibit diffusion under room temperature. When temperature rises, firstly the γ-phase transforms into β-phase, the difference between a and a' diminished, leading to an identical apical site no matter which orientation the I_i is lying. Then β-phase will gradually become α-phase through octahedral tilting, and the four defect sites become identical. The energy well created by different defect site disappeared, leading to an equally distributed diffusion on each {100} plane.

Moreover, researchers have also observed a similar planar motion behaviour of I_i in hybrid $MAPbI_3$ via classical MD simulation and concluded that this results in a fast effective diffusivity [175]. However, the anisotropy of crystal lattice under room temperature and the detailed influence of this planar confinement are not discussed in detail. In our study, we have performed a more accurate deep-learning-based MD simulation, which could capture more detailed dynamic information about perovskite lattice and defects. We thoroughly analysed the influence of planar-confined motion and lattice anisotropy on defect properties for I_i in all-inorganic $CsPbI_3$. Furthermore, the current results show that the equatorial plane is energetically less favourable for I_i diffusion in all-inorganic $CsPbI_3$, making diffusion in that plane difficult. Meanwhile, although not explicitly mentioned, the results from the previous study seem to imply that I_i can still move along that specific plane in $MAPbI_3$ [175]. This could be attributed to the structure difference between the all-inorganic and organometal halide perovskites, but further systematic research is required to obtain a solid conclusion.

4.4. Conclusions

In summary, based on deep-learning MD simulation, our study explores the dynamics of iodide-related defects within all-inorganic $CsPbI_3$, particularly emphasizing the phenomenon of planar-confined motion. Our study reveals that iodide interstitials exhibit a quasi-two-dimensional self-diffusion behaviour, facilitating the rapid annihilation of newly formed anti-Frenkel defects and impeding the separation of interstitials from

vacancies. Moreover, the planar-confined motion can increase the defect lifetime in the long run. Furthermore, our analysis determined that the self-diffusion of defects is directionally dependent within the crystal lattice, with the ⟨100⟩ directions exhibiting the slowest diffusion, and the ⟨110⟩ directions the fastest. The study provides insights into directional anisotropy in defect mobility in $CsPbI_3$, which has significant implications for the design and performance optimization of perovskite optoelectronic devices, ionic capacitors, and resistive switching devices.

4.5. Methods

4.5.1. Deep Potential Molecular Dynamics

In order to improve the accuracy of classical MD close to that of density functional theory (DFT) calculations [206], the deep potential MD (DPMD) [252] framework was adopted to provide a deep learning model for the prediction of the potential energy surface of the simulation systems. We have adopted the deep learning force field of $CsPbI_3$ with defects from our previous research [1]. In general, the dataset contains DFT calculation results by the Vienna Ab-initio Simulation Package (VASP) [205]. The projector-augmented wave (PAW) pseudopotentials [206] and Perdew-Burker-Ernzerhof (PBE) parametrization [122] were adopted for the static single-point calculations and ab initio molecular dynamics (AIMD) simulations. Van der Waals interaction was considered via the DFT-D3 method with the Becke-Johnson (BJ) damping function [207]. The model effectively distinguishes between α, β, and γ phases, and has been trained in a variety of defect systems. The detailed computational setup, dataset construction process, and accuracy can be found in that research.

Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) code [214] implemented with the DeepMD-kit interface was used to perform MD simulations in this study. An 8 × 8 × 8 cubic supercell containing 512 $CsPbI_3$ units and one defect is adopted for simulating halide diffusion, which could be thermally relaxed to tetragonal and orthorhombic phases. The simulations were done under NPT ensembles with the canonical sampling thermostat using global velocity rescaling with Hamiltonian dynamics [253] and

Parrinello-Rahman barostat [148]. The timestep is 5 fs. We first performed the energy minimization to the initial configurations. Then a 50 ps NVT and a 50 ps NPT simulation were done to obtain the thermal equilibration.

4.5.2. First-principles calculations

The DFT calculations in our study are all based on the VASP code with PAW pseudopotentials and PBE parametrization. DFT-D3(BJ) was selected for the van der Waals interaction. The supercells for geometry optimization and energy calculations contain $2\sqrt{2} \times 2\sqrt{2} \times 2$ orthorhombic unit cells, equivalent to a 4×4×4 cubic supercell. The calculations used a plane-wave cutoff of 500 eV and a Γ-only K-point mesh. The energy threshold for electron self-consistent calculations is 10^{-6} eV and the force threshold on each atom for the geometry optimization is 0.02 eV/Å.

4.5.3. Postprocessing of MD trajectories

The positions of defects were extracted from MD results with the Python interface of Ovito [254] by determining the centre of mass of the Pb atoms with the coordination number of 7 (for I_i) or 5 (for V_I). Moreover, the normal direction of I_i was obtained by the cross product of two adjacent Pb-I bond directions of the I_i. The orientation indicators were obtained by first taking the absolute values of each component of the normal, and then calculating the softmax value to get the final values.

The 3D diffusion coefficients for iodides were obtained from the linear fitting of the time derivative of MSD of iodides (d=3 for 3D):

$$D = \frac{1}{2d} \lim_{t \to \infty} \frac{d}{dt} \mathrm{MSD}(r) \tag{47}$$

The MSD function is obtained via the Einstein formula:

$$\mathrm{MSD}(r) = \left\langle \frac{1}{N} \sum_{i=1}^{N} |\boldsymbol{r} - \boldsymbol{r}(t_0)|^2 \right\rangle_{t_0} \tag{48}$$

In which N is the number of iodides in the system, r is the coordinate of the iodides. The diffusion coefficients were derived with the Python package MDAnalysis [255]. Then the diffusion activation energies E_a^D were derived from the data. Since the diffusion under lower temperatures is significantly slower than that at higher temperatures, we have used different simulation time for different temperatures to obtain accurate MSD, as listed in **Table 4.3**.

Table 4.3 The simulation length for different temperatures.

Temperature (K)	I_i (ns)	V_I (ns)
350	200	175
375	150	150
400	25	50
≥ 425	25	25

The hopping activation energies E_a^v were obtained by a similar method but using different types of data. The trajectories were processed using breakpoint analysis to obtain the numbers of hopping events of I_i IP and V_I hopping. The OP hopping events were obtained by analyzing the orientation following the same process. The E_a^v for different events were then extracted via the Arrhenius plots.

4.5.4. Rejection-free kinetic Monte Carlo simulation

The rf-kMC code was written in Python with the parallel programming language Taichi [256]. The activation energies in the Monte Carlo simulation were set according to the fitting results obtained in DPMD simulations. All the simulations were performed on grid points similar to I sites in $CsPbI_3$. Two events were considered: IP and OP hopping. The code can randomly choose one event from the two according to the possibility calculated via Boltzmann distribution. If the IP hopping event was chosen, the defect would choose one over four IP neighbour sites with the same probability. On the other hand, if the OP hopping event was chosen, the defect would change its orientation. The time of each event can be estimated by the formula:

$$t = -\frac{\ln r}{k} \tag{49}$$

In which r is a random number in the interval [0, 1], and k is the rate constant for one single event. Then the total simulation time can be estimated by summing over the time of all individual events.

To quantitively analyze the diffusion along a specific crystal direction family, the projected MSD was derived from kMC trajectories. Meanwhile, due to the presence of multiple identical crystal directions within each lattice family, the projection relying on a single direction cannot represent the whole picture and lead to underestimations of the diffusion. For instance, the ⟨110⟩ and ⟨111⟩ families feature twelve and eight equivalent directions, respectively, in a cubic lattice. Thus, we simultaneously collect all the equivalent directions and choose the direction with the largest MSD in each time step as the representation for the diffusion of that crystal direction family.

4.6. Supplementary Section

4.6.1. Additional figures

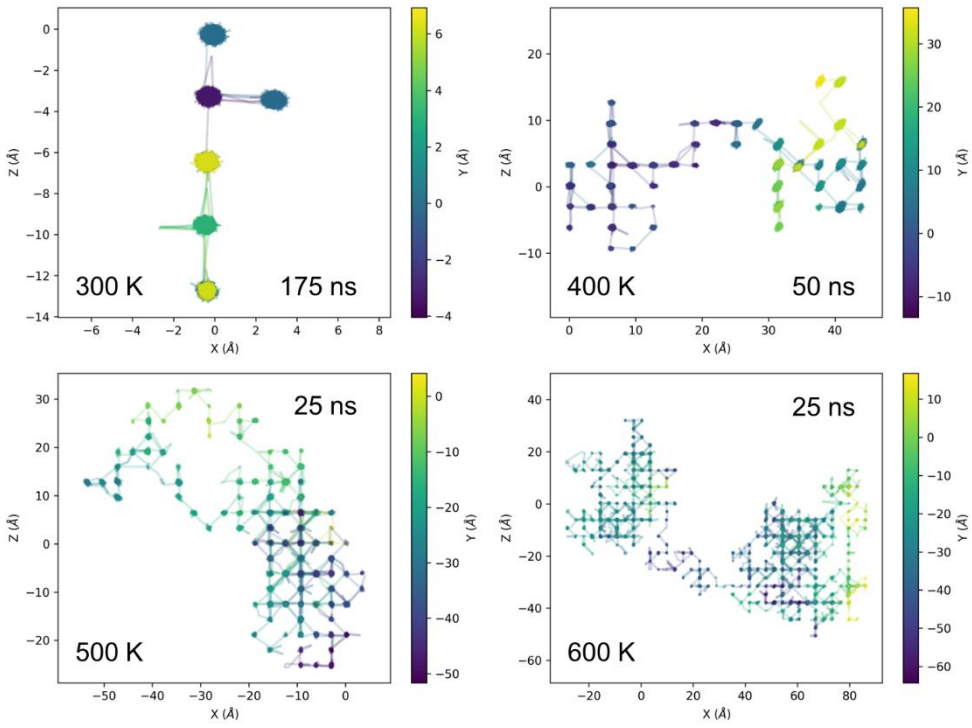

Figure 4.8 The unwrapped projected trajectories of V_I under 300, 400, 500, and 600 K.

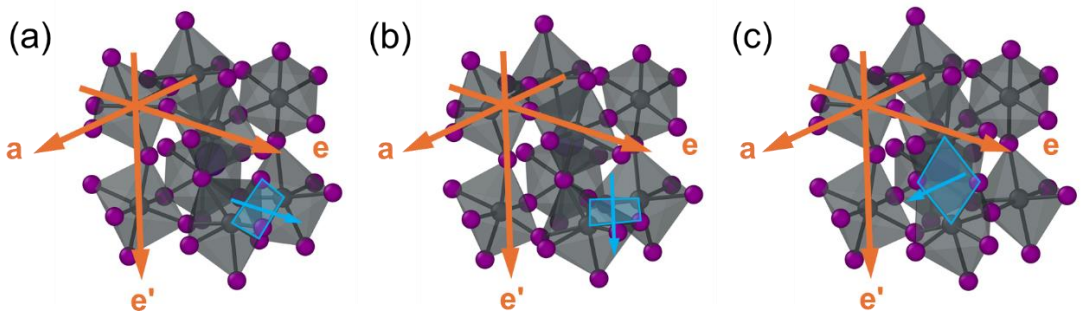

Figure 4.9 Schematics for the initial structures used in **Figure 4.6**. The orange axes represent the direction for the apical (denoted as a) and the two equatorial (denoted as e and e') directions. The blue rhombi denote the I_i position, and the blue arrows denote the orientation of the defect.

4.6.2. Derivation of 3D random walk properties

In a 3D simple cubic lattice, all sites have 6 identical neighbours: up, down, front, back, left, and right. In each time step, the walker will randomly choose one neighbour to jump in. We define the N-step random walk process as a list of displacement $\{D_1, D_2, \dots, D_N\}, D_i \in \{X^+, X^-, Y^+, Y^-, Z^+, Z^-\}$, in which X, Y, and Z denote the axis of displacement and +/- denote the direction of the step. For every single step, we need to choose 1 of the 6 elements. Thus, the total number of possible paths for an N-step random walk is 6^N. The total displacement is defined as the summation of all displacement in the random walk process:

$$S_N = \sum_{i=1}^{N} D_i \tag{50}$$

In order to return to the original place, S_N = 0, the total number of steps of "+" jumps should be equal to the "-" jumps in each direction. Thus, only even steps of move can make it home. Let us only consider 2N steps. Suppose there are i, j, and k times of X^+, Y^+, Z^+ walk, it must fulfill the relation i+j+k=N to go back to the origin. The returning probability can be calculated with the following procedure:

1. Choose N “+” movements from the 2N steps.
2. Choose i X movements from the N steps (same for “-”).
3. Choose j Y movements from the N-i steps (same for “-”).
4. Choose k Z movements from the N-i-j=k steps, which is equal to one (same for “-”).

We obtain the returning probability by summing over all possible i, j, and k:

$$(2N) = \frac{1}{6^{2N}} \sum_{i=0}^{N} \sum_{j=0}^{N-i} \binom{2N}{N} \binom{N}{i}^2 \binom{N-i}{j}^2 \tag{51}$$

Using Chu–Vandermonde identity, we can simplify the second summation of eq. (6):

$$P_3(2N) = \frac{1}{6^{2N}} \binom{2N}{N} \sum_{i=0}^{N} \binom{N}{i}^2 \binom{2(N-i)}{N-i} \tag{52}$$

Since the binomial coefficient is symmetric and $\binom{N}{i} = \binom{N}{N-i}$, we can get the result by reversing the summation index:

$$P_3(2N) = \frac{1}{6^{2N}} \binom{2N}{N} \sum_{i=0}^{N} \binom{N}{i}^2 \binom{2i}{i} \tag{53}$$

Chapter 5.

Deep Charge: A Deep Learning Model of Electron Density from One-Shot Density Functional Theory Calculation

5.1. Abstract

Electron charge density is a fundamental physical quantity, determining various properties of matter. In this study, we have proposed a deep-learning model for accurate charge density prediction. Our model naturally preserves physical symmetries and can be effectively trained from one-shot density functional theory calculation toward high accuracy. It captures detailed atomic environment information, ensuring accurate predictions of charge density across bulk, surface, molecules, and amorphous structures. This implementation exhibits excellent scalability and provides efficient analyses of material properties in large-scale condensed matter systems.

5.2. Introduction

In the vast domain of materials science and condensed matter physics, electron charge density is a fundamental property that describes the spatial distribution of electrons in a material. It provides invaluable insights into overall material behaviors and is an essential quantity in various quantum mechanical frameworks, especially in density functional theory (DFT) [114], where the ground-state properties of a many-body system can be described as functionals of the electron charge density. Commonly, DFT solves the Kohn-Sham (KS) equations [26] via the self-consistent loop [205] to adopt the ground state

charge density $\varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^n)$, where charge density ϱ is a function of the space position $\boldsymbol{r}$ and the positions of all n atoms in the system $\{\boldsymbol{R}_i\}_{i=1}^n$. A significant challenge of this approach is the computational complexity of $O(N^3)$ with respect to the number of electrons. The harshness of density functional optimization impedes the application of DFT in mesoscopic systems with thousands or even millions of atoms. This limitation has spurred the search for alternative approaches that can efficiently represent charge density without compromising accuracy.

The electron density generally obtains "nearsightedness" and screening effect in many-atom systems [257,258]. These effects make it possible to introduce a radius r_c around a space point, beyond which any perturbation on the atomic environment only provides negligible impacts on the electron density at that point. This is the locality property of electron density, which does not apply to the KS wavefunctions. In that case, the original time-consuming KS self-consistent calculation can be turned into mapping local atomic environment information to the electron density. Such a mapping can naturally be obtained via techniques such as machine learning. A well-trained machine learning model can be used for predicting charge density in different atomic systems, and even be scalable for determining large-scale material systems. Currently, several machine-learning models have been proposed to predict electron charge density as alternatives to first-principles calculations, as listed in **Table 5.**1.

One commonly used strategy is to express the charge density as the summation of a set of basis functions such as spherical harmonics, with the coefficients fitted via machine learning [259–265]. The rotational invariance of charge density can be automatically protected with a reasonable basis set. These methods are efficient but mainly used in charge-localized systems, and the accuracy is highly dependent on the quality and complexity of basis functions and usually requires DFT calculations of plenty of different atomic structures to train. On the other hand, the grid-based methods can be adopted to obtain more flexibility [111,266–270]. These methods focus on mapping every unique local environment around real space points to non-degenerate values in the latent space, which thus contains enough information for charge density representation. The mapping should generally be invariant under three physical symmetry operations: translation, rotation, and permutation, as will be discussed in detail below. Achieving these without losing structural detail information is a nontrivial task. Moreover, grid point methods

usually require large numbers of data points from various structures to obtain satisfactory accuracy.

Table 5.1 The published works on machine-learning charge density. The top 7 methods belong to basis function methods, while the bottom 6 are grid-point methods. The unit for MAE and RMSE values is e/Å^3.

Method	Systems	Accuracy	Data size	Ref.
SA-GPR	Small molecules	MAE: 1.0% ~ 1.2%	~1000 structures	[259]
SA-GPR	Organic dimers	rMAE[b]: 0.3%~1.8%	~2300 structures	[260]
SALTED	Al, Si, I_h-Ice	%RMSE[c]: ~1.5%	100 structures	[261]
EGNN[a]	DNA	rMAE ~1%	922 structures	[262]
SchNOrb	Small molecules	\	25000 structures	[263]
EGNN	Water clusters	rMAE: ~1.4%	6000 structures	[264]
ML-HK map	Small molecules	\	100~350 structures	[265]
Gaussians	PE and Al thin films	RMSE[d]: ~5×10^{-4}	> 10^7 points	[266]
CGCNN	Polymer, zeolite	RMSE: 0.04~0.18	>3×10^5 points	[267]
Deep Density	Molecules, Al	rRMSE[e]: 0.5%~2.2%	10^5~10^6 points	[268]
Gaussians	Organics	RMSE: 0.03~1.52	~1.5×10^9 points	[269]
EGNN	Molecules, cathode	rMAE: 0.06%~0.27%	10^3~10^5 structures	[270]
JLCDMs	Benzene, Al, MoS_2	RMSE: 0.006~0.008	~1.6×10^6 points	[111]
This work	Si, Al, Al-Mg, water	MAE: 0.05%~0.4%	>10^5 points	

[a] Equivariant graph neural network.

[b] Relative mean absolute error. $\sum_i^N|\rho_i-\hat{\rho}_i|/\sum_i^N|\rho_i|$, in which ρ_i is the real charge density of the ith grid point, $\hat{\rho}_i$ is the represented value of that point.

[c] Percentage root means square error, defined in the original article.

[d] Root mean square error. $\sqrt{\sum_i^N(\rho_i-\hat{\rho}_i)^2}/N$, in which N is the total number of grid points.

[e] Relative root means square error. $\sqrt{\sum_i^N(\rho_i-\hat{\rho}_i)^2}\Big/\sqrt{\sum_i^N\rho_i^2}$.

Herein, to overcome the above challenges for the grid-based method, we introduce a deep-learning charge model (Deep Charge) to represent electron charge density. We have adopted the embedding framework from Deep Potential (DP) [157] to extract the local atomic environment, which naturally preserves the system's symmetry. This implementation

requires very few structures for training. Remarkably, a one-shot DFT calculation is sufficient for simple systems like crystalline silicon and aluminium to reach a mean absolute error (MAE) below 7×10^{-4} e/Å^3. The model provides a scalable representation of charge density in bulk, surface, and complex systems such as alloy and amorphous, involving semiconducting and metallic systems. Our Deep Charge model can thus efficiently assist in the prediction and analysis of charge density in large-scale condensed matter systems with DFT accuracy.

5.3. Deep Charge model

The most important reasons why the Deep Charge model can effectively capture the physics consistent with KS-DFT are the nearsightedness [258] and the screening effect [271,272] of materials. When a perturbation in the atomic structure is introduced within a local environment of a material, the ground state distribution of electrons adjusts accordingly. The magnitude of the charge density change diminishes with increasing distance from the perturbation due to the destructive interference of the density amplitude concerning single-particle eigenstates. In gapped systems or disordered systems, the decay is exponential, causing a rapid decrease in influence. Meanwhile, in ordered gapless systems such as metals, the decay follows power laws rather than exponential, breaking the nearsightedness to some extent. On the other hand, the screening effect of metallic electrons shields the long-range Columb potential from the perturbed atoms, forcing the system to be nearsighted. These results ensure that almost all the information about local charge density is encoded in the atomic structures within a suitable cutoff radius. The Deep Charge model can thus be trained to establish the relationship between the charge density and atomic structures approaching the DFT accuracy.

The continuous charge density is discretized into a space grid in grid point methods. Suppose there are n atoms in the system, the charge density value at position $\mathbf{r}$ can be written as a function of the coordinates of atoms $\{\boldsymbol{R}_i\}_{i=1}^n$,

$$\rho = \varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^n) \tag{54}$$

where all the information of the system is required to determine the density at a single point. Instead, we can use only the coordinates of neighbouring atoms $\{\boldsymbol{R}_i | i \in n_{r_c}(\boldsymbol{r})\}$, where $n_{r_c}(\boldsymbol{r})$ denotes the set of atoms within the cutoff radius r_c surrounding $\mathbf{r}$ with a size of n_c.

$$\varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n}) \cong \varrho\big(\boldsymbol{r}, \{\boldsymbol{R}_i | i \in n_{r_c}(\boldsymbol{r})\}\big) \tag{55}$$

Ideally, this charge density representation should be designed carefully to fulfil the invariance under three symmetry operations:

1. Any translation $\hat{\mathcal{T}}_{\boldsymbol{d}}$ of the whole system $\hat{\mathcal{T}}_{\boldsymbol{d}}\varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n}) = \varrho(\boldsymbol{r}+\boldsymbol{d}, \{\boldsymbol{R}_i+\boldsymbol{d}\}_{i=1}^{n})$ should result in the same value as $\varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n})$;
2. Any operation $\hat{O} \in \mathrm{O}(3)$ centered at $\mathbf{r}$ should obey $\hat{O}\varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n}) = \varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n})$;
3. Any permutation $\hat{\pi}$ on labels of the same atomic species fulfils $\varrho\left(\boldsymbol{r}, \{\boldsymbol{R}_{\hat{\pi}(i)}\}_{i=1}^{n}\right) = \varrho(\boldsymbol{r}, \{\boldsymbol{R}_i\}_{i=1}^{n})$.

To preserve the translational invariance of the density, the relative coordinates $\boldsymbol{r}_i = \boldsymbol{r} - \boldsymbol{R}_i = (x_i, y_i, z_i)$ are adopted:

$$\varrho\big(\boldsymbol{r}, \{\boldsymbol{R}_i | i \in n_{r_c}(\boldsymbol{r})\}\big) = \varrho\big(\{\boldsymbol{r}_i | i \in n_{r_c}(\boldsymbol{r})\}\big) \tag{56}$$

The density can be represented via neural networks as

$$\varrho\big(\{\boldsymbol{r}_i | i \in n_{r_c}(\boldsymbol{r})\}\big) = \mathcal{F}\big(\mathcal{D}\big(\{\boldsymbol{r}_i | i \in n_{r_c}(\boldsymbol{r})\}; \boldsymbol{\theta}_d\big); \boldsymbol{\theta}_f\big) \tag{57}$$

where $\mathcal{D}$ represents the descriptor that extracts all the local environment information, and $\mathcal{F}$ represents the fitting net for charge density. The trainable neural network parameters for them are $\boldsymbol{\theta}_d$ and $\boldsymbol{\theta}_f$, respectively. In the implementation of Deep Charge, the descriptor is constructed to conserve the rotation and permutation symmetry. Herein, we use the two-body embedding DeepPot-SE descriptor with radial and angular information [160]:

$$\mathcal{D} = \frac{1}{n_c^2}\mathcal{G}^T\mathcal{R}\mathcal{R}^T\mathcal{G}_< \tag{58}$$

in which $\mathcal{R} \in \mathbb{R}^{n_c \times \{1,4\}}$ is the coordinate matrix with the form of

$$\mathcal{R}_i = s(r_i)\left(1, \frac{x_i}{r_i}, \frac{y_i}{r_i}, \frac{z_i}{r_i}\right) \tag{59}$$

in which $r_i = \|\boldsymbol{r}_i\|$ is the norm of relative coordinates. $s(r)$ is the switching function defined with a fifth-order polynomial:

$$s(r) = \begin{cases} \dfrac{1}{r}, & r < r_s \\ \dfrac{[u^3(-6u^2 + 15u - 10) + 1]}{r}, & r_s \le r < r_c \\ 0, & r \ge r_c \end{cases}, \quad u = \frac{r - r_s}{r_c - r_s} \tag{60}$$

where r_s is the radius where smoothing starts. The construction of $\mathcal{R}\mathcal{R}^T$ is naturally invariant under O(3) symmetry operation. The embedding matrix $\mathcal{G}$ is for obtaining the permutation invariance according to ref. [273], and $\mathcal{G}_<$ is the first few columns of $\mathcal{G}$ (10 columns in our case). The form of $\mathcal{G}$ is trained with neural networks $\mathcal{N}$:

$$\mathcal{G}_i = \mathcal{N}\big(s(r_i)\big) \tag{61}$$

The fitting net $\mathcal{F}$ obtains a simple fully connected feed-forward network structure. The detailed charge density prediction workflow is shown in **Figure 5.1**. The continuous space is discretized into grids. The model will take the atomic configuration and grid points as input and find the local environment for each grid point in parallel, which will be used in the embedding networks to construct the descriptor and then input to the fitting networks. The final charge density can thus be collected from the grid points.

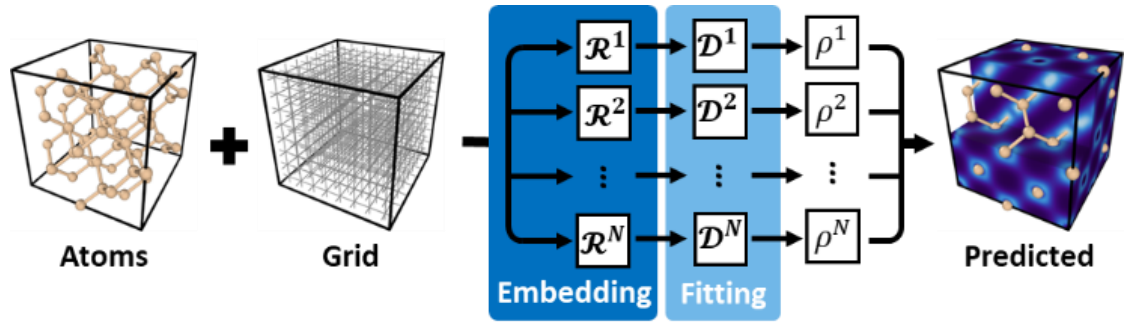

Figure 5.1 Workflow of Deep Charge prediction. Suppose there are *N* grid points in the system, each grid point goes through one network separately.

Other similar grid-point methods with high accuracy, such as those that use Gaussian fingerprints [266] and JLCDMs [111], usually require a pre-defined set of filter functions to construct fingerprints. These functions are applied to the relative distances between the grid points and each surrounding atom and are then summed over the results from all the atoms to obtain the fingerprint. Consequently, the number of filter functions corresponds to the number of consistent fingerprints. This construction method has the advantage of reducing the complexity of training, as environmental information is extracted via manually designed functions. Meanwhile, the quality of extracted information can highly depend on the quality and number of the filters and may lack flexibility in very complex circumstances. In the case of Deep Charge, on the other hand, the descriptor is trainable, providing additional degrees of freedom to map local environmental information into latent space. This implementation endowed the model with more adaptability to complex systems such as amorphous structures, as demonstrated and discussed later in this article.

5.4. Results and discussion

5.4.1. Performance in bulk systems

In this section, we examined the convergence relation between the number of data points for training and the accuracy in bulks. We chose crystalline silicon with 64 atoms for this study. Silicon is an extensively studied semiconductor with localized valence electrons at bond positions. The training sets were based on the charge density obtained from only one structure, while the testing was performed on all the grid points from an unseen structure. The results are shown in **Figure 5.2**. In addition, we compared the convergence with models trained on 5 different structures, 10^5 data points for each structure selected, displayed with blue dashed lines in **Figure 5.2(a)**. The model can reach an MAE of 6.7×10^{-4} e/Å^3 trained with 10^5 data points. It is noticeable that the one-structure models exhibit similar accuracy to the five-structure models, indicating that a little charge density information of one single structure of structure is enough to learn the patterns in crystals. The parity plot of **Figure 5.2(b)** has a coefficient of determination (R^2) of 0.999957, also indicating a nearly linear relationship. Meanwhile, as shown in **Figure 5.2(c)-(f)**, the absolute charge density errors are larger around the centre of atoms, where the fluctuation and charge density are also

higher. The absolute error distribution is also shown in **Figure 5.7(a)** in **Section 5.6**(see also references [122,159,205,206,274–277] therein).

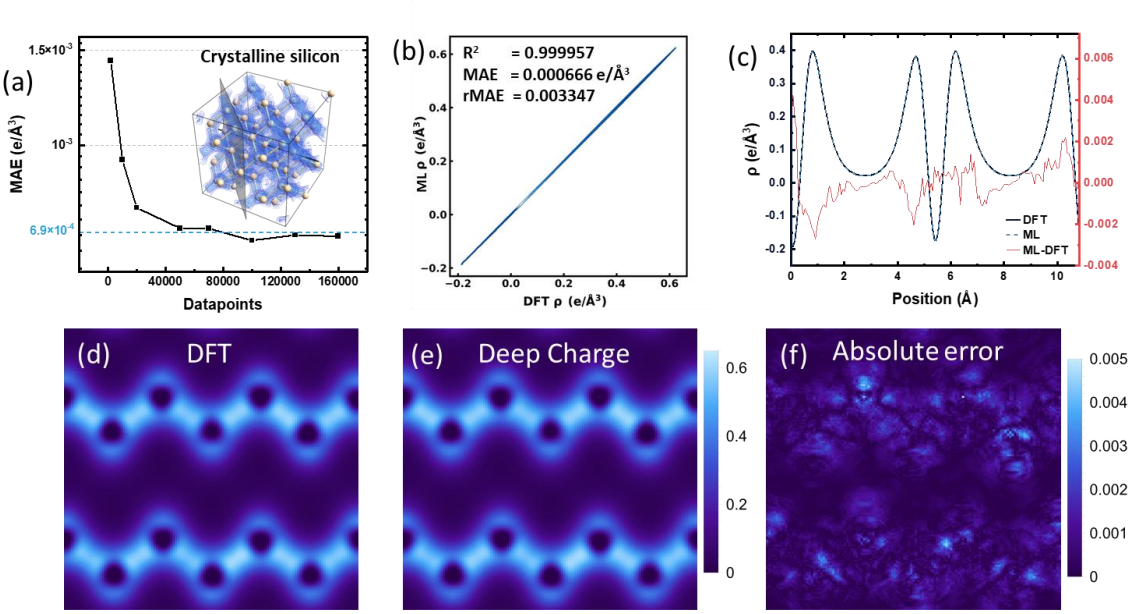

Figure 5.2 The performance of the Deep Charge model on bulk silicon. (a) The convergence relation between the number of data points used for training and the MAE values for the testing set. (b) The scatter plot of predicted and test densities from a whole testing structure. Notice that the negative values of charge density are induced by the on-site correction terms in projector augmented wave (PAW) pseudopotentials. (c) The charge density along the line is indicated in the inset of (a). (d) and (e) is the charge density computed from DFT and machine learning, respectively. The slice is denoted in the inset of (a). (f) The absolute error. Colour bars are in the unit of e/Å^3.

5.4.2. Performance in surface systems

To test the capability of the model under different atomic environments, we further evaluated the charge density representation in surface structures. The aluminium surface with 144 atoms was chosen, because it is a benchmark material for testing charge density representation with highly delocalized valence electrons[8,13,15,18]. As shown in **Figure 5.3(a) and (b)**, the accuracy of our model approached below 6×10^{-4} e/Å^3 after training with 10^5 data points, which is quite similar to the five-structure model. We have extracted the charge density along the line across the surface and bond positions, as shown in **Figure**

5.3(c). The charge density near the surface is lower than that in the bulk. This is owing to the surface relaxation, the atoms near the surface are loosely bonded, reducing the charge density between atoms. Moreover, the charge density inside the film also varies considerably with the disturbance of atomic structures than that in silicon, which is also clearly shown in **Figure 5.4**. The results from the Deep Charge model greatly overlap those from the DFT calculation, meaning that the embedding approach can accurately capture the subtle changes in the environment, with the error fluctuating under 0.004 e/Å^3 along the line.

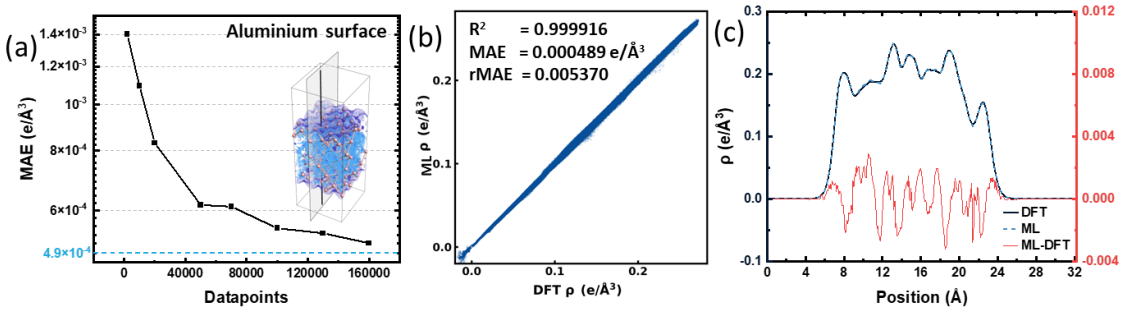

Figure 5.3 (a) The convergence relation of machine learning model in aluminium surface. (b) the scatter plot of predicted and test densities. (c) The charge density along the line is indicated in the inset of (a).

In addition, many other works have tested their model in aluminium with different system sizes and structures. Thus, we herein briefly compare other models with ours. The SALTED [261], Deep Density [268], and JLCDM [111] models have been examined in small bulk aluminium systems, in which the errors are ~1% (%RMSE, 4 atoms), 2×10^{-2} eV/Å^3 (MAE, 32 atoms), 4.81×10^{-4} eV/Å^3 (MAE, 32 atoms), respectively. For comparison, we trained and tested our model in the same systems and obtained an MAE of 3.37×10^{-4} eV/Å^3, lower than their values. Meanwhile, Gaussian fingerprints [266] can reach an RMSE of about 10^{-3} eV/Å^3 in 6-layer thin film systems (144 atoms) with about 10^5 datapoints, while the error can be further decreased to 6×10^{-4} eV/Å^3 but requires two orders of magnitude more data than us. Overall, the Deep Charge model can provide an efficient and accurate way of representing charge density in aluminium.

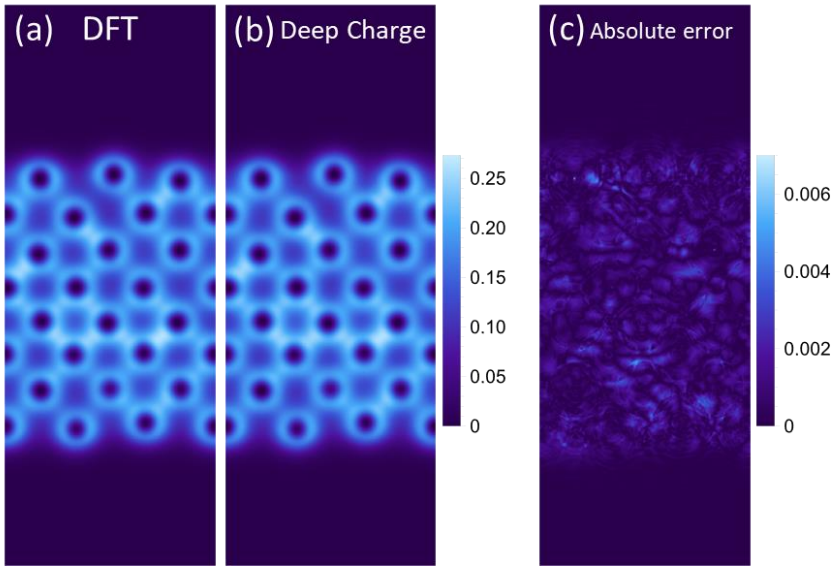

Figure 5.4 (a) and (b) are the computed charge density from DFT and machine learning, respectively. The slice is denoted in **Figure 5.3(a)**. (c) the absolute error.

5.4.3. Performance and scalability in complex structures

In this section, we examined the scalability of the machine learning model in complex structures. Scalability is one of the most valuable aims of machine learning, since complex structures normally require a large system size in practical applications, while the ordinary DFT calculations for large supercells are extremely challenging. Therefore, we have studied the model in alloy and amorphous materials. Specifically, aluminium-magnesium alloy (108 atoms), amorphous silicon (64 atoms), and water (48 atoms) were tested.

First, a face-centred cubic structure with randomly placed aluminium and magnesium was used for the alloy study. The training convergence towards the number of structures trained on was tested, and 10^5 grid points from each structure were chosen for training, as shown in **Figure 5.5(a)**. The MAE approaches 8.0×10^{-4} e/Å^3 with more than 5 structures of training data. **Figure 5.5(b)** and **(c)** indicate that the increase in system size barely influences the predictability of the machine learning model. The MAE in the system with 256 atoms is 8.14×10^{-4} e/Å^3, slightly smaller than 7.66×10^{-4} e/Å^3 from the system with 108 atoms. The detailed charge density is shown in the first row of **Figure 5.6**. The absolute

errors are mostly contributed from the centre area of magnesium atoms, with a value of around 0.06 e/Å^3.

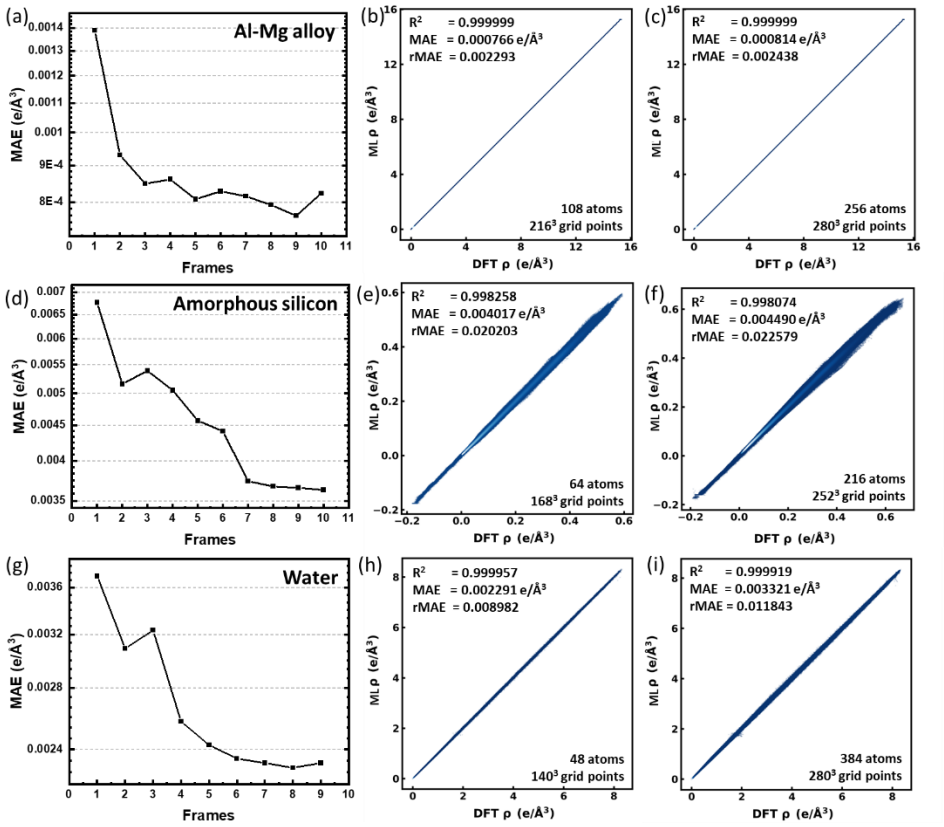

Figure 5.5 (a), (d) and (g) are the convergence relation between the number of structures used for training and the MAE values for the testing set of aluminium-magnesium alloy, amorphous silicon, and water, respectively. (b), (e) and (h) are the corresponding scatter plots for charge density calculated by DFT and Deep Charge for a whole grid of data from a testing structure. (c), (f) and (i) are the corresponding parity plots in larger supercells.

Next, amorphous silicon and water reveal similar results, as shown in **Figure 5.5(d)-(i)** and **Figure 5.6**. For amorphous silicon, the MAE reached below 0.004 e/Å^3 after training with 7 structures of data. The max absolute error for the test structure is 0.037 e/Å^3, nearing the silicon atoms. Moreover, we have also examined the performance of our model in the silicon melting process, which obtains MAEs of about 0.005 e/Å^3. The details are discussed in the **Section 5.6**. In the case of water, the MAE decreases below 0.0023 e/Å^3 after training with 6 structures. The predictions in silicon with 216 atoms and water with 384 atoms are

as good as that in smaller systems. However, with the increase in system size, the accuracy for water decreases relatively large, which may be induced by the rather small training structures. These results show that the accuracy of Deep Charge in amorphous structures is relatively lower than that of crystalline but can still be a good representation. Besides, the larger error around atom centres could be related to the cutoff function we used, which increases the weight of adjacent atoms but probably not be the best description of the electron distributions. A reasonable design of the cutoff function by reference to the near-core charge density should improve the predictability of machine learning models. The extremely high scalability will pave the way for predicting charge density in large-scale simulations.

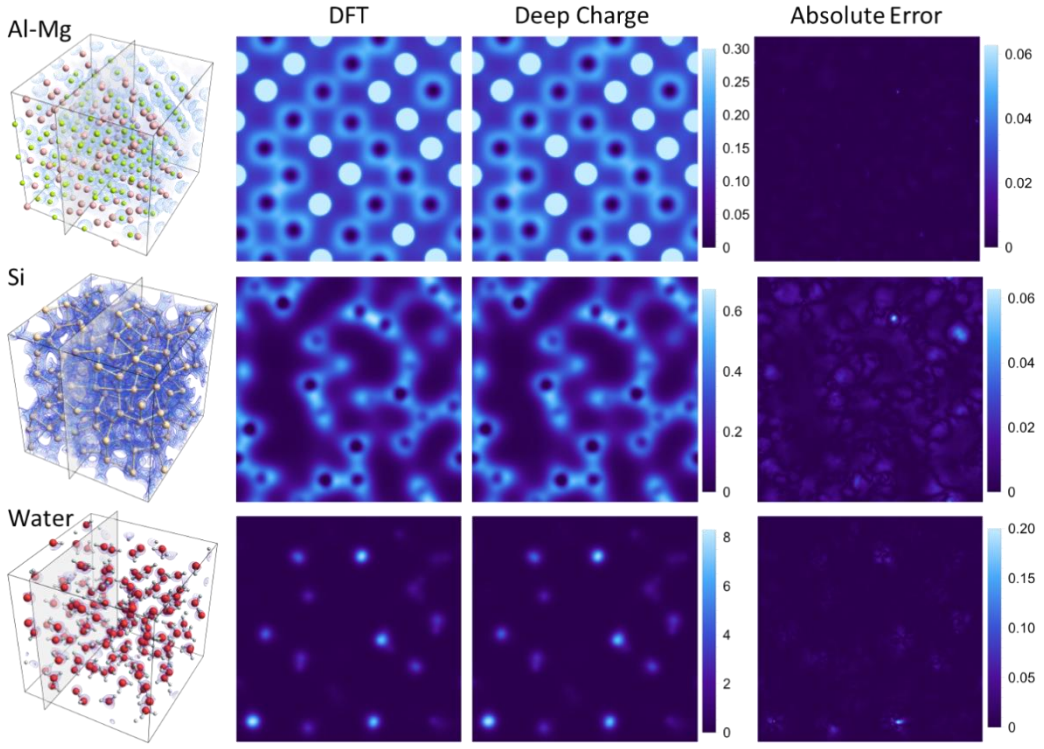

Figure 5.6 Charge density in large aluminium-magnesium alloy, amorphous silicon, and water supercells. Noticed that the maximum value of charge density for aluminium-magnesium alloy is above 15.3 e/Å^3. We only display the charge density values in the interval of 0~0.3 e/Å^3 to show the details clearly.

From all the results above, we can see that the Deep Charge model can effectively capture the radial distribution of the charge near atoms as well as the angular distribution around bond positions, due to two factors. First, the construction of the local environmental matrix $\mathcal{R}$ of each grid point has included the generalized coordinates with angular information for every neighbouring atom. Second, the embedding net $\mathcal{G}$ of each grid point works as filters that add weights to each neighbouring atom according to the distance and chemical species. From a physical point of view, this can be regarded as the contribution of each atom to the charge density representation in latent space. Different from the above-mentioned filters such as Gaussian, the parameters of embedding nets are adaptive and can be learned to align with the DFT results via training. Together with the weight of atomic contribution and detailed generalized coordinates, the model obtains its ability to represent the charge density of arbitrary material systems.

5.5. Conclusion

In summary, we have realized a new charge density representation with machine learning techniques. Our Deep Charge model can accurately represent charge density in semiconductor, metal, and molecular systems. The symmetry-preserving descriptor enhances the training efficiency. Only one single DFT calculation is enough for training extremely high-accuracy models for simple crystal structures. It can also capture the fluctuations of the environment and represent complex systems by training with fewer data. Meanwhile, further improvement in the accuracy and efficiency of this model is still possible. A reasonable design of the cutoff function and loss function may increase the charge density accuracy near the atoms. Moreover, implementing a concurrent learning strategy can automate the sampling process of data points used for training, leading to more efficient learning and less human intervention.

In addition, reasonable modifications to the model can extend its application. Apart from the descriptor employed in our study, the attention-based model [77] offers the potential to create a universal charge density representation covering more elements. Furthermore, the model can be adapted to predict the spatial distribution of other fields, such as spin-charge density and the local density of states on grids. Achieving this requires suitable

modifications to the network architecture and transitioning from scalar output values to two- or multi-dimensional vectors. The feasibility and implementation will be carried out in future works.

Our model is easy to implement and can provide an approach for fast charge density prediction avoiding the usage of KS quasi-particle wavefunctions. Looking forward, with the predicted charge density, many other physical quantities like ground-state energy and force can be derived theoretically based on the orbital-free DFT frameworks [66], which thus assist the electronic structure analysis, stability prediction, elastic behaviour, etc. in large-scale condensed matter systems with the first-principal accuracy. Additionally, the predicted charge density can be integrated into physics-informed neural networks to efficiently derive these quantities. This approach further enhances the versatility and efficiency of our model in capturing and analyzing the complex physics of condensed matter systems.

5.6. Supplementary Section

5.6.1. Density functional theory calculations

The charge density data are all based on the DFT calculations using the Vienna Ab initio Simulation Package (VASP) [205]. We used the Perdew-Burke-Ernzerhof (PBE) [122] exchange and correlation functional with PAW pseudopotentials to handle the core electrons [206]. The charge density data for training only contains the valence electrons. Explicitly, the electrons in 3s and 3p orbitals for silicon and aluminium, 2p and 3s orbitals for magnesium, 2s and 2p orbitals for oxygen were treated as valance electrons. Our models were then based on the charge density from valance electrons.

The electronic self-consistent calculation stops when the difference between the current and previous iteration is less than 10^{-6} eV for all the single-point calculations. For the accurate charge density calculation of silicon, the tetrahedron method with Blöchl corrections was adopted, with a 2×2×2 k-point mesh and a plane wave energy cutoff of 520 eV. A relatively low criterion was adopted for AIMD and geometry optimization, in which we set Gaussian smearing with a sigma value of 0.05, the electronic stop condition as 0.001

eV, and the ionic stop condition as -0.08 eV/Å. The timestep for the AIMD simulation was set to 3 fs. For aluminium, the first-order Methfessel-Paxton method was used for smearing, with a sigma of 0.22. The energy cutoff was 650 and Monkhorst-Pack k-point mesh with a k-spacing of 0.1 $Å^{-1}$ was adopted. The AIMD simulations of aluminium were the same as that of silicon but with a timestep of 5 fs. The settings for aluminium-magnesium alloy are similar to that of aluminium, but with a k-spacing of 0.2 $Å^{-1}$ for faster calculation. For H_2O, the SCAN functional [274] was used to model the exchange-correlation energy. The energy cutoff was set to 680 eV and the k-point spacing is 0.5 $Å^{-1}$.

5.6.2. Datasets construction

The structures for constructing datasets were sampled from AIMD or deep potential molecular dynamics (DPMD). We used AIMD under 300 K for crystalline silicon and aluminium. In amorphous silicon, we first ran an AIMD under 2000 K to obtain the randomly distributed atoms. Then geometry optimizations were performed to obtain the amorphous structures. On the other hand, we adopted DPMD with a Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [214] for H_2O and aluminium-magnesium alloy, for which the Deep Charge models were available from previous research [276,277]. The NPT simulations were under 300 K and 100 kPa. The initial structures of aluminium-magnesium alloy were generated by randomly replacing half of the aluminium atoms with magnesium in the face-centred cubic lattice. Then the randomly generated structures were relaxed for 0.5 ps with a timestep of 5 fs. For water, the timestep was set to 1 fs, and the structures were collected after simulating for more than 5 ps. The charge density data used in this work were then constructed from the obtained structures with DFT calculations described above. We collected the charge density generated from VASP. In total, we have collected 7 data frames for crystalline silicon, 109 data frames for amorphous silicon, 20 for aluminium, 20 for aluminium-magnesium alloy, and 10 for water. The final frame of data in each system was selected to be shown in this work as the testing structure. For each charge density file, we randomly chose 10^5 data points as default, except when testing the convergence of accuracy towards the number of data points. These data points are separated into 500 batches, each batch contains 200 points for training or validating.

5.6.3. Deep Charge model

To realize the training of the Deep Charge model, it is important to notice that our grid point strategy can be equivalent to the usage of pseudo-atoms in the DP scheme. The charge density at a specific grid point can be regarded as the "atomic energy" of that point, which can be represented by the DP model [159]. Thus, the Deep Charge model can output the distribution by inputting the real atoms in the system together with the pseudo-atoms on each grid point and then collecting the predicted scalar values of the pseudo-atoms. The r_c for all the models was set to 6.5 Å with an r_s of 0.0 Å, where the value of r_s did hardly affect the result. For H_2O, the training was operated for 2×10^6 steps, with a learning rate decreasing from 10^{-3} to 10^{-8} exponentially. We have chosen a three-layer ResNet structure of the form [20, 40, 80] for the embedding network and [200, 200, 200] for the fitting network. The same parameters are adopted for silicon, except for a longer training with 5×10^6 steps. For aluminium, the training was operated for 4×10^6 steps, with a learning rate decreasing from 10^{-3} to 10^{-7} exponentially. For aluminium-magnesium alloy, the training was operated for 3×10^6 steps, with a learning rate decreasing from 10^{-3} to 10^{-6} exponentially.

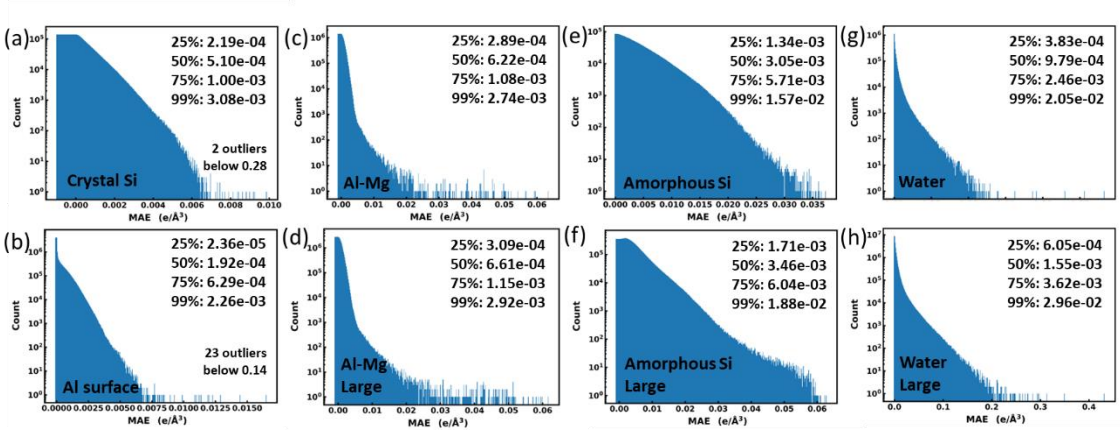

Figure 5.7 The absolute error distribution for grid points in the tested systems. There are two outliers not drawn located at 0.277 and 0.281 e/Å^3 in the crystalline silicon, and 23 outliers in aluminium with values below 0.14 e/Å^3. The text in each chart indicates how many percents of the errors are below a certain value.

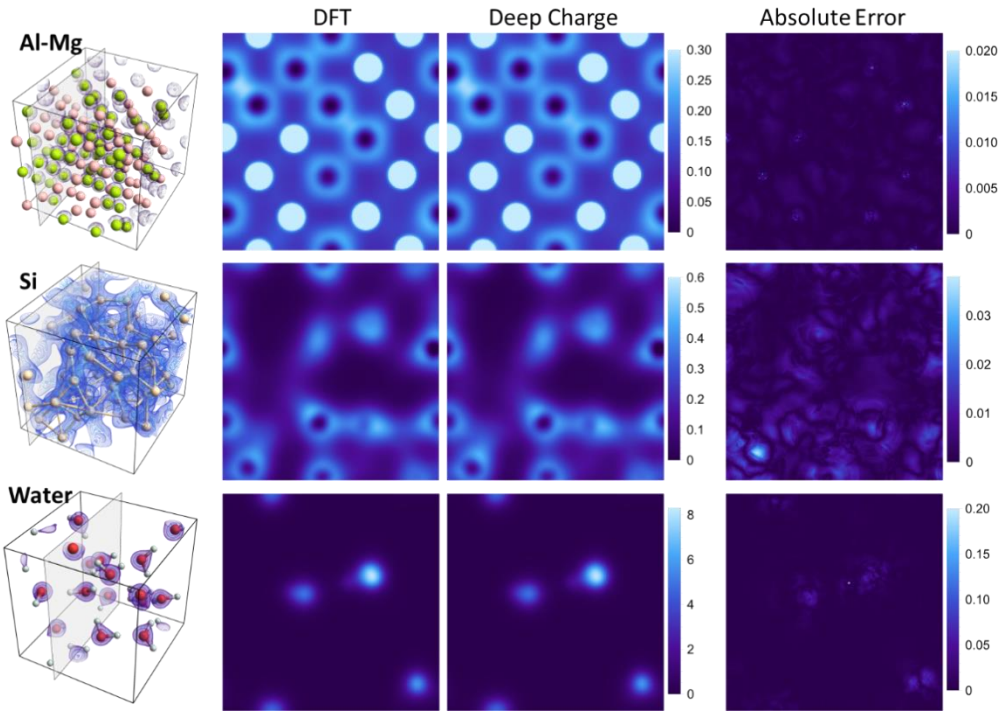

Figure 5.8 From top to bottom shows the comparison of charge density computed from DFT and Deep Charge in small systems of aluminium-magnesium alloy, amorphous silicon, and water, respectively. The first column shows the testing structures. The middle two columns display the charge density slices obtained from DFT and machine learning within the plane with max absolute error, respectively. Noticed that the maximum value of charge density for aluminium-magnesium alloy is above 15.3 e/Å^3. We only display the charge density values in the interval of 0~0.3 e/Å^3 to show the details clearly. The fourth column represents the absolute errors.

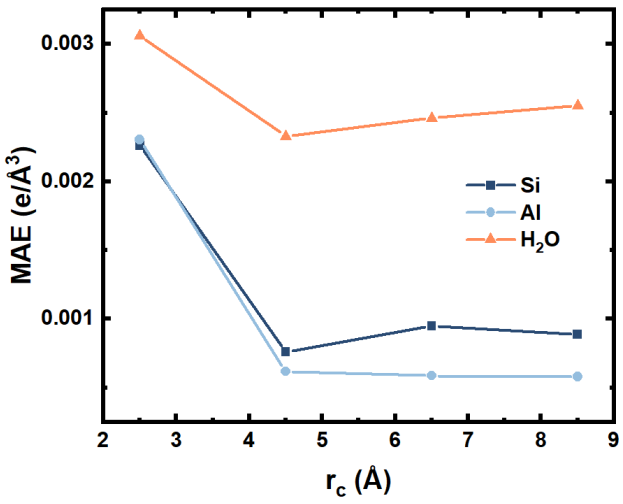

Figure 5.9 The convergence test between MAE and cutoff radius in silicon, aluminium and H_2O.

5.6.4. Prediction in transition states

We further tested our model in the melting process of silicon, as shown in FIGs. S4 and S5. The training data only contains the 64-atom datasets of 10 frames of amorphous silicon and 5 frames of crystal silicon under 300 K. The test frames were extracted from 1.5-ps AIMD simulations with temperature rises from 1800 to 2400 K in the 216-atom system. The MAE of the Deep Charge model is around 0.005 e/Å^3 for all the frames, slightly higher than the testing results in FIG. 5. Further enhancements are possible by fine-tuning our model with charge density data obtained in the specific range of temperatures.

In addition, we have also predicted the charge density of melting in 1000-atom systems. It can be seen that the 1000-atom system undergoes a clear nucleation process via melting, compared to the 216-atom system. The charge density varies significantly in the liquid phase, due to the rapid changes in silicon bond lengths. Electrons tend to accumulate between two silicon atoms at shorter distances. Conversely, the charge density is more consistent across the bonds in the solid phase. Our model can capture these details and thus indicates its predictability.

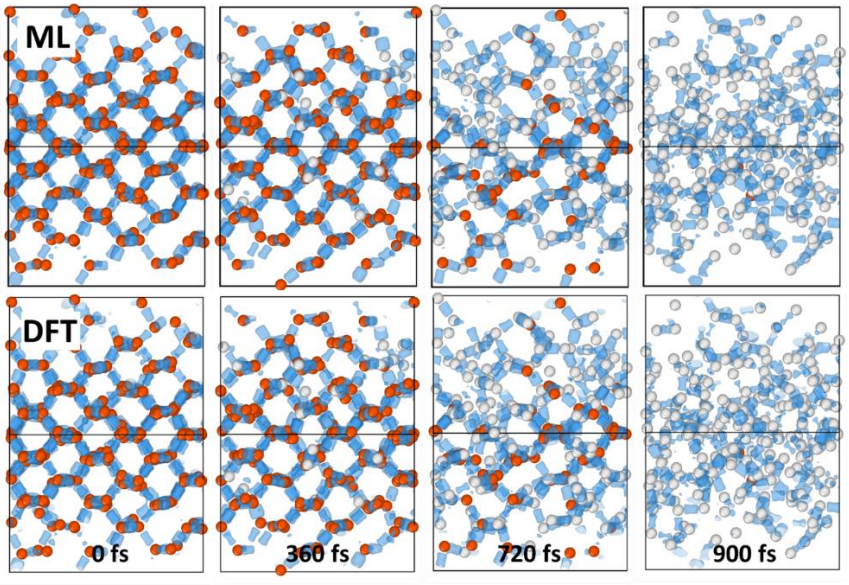

Figure 5.10 The charge density from Deep Charge prediction and DFT for the melting process of silicon in a large system with 216 atoms. The isosurface is at the value of 0.45 e/Å^3. The orange spheres represent the silicon atoms in a cubic diamond structure, while the white spheres represent the atoms in an amorphous structure.

TABLE S1. The prediction accuracy corresponds to frames in **Figure 5.10**.

Timestep (fs)	R^2	MAE (e/Å^3)	rMAE
0	0.997375	0.004844	0.024358
360	0.997196	0.005040	0.025340
720	0.997087	0.004944	0.024864
900	0.997522	0.004543	0.022848

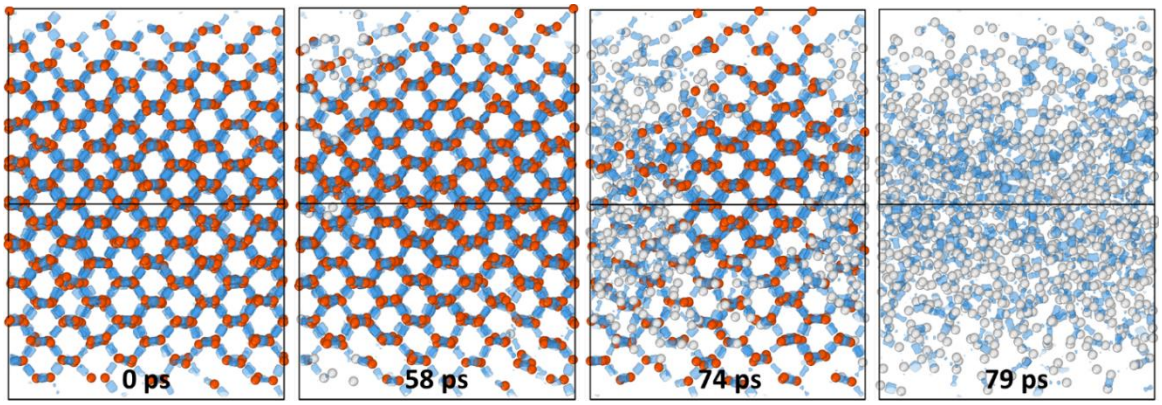

Figure 5.11 The charge density prediction of the silicon melting process in a 1000-atom system. The isosurface is at the value of 0.45 e/Å^3.

Chapter 6.

Summary and Prospective

6.1. Summary

In this thesis, the machine learning algorithm has been used to predict the potential energy surface and electron density for materials. I have constructed a deep learning potential for $CsPbI_3$ halide perovskite covering a wide temperature range and various defect states. The deep learning potential is further used in **Chapter 3** and **Chapter 4** for molecular dynamics simulations.

In **Chapter 3**, the long-time simulation trajectories of $CsPbI_3$ with iodide interstitials and vacancies under various temperatures were collected. The trajectories were used to sample rare defect hopping and annihilation events. They show that the existence of iodide vacancy, interstitial, and anti-Frenkel disorder can lower the creation barrier of additional anti-Frenkel disorders. Based on them, we found the anti-Frenkel disorders can elongate defect hopping distance and accelerate interstitial-vacancy annihilation. Moreover, the disorders can assist the hopping outside the moving plane of interstitials.

In **Chapter 4**, we found the planar confinement in the motion of iodide interstitials. The iodide interstitials can mainly move inside the {100} plane family of halide perovskite. This phenomenon can influence the statistical properties of halide diffusion, leading to a quasi-2D random walk pattern. This pattern will enhance the annihilation of nearby interstitial-vacancy pairs, but however, increase the lifetime of defects in the long run. Moreover, iodide interstitials can hardly locate and diffusion inside the equatorial plane and prefer to align with the apical direction. Based on the observations, the halide perovskite-based devices can be fine-tuned with different orientations selected.

Besides implementing machine learning in materials simulation, we have designed a

workflow for predicting the electron density of matter. In **Chapter 5**, the Deep Charge model was introduced. The model has been tested in plenty of physical systems covering metallic, semi-conducting and isolating materials. Numerous structures have been involved in the testing, such as the bulk and surface of crystalline materials, and more complex amorphous and alloy structures. The model shows high accuracy compared to the ab initio results. Especially for the crystalline bulk and surface, the Deep Charge model can obtain mean absolute errors in the magnitude of 10^{-4} e/Å^3, with only 10^5 datapoints from a single frame of ab initio result. It has also shown scalability towards complex configurations. The results indicate the ability of Deep Charge to efficiently represent the electron density with less data, which is suitable for the exploration of unknown systems.

In summary, the Deep Potential machine learning framework has been applied to the materials simulations and property predictions. It demonstrates the ability to explore unknown phenomena from existing datasets. The deep learning technics have a linear time complexity and can therefore be easily implemented for simulation and prediction in large systems for a long time. The deep learning approaches described throughout the thesis not only enhance our understanding of material properties at the microscopic level but also provide insights for optimizing material design and functionality in practical applications.

6.2. Prospective

Currently, AI for Science is active and still highly under development. Even though there are plenty of AI models that have been proposed for solving specific problems in materials science, there are still obstacles to be overcome.

For the material simulations, we generally require collecting a large number of datasets for a specific material system and then finding an accurate model structure before doing the truly interesting calculations. Moreover, a specific model can only be used in the corresponding systems. If we want to add new atom types in the simulation, new datasets containing the atom types and their possible interactions with other known atom types in the original datasets. This fact leaves machine learning potential energy surfaces inconvenient to use, especially in doped systems. For example, our model was trained with

the Cs, Pb, and I atom types and can be applied to $CsPbI_3$, CsI, and PbI_2 systems, but it is not possible to directly used in $CsPbBr_3$ systems. Therefore, it is necessary to collect and construct a database that covers the periodic table, based on which the large atomic model could be trained [278]. The ultimate goal of machine learning interatomic potential is to find a universal representation of the potential energy surface for material systems. Moreover, only a few machine learning potentials can include other degrees of freedom of material systems, such as spin, density of states, charged states, etc. which are all important for specific properties. Meanwhile, several methods that could attach to these degrees of freedom often lack generalization and are only applicable to certain systems. In the future, one goal is to implement the necessary degrees of freedom for accurate dynamic simulations.

For the Deep Charge model, the accuracy and universality should be further increased. The model holds promising potential not only for accelerating ab initio calculations. Since the density functional theory indicates that the physical properties of materials are encoded in the electron density, we can further develop tools based on the deep learning representation of electron density for accurate predictions of other properties efficiently. In addition, it is also necessary to design a universal electron density representation throughout the periodic table, so we do not require to collect new datasets before any calculation. This could possibly be done by implementing the attention mechanism to the current Deep Charge model.

In summary, after my Ph.D. degree, seeking for the collaboration of experiments is one promising direction to investigate the accuracy and predictability of the machine-learning-assisted property predictions. For example, the predicted charge density can be aligned with the scanning tunnelling microscope (STM) images. Besides, it is also vitally important to go beyond the boundaries of traditional computing methods. Developing efficient machine learning techniques to screening materials based on their fundamental compositions and structures is also necessary for the inverse design of new materials. Moreover, the machine learning technics can not only learn from DFT data but also experiments, which could also assist building new computational technics that is naturally aligned with the real-world physical systems.

Reference

[1] T. Lv, Y. Liang, F. Zeng, F. Li, X. Yang, J. Huang, and R. Zheng, *Kinetic Process with Anti-Frenkel Disorder in a CsPbI3 Perovskite*, J. Phys. Chem. Lett. **15**, 2929 (2024).

[2] T. Lv, Z. Zhong, Y. Liang, F. Li, J. Huang, and R. Zheng, *Deep Charge: Deep Learning Model of Electron Density from a One-Shot Density Functional Theory Calculation*, Phys Rev B **108**, 235159 (2023).

[3] T. M. Lenton, H. Held, E. Kriegler, J. W. Hall, W. Lucht, S. Rahmstorf, and H. J. Schellnhuber, *Tipping Elements in the Earth's Climate System*, Proc. Natl. Acad. Sci. **105**, 1786 (2008).

[4] M. Arvidsson, N. Lovsjö, and M. Keuschnigg, *Urban Scaling Laws Arise from Within-City Inequalities*, Nat. Hum. Behav. **7**, 365 (2023).

[5] A. Bahrini, M. Khamoshifar, H. Abbasimehr, R. J. Riggs, M. Esmaeili, R. M. Majdabadkohne, and M. Pasehvar, *ChatGPT: Applications, Opportunities, and Threats*, (2023).

[6] G. Team, R. Anil, S. Borgeaud, et al., *Gemini: A Family of Highly Capable Multimodal Models*, (2023).

[7] D. Patterson, J. Gonzalez, U. Hölzle, Q. Le, C. Liang, L.-M. Munguia, D. Rothchild, D. R. So, M. Texier, and J. Dean, *The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink*, Computer **55**, 18 (2022).

[8] A. de Vries, *The Growing Energy Footprint of Artificial Intelligence*, Joule **7**, 2191 (2023).

[9] C. Chen, J. Fu, and L. Lyu, *A Pathway Towards Responsible AI Generated Content*, (2023).

[10] M. M. Shulaker, G. Hills, N. Patil, H. Wei, H.-Y. Chen, H.-S. P. Wong, and S. Mitra, *Carbon Nanotube Computer*, Nature **501**, 526 (2013).

[11] J. A. Tulyagankhodjaev, P. Shih, J. Yu, et al., *Room-Temperature Wavelike Exciton Transport in a van Der Waals Superatomic Semiconductor*, Science **382**, 438 (2023).

[12] J. Zhao, P. Ji, Y. Li, et al., *Ultrahigh-Mobility Semiconducting Epitaxial Graphene on Silicon Carbide*, Nature **625**, 60 (2024).

[13] B. Lilia, R. Hennig, P. Hirschfeld, et al., *The 2021 Room-Temperature Superconductivity Roadmap*, J. Phys. Condens. Matter **34**, 183002 (2022).

[14] L. Si and K. Held, *Electronic Structure of the Putative Room-Temperature Superconductor \textPb_9\textCu(\textPO_4)_6\textO*, Phys Rev B **108**, L121110

(2023).

[15] A. Kojima, K. Teshima, Y. Shirai, and T. Miyasaka, *Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells*, J. Am. Chem. Soc. **131**, 6050 (2009).

[16] A. K. Jena, A. Kulkarni, and T. Miyasaka, *Halide Perovskite Photovoltaics: Background, Status, and Future Prospects*, Chem. Rev. **119**, 3036 (2019).

[17] W. Peng, K. Mao, F. Cai, et al., *Reducing Nonradiative Recombination in Perovskite Solar Cells with a Porous Insulator Contact*, Science **379**, 683 (2023).

[18] S. M. Park, M. Wei, J. Xu, et al., *Engineering Ligand Reactivity Enables High-Temperature Operation of Stable Perovskite Solar Cells*, Science **381**, 209 (2023).

[19] M. A. Green, E. D. Dunlop, M. Yoshita, N. Kopidakis, K. Bothe, G. Siefer, and X. Hao, *Solar Cell Efficiency Tables (Version 63)*, Prog. Photovolt. Res. Appl. **32**, 3 (2024).

[20] L. Qiu, S. He, L. K. Ono, S. Liu, and Y. Qi, *Scalable Fabrication of Metal Halide Perovskite Solar Cells and Modules*, ACS Energy Lett. **4**, 2147 (2019).

[21] B. Sanchez-Lengeling and A. Aspuru-Guzik, *Inverse Molecular Design Using Machine Learning: Generative Models for Matter Engineering*, Science **361**, 360 (2018).

[22] A. Zunger, *Inverse Design in Search of Materials with Target Functionalities*, Nat. Rev. Chem. **2**, 0121 (2018).

[23] B. Park and S. I. Seok, *Intrinsic Instability of Inorganic–Organic Hybrid Halide Perovskite Materials*, Adv. Mater. **31**, 1805337 (2019).

[24] W. Kohn, A. D. Becke, and R. G. Parr, *Density Functional Theory of Electronic Structure*, J. Phys. Chem. **100**, 12974 (1996).

[25] B. J. Alder and T. E. Wainwright, *Studies in Molecular Dynamics. I. General Method*, J. Chem. Phys. **31**, 459 (1959).

[26] W. Kohn and L. J. Sham, *Self-Consistent Equations Including Exchange and Correlation Effects*, Phys. Rev. **140**, A1133 (1965).

[27] F. HERMAN, *Theoretical Investigation of the Electronic Energy Band Structure of Solids*, Rev. Mod. Phys. **30**, 102 (1958).

[28] G. Vignale and M. Rasolt, *Density-Functional Theory in Strong Magnetic Fields*, Phys. Rev. Lett. **59**, 2360 (1987).

[29] M. Gajdoš, K. Hummer, G. Kresse, J. Furthmüller, and F. Bechstedt, *Linear Optical Properties in the Projector-Augmented Wave Methodology*, Phys. Rev. B **73**, 045112

(2006).

[30] R. Iftimie, P. Minary, and M. E. Tuckerman, *Ab Initio Molecular Dynamics: Concepts, Recent Developments, and Future Trends*, Proc. Natl. Acad. Sci. **102**, 6654 (2005).

[31] M. E. Tuckerman and G. J. Martyna, *Understanding Modern Molecular Dynamics: Techniques and Applications*, J. Phys. Chem. B **104**, 159 (2000).

[32] N. Mardirossian and M. Head-Gordon, *Thirty Years of Density Functional Theory in Computational Chemistry: An Overview and Extensive Assessment of 200 Density Functionals*, Mol. Phys. **115**, 2315 (2017).

[33] D. C. Langreth and M. J. Mehl, *Beyond the Local-Density Approximation in Calculations of Ground-State Electronic Properties*, Phys Rev B **28**, 1809 (1983).

[34] A. D. Becke, *Density-Functional Exchange-Energy Approximation with Correct Asymptotic Behavior*, Phys Rev A **38**, 3098 (1988).

[35] H. Wang, T. Fu, Y. Du, et al., *Scientific Discovery in the Age of Artificial Intelligence*, Nature **620**, 47 (2023).

[36] F. -Y. Wang, *Toward a Revolution in Transportation Operations: AI for Complex Systems*, IEEE Intell. Syst. **23**, 8 (2008).

[37] A. Birhane, A. Kasirzadeh, D. Leslie, and S. Wachter, *Science in the Age of Large Language Models*, Nat. Rev. Phys. **5**, 277 (2023).

[38] J. Jumper, R. Evans, A. Pritzel, et al., *Highly Accurate Protein Structure Prediction with AlphaFold*, Nature **596**, 583 (2021).

[39] E. Callaway, *'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein*, Nature **608**, (2022).

[40] S. Greydanus, M. Dzamba, and J. Yosinski, *Hamiltonian Neural Networks*.

[41] M. Cranmer, S. Greydanus, S. Hoyer, P. Battaglia, D. Spergel, and S. Ho, *Lagrangian Neural Networks*.

[42] S.-M. Udrescu and M. Tegmark, *AI Feynman: A Physics-Inspired Method for Symbolic Regression*, Sci. Adv. **6**, eaay2631 (n.d.).

[43] Z. Liu and M. Tegmark, *Machine Learning Hidden Symmetries*, Phys Rev Lett **128**, 180201 (2022).

[44] P. Lemos, N. Jeffrey, M. Cranmer, S. Ho, and P. Battaglia, *Rediscovering Orbital Mechanics with Machine Learning*.

[45] W. S. McCulloch and W. Pitts, *A Logical Calculus of the Ideas Immanent in Nervous Activity*, Bull. Math. Biophys. **5**, 115 (1943).

[46] A. Abraham, *Artificial Neural Networks*, in *Handbook of Measuring System Design* (John Wiley & Sons, Ltd, 2005).

[47] F. Kamiran and T. Calders, *Data Preprocessing Techniques for Classification without Discrimination*, Knowl. Inf. Syst. **33**, 1 (2012).

[48] S. García, S. Ramírez-Gallego, J. Luengo, J. M. Benítez, and F. Herrera, *Big Data Preprocessing: Methods and Prospects*, Big Data Anal. **1**, 9 (2016).

[49] E. A. Felix and S. P. Lee, *Systematic Literature Review of Preprocessing Techniques for Imbalanced Data*, IET Softw. **13**, 479 (2019).

[50] R. HECHT-NIELSEN, *III.3 - Theory of the Backpropagation Neural Network**Based on "Nonindent" by Robert Hecht-Nielsen, Which Appeared in Proceedings of the International Joint Conference on Neural Networks 1, 593–611, June 1989. © 1989 IEEE.*, in *Neural Networks for Perception*, edited by H. Wechsler (Academic Press, 1992), pp. 65–93.

[51] P. J. Werbos, *Backpropagation through Time: What It Does and How to Do It*, Proc. IEEE **78**, 1550 (1990).

[52] S. Amari, *Backpropagation and Stochastic Gradient Descent Method*, Neurocomputing **5**, 185 (1993).

[53] O. Isayev, C. Oses, C. Toher, E. Gossett, S. Curtarolo, and A. Tropsha, *Universal Fragment Descriptors for Predicting Properties of Inorganic Crystals*, Nat. Commun. **8**, 15679 (2017).

[54] K. T. Schütt, F. Arbabzadah, S. Chmiela, K. R. Müller, and A. Tkatchenko, *Quantum-Chemical Insights from Deep Tensor Neural Networks*, Nat. Commun. **8**, 13890 (2017).

[55] K. T. Schütt, P.-J. Kindermans, H. E. Sauceda, S. Chmiela, A. Tkatchenko, and K.-R. Müller, *SchNet: A Continuous-Filter Convolutional Neural Network for Modeling Quantum Interactions*, in *Proceedings of the 31st International Conference on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA, 2017), pp. 992–1002.

[56] D. Jha, L. Ward, A. Paul, W. Liao, A. Choudhary, C. Wolverton, and A. Agrawal, *ElemNet: Deep Learning the Chemistry of Materials From Only Elemental Composition*, Sci. Rep. **8**, 17593 (2018).

[57] K. Das, B. Samanta, P. Goyal, S.-C. Lee, S. Bhattacharjee, and N. Ganguly, *CrysXPP: An Explainable Property Predictor for Crystalline Materials*, Npj Comput. Mater. **8**, 43 (2022).

[58] S. Kong, D. Guevarra, C. P. Gomes, and J. M. Gregoire, *Materials Representation and Transfer Learning for Multi-Property Prediction*, Appl. Phys. Rev. **8**, 021409 (2021).

[59] D. Jha, K. Choudhary, F. Tavazza, W. Liao, A. Choudhary, C. Campbell, and A. Agrawal, *Enhancing Materials Property Prediction by Leveraging Computational and Experimental Data Using Deep Transfer Learning*, Nat. Commun. **10**, 5316 (2019).

[60] J. E. Saal, S. Kirklin, M. Aykol, B. Meredig, and C. Wolverton, *Materials Design and Discovery with High-Throughput Density Functional Theory: The Open Quantum Materials Database (OQMD)*, JOM **65**, 1501 (2013).

[61] D. Jha, L. Ward, Z. Yang, C. Wolverton, I. Foster, W. Liao, A. Choudhary, and A. Agrawal, *IRNet: A General Purpose Deep Residual Regression Framework for Materials Discovery*, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Association for Computing Machinery, New York, NY, USA, 2019), pp. 2385–2393.

[62] D. Jha, V. Gupta, W. Liao, A. Choudhary, and A. Agrawal, *Moving Closer to Experimental Level Materials Property Prediction Using AI*, Sci. Rep. **12**, 11953 (2022).

[63] L. Ward, A. Dunn, A. Faghaninia, et al., *Matminer: An Open Source Toolkit for Materials Data Mining*, Comput. Mater. Sci. **152**, 60 (2018).

[64] H. Li, Z. Wang, N. Zou, M. Ye, R. Xu, X. Gong, W. Duan, and Y. Xu, *Deep-Learning Density Functional Theory Hamiltonian for Efficient Ab Initio Electronic-Structure Calculation*, Nat. Comput. Sci. **2**, 367 (2022).

[65] G. Hegde and R. C. Bowen, *Machine-Learned Approximations to Density Functional Theory Hamiltonians*, Sci. Rep. **7**, 42669 (2017).

[66] G. S. Ho, V. L. Lignères, and E. A. Carter, *Introducing PROFESS: A New Program for Orbital-Free Density Functional Theory Calculations*, Comput. Phys. Commun. **179**, 839 (2008).

[67] K. Ryczko, S. J. Wetzel, R. G. Melko, and I. Tamblyn, *Toward Orbital-Free Density Functional Theory with Small Data Sets and Deep Learning*, J. Chem. Theory Comput. **18**, 1122 (2022).

[68] S. Kumar, X. Jing, J. E. Pask, A. J. Medford, and P. Suryanarayana, *Kohn–Sham Accuracy from Orbital-Free Density Functional Theory via Δ-Machine Learning*, J. Chem. Phys. **159**, 244106 (2023).

[69] P. Friederich, F. Häse, J. Proppe, and A. Aspuru-Guzik, *Machine-Learned Potentials for next-Generation Matter Simulations*, Nat. Mater. **20**, 750 (2021).

[70] F. Ercolessi and J. B. Adams, *Interatomic Potentials from First-Principles Calculations: The Force-Matching Method*, Europhys. Lett. **26**, 583 (1994).

[71] T. B. Blank, S. D. Brown, A. W. Calhoun, and D. J. Doren, *Neural Network Models of Potential Energy Surfaces*, J. Chem. Phys. **103**, 4129 (1995).

[72] A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons*, Phys Rev Lett **104**, 136403 (2010).

[73] V. L. Deringer, N. Bernstein, G. Csányi, C. Ben Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, and S. R. Elliott, *Origins of Structural and Electronic Transitions in Disordered Silicon*, Nature **589**, 59 (2021).

[74] J. S. Smith, O. Isayev, and A. E. Roitberg, *ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost*, Chem Sci **8**, 3192 (2017).

[75] H. Wang, L. Zhang, J. Han, and W. E, *DeePMD-Kit: A Deep Learning Package for Many-Body Potential Energy Representation and Molecular Dynamics*, Comput. Phys. Commun. **228**, 178 (2018).

[76] D. Lu, H. Wang, M. Chen, L. Lin, R. Car, W. E, W. Jia, and L. Zhang, *86 PFLOPS Deep Potential Molecular Dynamics Simulation of 100 Million Atoms with Ab Initio Accuracy*, Comput. Phys. Commun. **259**, 107624 (2021).

[77] D. Zhang, H. Bi, F.-Z. Dai, W. Jiang, L. Zhang, and H. Wang, *DPA-1: Pretraining of Attention-Based Deep Potential Model for Molecular Simulation*, (2023).

[78] C. M. Handley and P. L. A. Popelier, *Dynamically Polarizable Water Potential Based on Multipole Moments Trained by Machine Learning*, J. Chem. Theory Comput. **5**, 1474 (2009).

[79] S. Houlding, S. Y. Liem, and P. L. A. Popelier, *A Polarizable High-Rank Quantum Topological Electrostatic Potential Developed Using Neural Networks: Molecular Dynamics Simulations on the Hydrogen Fluoride Dimer*, Int. J. Quantum Chem. **107**, 2817 (2007).

[80] M. Gastegger, J. Behler, and P. Marquetand, *Machine Learning Molecular Dynamics for the Simulation of Infrared Spectra*, Chem Sci **8**, 6924 (2017).

[81] L. Zhang, H. Wang, M. C. Muniz, A. Z. Panagiotopoulos, R. Car, and W. E, *A Deep Potential Model with Long-Range Electrostatic Interactions*, J. Chem. Phys. **156**, 124107 (2022).

[82] P. Raccuglia, K. C. Elbert, P. D. F. Adler, C. Falk, M. B. Wenny, A. Mollo, M. Zeller, S. A. Friedler, J. Schrier, and A. J. Norquist, *Machine-Learning-Assisted Materials Discovery Using Failed Experiments*, Nature **533**, 73 (2016).

[83] K. Muraoka, Y. Sada, D. Miyazaki, W. Chaikittisilp, and T. Okubo, *Linking Synthesis and Structure Descriptors from a Large Collection of Synthetic Records of Zeolite Materials*, Nat. Commun. **10**, 4459 (2019).

[84] A. R. Kitahara and E. A. Holm, *Microstructure Cluster Analysis with Transfer Learning and Unsupervised Learning*, Integrating Mater. Manuf. Innov. **7**, 148 (2018).

[85] S. M. Azimi, D. Britz, M. Engstler, M. Fritz, and F. Mücklich, *Advanced Steel Microstructural Classification by Deep Learning Methods*, Sci. Rep. **8**, 2128 (2018).

[86] J. Madsen, P. Liu, J. Kling, J. B. Wagner, T. W. Hansen, O. Winther, and J. Schiøtz, *A Deep Learning Approach to Identify Local Structures in Atomic-Resolution Transmission Electron Microscopy Images*, Adv. Theory Simul. **1**, 1800037 (2018).

[87] D. Franke, C. M. Jeffries, and D. I. Svergun, *Machine Learning Methods for X-Ray Scattering Data Analysis from Biomacromolecular Solutions*, Biophys. J. **114**, 2485 (2018).

[88] S. B. Torrisi, M. R. Carbone, B. A. Rohr, J. H. Montoya, Y. Ha, J. Yano, S. K. Suram, and L. Hung, *Random Forest Machine Learning Models for Interpretable X-Ray Absorption near-Edge Structure Spectrum-Property Relationships*, Npj Comput. Mater. **6**, 109 (2020).

[89] Y. Li, Y. Wei, Z. Wang, et al., *Quantitative Three-Dimensional Imaging of Chemical Short-Range Order via Machine Learning Enhanced Atom Probe Tomography*, Nat. Commun. **14**, 7410 (2023).

[90] S. Madireddy, D.-W. Chung, T. Loeffler, S. K. R. S. Sankaranarayanan, D. N. Seidman, P. Balaprakash, and O. Heinonen, *Phase Segmentation in Atom-Probe Tomography Using Deep Learning-Based Edge Detection*, Sci. Rep. **9**, 20140 (2019).

[91] X. Zhou, Y. Wei, M. Kühbach, H. Zhao, F. Vogel, R. Darvishi Kamachali, G. B. Thompson, D. Raabe, and B. Gault, *Revealing In-Plane Grain Boundary Composition Features through Machine Learning from Atom Probe Tomography Data*, Acta Mater. **226**, 117633 (2022).

[92] B. Burger, P. M. Maffettone, V. V. Gusev, et al., *A Mobile Robotic Chemist*, Nature **583**, 237 (2020).

[93] P. S. Gromski, J. M. Granda, and L. Cronin, *Universal Chemical Synthesis and Discovery with 'The Chemputer,'* Trends Chem. **2**, 4 (2020).

[94] D. A. Boiko, R. MacKnight, B. Kline, and G. Gomes, *Autonomous Chemical Research with Large Language Models*, Nature **624**, 570 (2023).

[95] D. E. Scaife, P. F. Weller, and W. G. Fisher, *Crystal Preparation and Properties of Cesium Tin(II) Trihalides*, J. Solid State Chem. **9**, 308 (1974).

[96] J. Mizusaki, K. Arai, and K. Fueki, *Ionic Conduction of the Perovskite-Type Halides*, Solid State Ion. **11**, 203 (1983).

[97] R. E. Wasylishen, O. Knop, and J. B. Macdonald, *Cation Rotation in Methylammonium Lead Halides*, Solid State Commun. **56**, 581 (1985).

[98] M. Wuttig, C.-F. Schön, M. Schumacher, J. Robertson, P. Golub, E. Bousquet, C. Gatti, and J.-Y. Raty, *Halide Perovskites: Advanced Photovoltaic Materials*

Empowered by a Unique Bonding Mechanism, Adv. Funct. Mater. **32**, 2110166 (2022).

[99] Q. A. Akkerman and L. Manna, *What Defines a Halide Perovskite?*, ACS Energy Lett. **5**, 604 (2020).

[100] R. H. Mitchell, M. D. Welch, and A. R. Chakhmouradian, *Nomenclature of the Perovskite Supergroup: A Hierarchical System of Classification Based on Crystal Structure and Composition*, Mineral. Mag. **81**, 411 (2017).

[101] V. M. Goldschmidt, *Die Gesetze Der Krystallochemie*, Naturwissenschaften **14**, 477 (1926).

[102] Z. Li, M. Yang, J.-S. Park, S.-H. Wei, J. J. Berry, and K. Zhu, *Stabilizing Perovskite Structures by Tuning Tolerance Factor: Formation of Formamidinium and Cesium Lead Iodide Solid-State Alloys*, Chem. Mater. **28**, 284 (2016).

[103] C. J. Bartel, C. Sutton, B. R. Goldsmith, R. Ouyang, C. B. Musgrave, L. M. Ghiringhelli, and M. Scheffler, *New Tolerance Factor to Predict the Stability of Perovskite Oxides and Halides*, Sci. Adv. **5**, eaav0693 (n.d.).

[104] C. Wang, D. Qu, B. Zhou, C. Shang, X. Zhang, Y. Tu, and W. Huang, *Self-Healing Behavior of the Metal Halide Perovskites and Photovoltaics*, Small **20**, 2307645 (2024).

[105] W. Peng, C. Aranda, O. M. Bakr, G. Garcia-Belmonte, J. Bisquert, and A. Guerrero, *Quantification of Ionic Diffusion in Lead Halide Perovskite Single Crystals*, ACS Energy Lett. **3**, 1477 (2018).

[106] P. V. Kamat and M. Kuno, *Halide Ion Migration in Perovskite Nanocrystals and Nanostructures*, Acc. Chem. Res. **54**, 520 (2021).

[107] C.-X. Zhang, T. Shen, D. Guo, L.-M. Tang, K. Yang, and H.-X. Deng, *Reviewing and Understanding the Stability Mechanism of Halide Perovskite Solar Cells*, InfoMat **2**, 1034 (2020).

[108] Y. Yuan and J. Huang, *Ion Migration in Organometal Trihalide Perovskite and Its Impact on Photovoltaic Efficiency and Stability*, Acc. Chem. Res. **49**, 286 (2016).

[109] H. Zai, Y. Ma, Q. Chen, and H. Zhou, *Ion Migration in Halide Perovskite Solar Cells: Mechanism, Characterization, Impact and Suppression*, J. Energy Chem. **63**, 528 (2021).

[110] H. Jónsson, G. Mills, and K. W. Jacobsen, *Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions*, in *Classical and Quantum Dynamics in Condensed Phase Simulations* (Proceedings of the International School of Physics, Villa Marigola, 1998), pp. 385–404.

[111] B. Focassio, M. Domina, U. Patil, A. Fazzio, and S. Sanvito, *Linear Jacobi-*

Legendre Expansion of the Charge Density for Machine Learning-Accelerated Electronic Structure Calculations, Npj Comput. Mater. **9**, 87 (2023).

[112] T. Koker, K. Quigley, E. Taw, K. Tibbetts, and L. Li, *Higher-Order Equivariant Neural Networks for Charge Density Prediction in Materials*.

[113] R. Hong, Y.-X. Xiao, J. Hu, A.-C. Ji, and S.-J. Ran, *Functional Tensor Network Solving Many-Body Schrödinger Equation*, Phys Rev B **105**, 165116 (2022).

[114] R. O. Jones, *Density Functional Theory: Its Origins, Rise to Prominence, and Future*, Rev Mod Phys **87**, 897 (2015).

[115] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics*, 3rd ed. (Cambridge University Press, Cambridge, 2018).

[116] X. Gao and L.-M. Duan, *Efficient Representation of Quantum Many-Body States with Deep Neural Networks*, Nat. Commun. **8**, 662 (2017).

[117] M. Born and R. Oppenheimer, *Zur Quantentheorie Der Molekeln*, Ann. Phys. **389**, 457 (1927).

[118] A. E. Feiguin, *Phys 5870: Modern Computational Methods in Solids*, Northeast. Univ. (2012).

[119] P. Hohenberg and W. Kohn, *Inhomogeneous Electron Gas*, Phys Rev **136**, B864 (1964).

[120] M. Levy, *Universal Variational Functionals of Electron Densities, First-Order Density Matrices, and Natural Spin-Orbitals and Solution of the* v*-Representability Problem*, Proc. Natl. Acad. Sci. **76**, 6062 (1979).

[121] *The Hohenberg-Kohn Theorems*, in *A Chemist's Guide to Density Functional Theory* (John Wiley & Sons, Ltd, 2001), pp. 33–40.

[122] J. P. Perdew, K. Burke, and M. Ernzerhof, *Generalized Gradient Approximation Made Simple*, Phys. Rev. Lett. **77**, 3865 (1996).

[123] J. P. Perdew, J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais, *Atoms, Molecules, Solids, and Surfaces: Applications of the Generalized Gradient Approximation for Exchange and Correlation*, Phys Rev B **46**, 6671 (1992).

[124] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew, and J. Sun, *Construction of Meta-GGA Functionals through Restoration of Exact Constraint Adherence to Regularized SCAN Functionals*, J. Chem. Phys. **156**, 034109 (2022).

[125] D. Mejía-Rodríguez and S. B. Trickey, *Meta-GGA Performance in Solids at Almost GGA Cost*, Phys Rev B **102**, 121109 (2020).

[126] J. H. Yang, D. A. Kitchaev, and G. Ceder, *Rationalizing Accurate Structure Prediction in the Meta-GGA SCAN Functional*, Phys Rev B **100**, 035132 (2019).

[127] G. Ciccotti, C. Dellago, M. Ferrario, E. R. Hernández, and M. E. Tuckerman, *Molecular Simulations: Past, Present, and Future (a Topical Issue in EPJB)*, Eur. Phys. J. B **95**, 3 (2022).

[128] E. Fermi, P. Pasta, S. Ulam, and M. Tsingou, STUDIES OF THE NONLINEAR PROBLEMS, 1955.

[129] J. B. Gibson, A. N. Goland, M. Milgram, and G. H. Vineyard, *Dynamics of Radiation Damage*, Phys. Rev. **120**, 1229 (1960).

[130] A. Rahman, *Correlations in the Motion of Atoms in Liquid Argon*, Phys. Rev. **136**, A405 (1964).

[131] J. A. McCammon, B. R. Gelin, and M. Karplus, *Dynamics of Folded Proteins*, Nature **267**, 585 (1977).

[132] L. Ya-zhou, L. Yun, S. Shuo, S. Yan-yu, H. Sheng-peng, S. Xiao-guo, G. Ning, and L. Wei-min, *Molecular Dynamics Simulation of Phase Transition and Crack Propagation in Metastable High Entropy Alloy*, Mater. Today Commun. **33**, 104642 (2022).

[133] J. R. Gissinger, B. D. Jensen, and K. E. Wise, *Modeling Chemical Reactions in Classical Molecular Dynamics Simulations*, Polymer **128**, 211 (2017).

[134] S. V. S. Martin H. Müser and L. Pastewka, *Interatomic Potentials: Achievements and Challenges*, Adv. Phys. X **8**, 2093129 (2023).

[135] J. E. Jones and S. Chapman, *On the Determination of Molecular Fields. —II. From the Equation of State of a Gas*, Proc. R. Soc. Lond. Ser. Contain. Pap. Math. Phys. Character **106**, 463 (1997).

[136] J. P. Wagner and P. R. Schreiner, *London Dispersion in Molecular Chemistry—Reconsidering Steric Effects*, Angew. Chem. Int. Ed. **54**, 12274 (2015).

[137] R. A. Buckingham and J. E. Lennard-Jones, *The Classical Equation of State of Gaseous Helium, Neon and Argon*, Proc. R. Soc. Lond. Ser. Math. Phys. Sci. **168**, 264 (1997).

[138] M. Matsui, *Molecular Dynamics Study of MgSiO3 Perovskite*, Phys. Chem. Miner. **16**, 234 (1988).

[139] B. Mortazavi, X. Zhuang, T. Rabczuk, and A. V. Shapeev, *Atomistic Modeling of the Mechanical Properties: The Rise of Machine Learning Interatomic Potentials*, Mater Horiz **10**, 1956 (2023).

[140] A. P. Bartók, R. Kondor, and G. Csányi, *On Representing Chemical Environments*,

Phys Rev B **87**, 184115 (2013).

[141] J. W. Gibbs, editor , *ELEMENTARY PRINCIPLES IN STATISTICAL MECHANICS*, in *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics* (Cambridge University Press, Cambridge, 2010), pp. 1–2.

[142] P. H. Hünenberger, *Thermostat Algorithms for Molecular Dynamics Simulations*, in *Advanced Computer Simulation: Approaches for Soft Matter Sciences I*, edited by C. Dr. Holm and K. Prof. Dr. Kremer (Springer Berlin Heidelberg, Berlin, Heidelberg, 2005), pp. 105–149.

[143] W. G. Hoover, A. J. C. Ladd, and B. Moran, *High-Strain-Rate Plastic Flow Studied via Nonequilibrium Molecular Dynamics*, Phys Rev Lett **48**, 1818 (1982).

[144] D. J. Evans, *Computer '"Experiment"' for Nonlinear Thermodynamics of Couette Flow*, J. Chem. Phys. **78**, 3297 (1983).

[145] H. C. Andersen, *Molecular Dynamics Simulations at Constant Pressure and/or Temperature*, J. Chem. Phys. **72**, 2384 (1980).

[146] W. Shinoda, M. Shiga, and M. Mikami, *Rapid Estimation of Elastic Constants by Molecular Dynamics Simulation under Constant Stress*, Phys Rev B **69**, 134103 (2004).

[147] G. J. Martyna, D. J. Tobias, and M. L. Klein, *Constant Pressure Molecular Dynamics Algorithms*, J. Chem. Phys. **101**, 4177 (1994).

[148] M. Parrinello and A. Rahman, *Polymorphic Transitions in Single Crystals: A New Molecular Dynamics Method*, J. Appl. Phys. **52**, 7182 (1981).

[149] M. Parrinello and A. Rahman, *Crystal Structure and Pair Potentials: A Molecular-Dynamics Study*, Phys Rev Lett **45**, 1196 (1980).

[150] G. KALIBAEVA, M. FERRARIO, and G. CICCOTTI, *Constant Pressure-Constant Temperature Molecular Dynamics: A Correct Constrained NPT Ensemble Using the Molecular Virial*, Mol. Phys. **101**, 765 (2003).

[151] R. Martoıfmmode \checkn\else ň\fiák, A. Laio, and M. Parrinello, *Predicting Crystal Structures: The Parrinello-Rahman Method Revisited*, Phys Rev Lett **90**, 075503 (2003).

[152] N. J. English, *Massively Parallel Molecular-Dynamics Simulation of Ice Crystallisation and Melting: The Roles of System Size, Ensemble, and Electrostatics*, J. Chem. Phys. **141**, 234501 (2014).

[153] S. C. Chowdhury, B. Z. (Gama) Haque, and J. W. Gillespie, *Molecular Dynamics Simulations of the Structure and Mechanical Properties of Silica Glass Using ReaxFF*, J. Mater. Sci. **51**, 10139 (2016).

[154] J. R. Ray and H. Zhang, *Correct Microcanonical Ensemble in Molecular Dynamics*, Phys Rev E **59**, 4781 (1999).

[155] T. Kraska, *Molecular-Dynamics Simulation of Argon Nucleation from Supersaturated Vapor in the NVE Ensemble*, J. Chem. Phys. **124**, 054507 (2006).

[156] T. Developers, *TensorFlow*, (2024).

[157] J. Zeng, D. Zhang, D. Lu, et al., *DeePMD-Kit v2: A Software Package for Deep Potential Models*, J. Chem. Phys. **159**, 054801 (2023).

[158] K. He, X. Zhang, S. Ren, and J. Sun, *Identity Mappings in Deep Residual Networks*, in *Computer Vision – ECCV 2016*, edited by B. Leibe, J. Matas, N. Sebe, and M. Welling (Springer International Publishing, Cham, 2016), pp. 630–645.

[159] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, *Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics*, Phys. Rev. Lett. **120**, 143001 (2018).

[160] L. Zhang, J. Han, H. Wang, W. A. Saidi, R. Car, and E. Weinan, *End-to-End Symmetry Preserving Inter-Atomic Potential Energy Model for Finite and Extended Systems*, in *Proceedings of the 32nd International Conference on Neural Information Processing Systems* (Curran Associates Inc., Red Hook, NY, USA, 2018), pp. 4441–4451.

[161] X. Wang, Y. Wang, L. Zhang, F. Dai, and H. Wang, *A Tungsten Deep Neural-Network Potential for Simulating Mechanical Property Degradation under Fusion Service Environment*, Nucl. Fusion **62**, 126013 (2022).

[162] Y. Zhang, H. Wang, W. Chen, J. Zeng, L. Zhang, H. Wang, and W. E, *DP-GEN: A Concurrent Learning Platform for the Generation of Reliable Deep Learning Based Potential Energy Models*, Comput. Phys. Commun. **253**, 107206 (2020).

[163] J. Siekmann, S. Ravishankar, and T. Kirchartz, *Apparent Defect Densities in Halide Perovskite Thin Films and Single Crystals*, ACS Energy Lett **6**, 3244 (2021).

[164] M. De Keersmaecker, N. R. Armstrong, and E. L. Ratcliff, *Defect Quantification in Metal Halide Perovskites: The Solid-State Electrochemical Alternative*, Energy Env. Sci **14**, 4840 (2021).

[165] S. G. Motti, D. Meggiolaro, S. Martani, R. Sorrentino, A. J. Barker, F. De Angelis, and A. Petrozza, *Defect Activity in Lead Halide Perovskites*, Adv Mater **31**, 1901183 (2019).

[166] S. P. Senanayak, M. Abdi-Jalebi, V. S. Kamboj, et al., *A General Approach for Hysteresis-Free, Operationally Stable Metal Halide Perovskite Field-Effect Transistors*, Sci. Adv. **6**, eaaz4948 (n.d.).

[167] J. M. Azpiroz, E. Mosconi, J. Bisquert, and F. De Angelis, *Defect Migration in*

Methylammonium Lead Iodide and Its Role in Perovskite Solar Cell Operation, Energy Env. Sci **8**, 2118 (2015).

[168] H. Yu, H. Lu, F. Xie, S. Zhou, and N. Zhao, *Native Defect-Induced Hysteresis Behavior in Organolead Iodide Perovskite Solar Cells*, Adv Funct Mater **26**, 1411 (2016).

[169] H. Zhu, A. Liu, K. I. Shim, et al., *High-Performance Hysteresis-Free Perovskite Transistors through Anion Engineering*, Nat. Commun. **13**, 1741 (2022).

[170] S. Kosar, A. J. Winchester, T. A. S. Doherty, et al., *Unraveling the Varied Nature and Roles of Defects in Hybrid Halide Perovskites with Time-Resolved Photoemission Electron Microscopy*, Energy Env. Sci **14**, 6320 (2021).

[171] J. Choi, S. Park, J. Lee, et al., *Organolead Halide Perovskites for Low Operating Voltage Multilevel Resistive Switching*, Adv. Mater. **28**, 6562 (2016).

[172] R. Gao, A. C. P. Jain, S. Pandya, et al., *Designing Optimal Perovskite Structure for High Ionic Conduction*, Adv. Mater. **32**, 1905178 (2020).

[173] C. Li, A. Guerrero, S. Huettner, and J. Bisquert, *Unravelling the Role of Vacancies in Lead Halide Perovskite through Electrical Switching of Photoluminescence*, Nat. Commun. **9**, 5113 (2018).

[174] C. Eames, J. M. Frost, P. R. Barnes, B. C. O'regan, A. Walsh, and M. S. Islam, *Ionic Transport in Hybrid Lead Iodide Perovskite Solar Cells*, Nat. Commun. **6**, 7497 (2015).

[175] P. Delugas, C. Caddeo, A. Filippetti, and A. Mattoni, *Thermally Activated Point Defect Diffusion in Methylammonium Lead Trihalide: Anisotropic and Ultrahigh Mobility of Iodine*, J Phys Chem Lett **7**, 2356 (2016).

[176] S. Tammireddy, S. Reichert, Q. An, A. D. Taylor, R. Ji, F. Paulus, Y. Vaynzof, and C. Deibel, *Temperature-Dependent Ionic Conductivity and Properties of Iodine-Related Defects in Metal Halide Perovskites*, ACS Energy Lett **7**, 310 (2022).

[177] M. Lai, A. Obliger, D. Lu, et al., *Intrinsic Anion Diffusivity in Lead Halide Perovskites Is Facilitated by a Soft Lattice*, Proc Natl Acad Sci **115**, 11929 (2018).

[178] A. Walsh, D. O. Scanlon, S. Chen, X. G. Gong, and S.-H. Wei, *Self-Regulation Mechanism for Charged Point Defects in Hybrid Halide Perovskites*, Angew. Chem. Int. Ed. **54**, 1791 (2015).

[179] M. J. Akhtar, Z.-U.-N. Akhtar, R. A. Jackson, and C. R. A. Catlow, *Computer Simulation Studies of Strontium Titanate*, J Am Ceram Soc **78**, 421 (1995).

[180] J. Crawford and P. Jacobs, *Point Defect Energies for Strontium Titanate: A Pair-Potentials Study*, J Solid State Chem **144**, 423 (1999).

[181] E. Mosconi, D. Meggiolaro, H. J. Snaith, S. D. Stranks, and F. De Angelis, *Light-Induced Annihilation of Frenkel Defects in Organo-Lead Halide Perovskites*, Energy Env. Sci **9**, 3180 (2016).

[182] D. Barboni and R. A. De Souza, *The Thermodynamics and Kinetics of Iodine Vacancies in the Hybrid Perovskite Methylammonium Lead Iodide*, Energy Env. Sci **11**, 3266 (2018).

[183] R. A. De Souza and D. Barboni, *Iodide-Ion Conduction in Methylammonium Lead Iodide Perovskite: Some Extraordinary Aspects*, Chem Commun **55**, 1108 (2019).

[184] M. Pols, V. Brouwers, S. Calero, and S. Tao, *How Fast Do Defects Migrate in Halide Perovskites: Insights from on-the-Fly Machine-Learned Force Fields*, Chem. Commun. **59**, 4660 (2023).

[185] D. Meggiolaro, S. G. Motti, E. Mosconi, A. J. Barker, J. Ball, C. Andrea Riccardo Perini, F. Deschler, A. Petrozza, and F. De Angelis, *Iodine Chemistry Determines the Defect Tolerance of Lead-Halide Perovskites*, Energy Env. Sci **11**, 702 (2018).

[186] A. Laio and M. Parrinello, *Escaping Free-Energy Minima*, Proc Natl Acad Sci **99**, 12562 (2002).

[187] N. Aihemaiti, Y. Jiang, Y. Zhu, and S. Peng, *Light-Induced Phase Segregation Evolution of All-Inorganic Mixed Halide Perovskites*, J. Phys. Chem. Lett. **14**, 267 (2023).

[188] W. Kaiser, K. Hussain, A. Singh, A. A. Alothman, D. Meggiolaro, A. Gagliardi, E. Mosconi, and F. De Angelis, *Defect Formation and Healing at Grain Boundaries in Lead-Halide Perovskites*, J Mater Chem A **10**, 24854 (2022).

[189] E. Maras, O. Trushin, A. Stukowski, T. Ala-Nissila, and H. Jónsson, *Global Transition Path Search for Dislocation Formation in Ge on Si(001)*, Comput Phys Commun **205**, 13 (2016).

[190] Z. Yao, W. Zhao, and S. (Frank) Liu, *Stability of the CsPbI3 Perovskite: From Fundamentals to Improvements*, J Mater Chem A **9**, 11124 (2021).

[191] R. A. Evarestov, A. Senocrate, E. A. Kotomin, and J. Maier, *First-Principles Calculations of Iodine-Related Point Defects in CsPbI 3*, Phys Chem Chem Phys **21**, 7841 (2019).

[192] J. He and S. B. Sinnott, *Ab Initio Calculations of Intrinsic Defects in Rutile TiO2*, J Am Ceram Soc **88**, 737 (2005).

[193] G. Pacchioni and G. Ieranò, *Ab Initio Formation Energies of Point Defects in Pure and Ge-Doped SiO 2*, Phys Rev B **56**, 7304 (1997).

[194] F. P. Sabino, G. M. Dalpian, and A. Zunger, *Light-Induced Frenkel Defect Pair Formation Can Lead to Phase-Segregation of Otherwise Miscible Halide Perovskite*

Alloys, Adv. Energy Mater. **13**, 2301539 (2023).

[195] Y. Cheng, X. Liu, Z. Guan, M. Li, Z. Zeng, H.-W. Li, S.-W. Tsang, A. G. Aberle, and F. Lin, *Revealing the Degradation and Self-Healing Mechanisms in Perovskite Solar Cells by Sub-Bandgap External Quantum Efficiency Spectroscopy*, Adv Mater **33**, 2006170 (2021).

[196] P. M. Piaggi and M. Parrinello, *Calculation of Phase Diagrams in the Multithermal-Multibaric Ensemble*, J Chem Phys **150**, 244119 (2019).

[197] S. Maheshwari, M. B. Fridriksson, S. Seal, J. Meyer, and F. C. Grozema, *The Relation between Rotational Dynamics of the Organic Cation and Phase Transitions in Hybrid Halide Perovskites*, J. Phys. Chem. C **123**, 14652 (2019).

[198] A. Peter Amalathas, L. Landová, Z. Hájková, L. Horák, M. Ledinsky, and J. Holovský, *Controlled Growth of Large Grains in CH3NH3PbI3 Perovskite Films Mediated by an Intermediate Liquid Phase without an Antisolvent for Efficient Solar Cells*, ACS Appl. Energy Mater. **3**, 12484 (2020).

[199] T.-B. Song, Z. Yuan, F. Babbe, D. P. Nenon, E. Aydin, S. De Wolf, and C. M. Sutter-Fella, *Dynamics of Antisolvent Processed Hybrid Metal Halide Perovskites Studied by In Situ Photoluminescence and Its Influence on Optoelectronic Properties*, ACS Appl. Energy Mater. **3**, 2386 (2020).

[200] N. Phung, A. Mattoni, J. A. Smith, et al., *Photoprotection in Metal Halide Perovskites by Ionic Defect Formation*, Joule **6**, 2152 (2022).

[201] D. Meggiolaro, E. Mosconi, and F. De Angelis, *Formation of Surface Defects Dominates Ion Migration in Lead-Halide Perovskites*, ACS Energy Lett. **4**, 779 (2019).

[202] N. Phung, A. Al-Ashouri, S. Meloni, A. Mattoni, S. Albrecht, E. L. Unger, A. Merdasa, and A. Abate, *The Role of Grain Boundaries on Ionic Defect Migration in Metal Halide Perovskites*, Adv. Energy Mater. **10**, 1903735 (2020).

[203] A. Mattoni and S. Meloni, *Defect Dynamics in MAPbI3 Polycrystalline Films: The Trapping Effect of Grain Boundaries*, Helv. Chim. Acta **103**, e2000110 (2020).

[204] M. Pols, T. Hilpert, I. A. W. Filot, A. C. T. van Duin, S. Calero, and S. Tao, *What Happens at Surfaces and Grain Boundaries of Halide Perovskites: Insights from Reactive Molecular Dynamics Simulations of CsPbI3*, ACS Appl. Mater. Interfaces **14**, 40841 (2022).

[205] G. Kresse and J. Furthmüller, *Efficient Iterative Schemes for Ab Initio Total-Energy Calculations Using a Plane-Wave Basis Set*, Phys Rev B **54**, 11169 (1996).

[206] G. Kresse and D. Joubert, *From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method*, Phys Rev B **59**, 1758 (1999).

[207] J. G. Brandenburg, J. E. Bates, J. Sun, and J. P. Perdew, *Benchmark Tests of a Strongly Constrained Semilocal Functional with a Long-Range Dispersion Correction*, Phys Rev B **94**, 115144 (2016).

[208] M. Pols, J. M. Vicent-Luna, I. Filot, A. C. T. van Duin, and S. Tao, *Atomistic Insights Into the Degradation of Inorganic Halide Perovskite CsPbI3: A Reactive Force Field Molecular Dynamics Study*, J. Phys. Chem. Lett. **12**, 5519 (2021).

[209] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *Hybrid Functionals Based on a Screened Coulomb Potential*, J. Chem. Phys. **118**, 8207 (2003).

[210] S. Nosé, *A Unified Formulation of the Constant Temperature Molecular Dynamics Methods*, J. Chem. Phys. **81**, 511 (1984).

[211] W. G. Hoover, *Canonical Dynamics: Equilibrium Phase-Space Distributions*, Phys Rev A **31**, 1695 (1985).

[212] T. Wang, Y. Zhang, W. Kong, et al., *Transporting Holes Stably under Iodide Invasion in Efficient Perovskite Solar Cells*, Science **377**, 1227 (2022).

[213] D. Lu, W. Jiang, Y. Chen, L. Zhang, W. Jia, H. Wang, and M. Chen, *DP Compress: A Model Compression Scheme for Generating Efficient Deep Potential Models*, J. Chem. Theory Comput. **18**, 5559 (2022).

[214] A. P. Thompson, H. M. Aktulga, R. Berger, et al., *LAMMPS - a Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales*, Comput. Phys. Commun. **271**, 108171 (2022).

[215] G. A. Tribello, M. Bonomi, D. Branduardi, C. Camilloni, and G. Bussi, *PLUMED 2: New Feathers for an Old Bird*, Comput. Phys. Commun. **185**, 604 (2014).

[216] F. Gao, Y. Zhao, X. Zhang, and J. You, *Recent Progresses on Defect Passivation toward Efficient Perovskite Solar Cells*, Adv. Energy Mater. **10**, 1902650 (2020).

[217] D. Yang, B. Zhao, T. Yang, R. Lai, D. Lan, R. H. Friend, and D. Di, *Toward Stable and Efficient Perovskite Light-Emitting Diodes*, Adv. Funct. Mater. **32**, 2109495 (2022).

[218] Y. Dong, Y. Zou, J. Song, X. Song, and H. Zeng, *Recent Progress of Metal Halide Perovskite Photodetectors*, J Mater Chem C **5**, 11369 (2017).

[219] Z. Liang, Y. Zhang, H. Xu, et al., *Homogenizing Out-of-Plane Cation Composition in Perovskite Solar Cells*, Nature **624**, 557 (2023).

[220] J. Park, J. Kim, H.-S. Yun, M. J. Paik, E. Noh, H. J. Mun, M. G. Kim, T. J. Shin, and S. I. Seok, *Controlled Growth of Perovskite Layers with Volatile Alkylammonium Chlorides*, Nature **616**, 724 (2023).

[221] T. A. Berhe, W.-N. Su, C.-H. Chen, C.-J. Pan, J.-H. Cheng, H.-M. Chen, M.-C. Tsai,

L.-Y. Chen, A. A. Dubale, and B.-J. Hwang, *Organometal Halide Perovskite Solar Cells: Degradation and Stability*, Energy Env. Sci **9**, 323 (2016).

[222] D. Zhang, D. Li, Y. Hu, A. Mei, and H. Han, *Degradation Pathways in Perovskite Solar Cells and How to Meet International Standards*, Commun. Mater. **3**, 58 (2022).

[223] J. Liang and Y. B. Qi, *Recent Progress on All-Inorganic Metal Halide Perovskite Solar Cells*, Mater. Today Nano **16**, 100143 (2021).

[224] J. Tian, Q. Xue, Q. Yao, N. Li, C. J. Brabec, and H.-L. Yip, *Inorganic Halide Perovskite Solar Cells: Progress and Challenges*, Adv. Energy Mater. **10**, 2000183 (2020).

[225] Y. Liu, N. Borodinov, L. Collins, M. Ahmadi, S. V. Kalinin, O. S. Ovchinnikova, and A. V. Ievlev, *Role of Decomposition Product Ions in Hysteretic Behavior of Metal Halide Perovskite*, ACS Nano **15**, 9017 (2021).

[226] C.-J. Tong, L. Li, L.-M. Liu, and O. V. Prezhdo, *Synergy between Ion Migration and Charge Carrier Recombination in Metal-Halide Perovskites*, J. Am. Chem. Soc. **142**, 3060 (2020).

[227] J. T. DuBose and P. V. Kamat, *Hole Trapping in Halide Perovskites Induces Phase Segregation*, Acc. Mater. Res. **3**, 761 (2022).

[228] H. Zhang, X. Fu, Y. Tang, H. Wang, C. Zhang, W. W. Yu, X. Wang, Y. Zhang, and M. Xiao, *Phase Segregation Due to Ion Migration in All-Inorganic Mixed-Halide Perovskite Nanocrystals*, Nat. Commun. **10**, 1088 (2019).

[229] S. R. G. Balestra, J. M. Vicent-Luna, S. Calero, S. Tao, and J. A. Anta, *Efficient Modelling of Ion Structure and Dynamics in Inorganic Metal Halide Perovskites*, J Mater Chem A **8**, 11824 (2020).

[230] J.-H. Yang, W.-J. Yin, J.-S. Park, and S.-H. Wei, *Fast Self-Diffusion of Ions in CH3NH3PbI3: The Interstiticaly Mechanism versus Vacancy-Assisted Mechanism*, J Mater Chem A **4**, 13105 (2016).

[231] W.-J. Yin, T. Shi, and Y. Yan, *Unusual Defect Physics in CH3NH3PbI3 Perovskite Solar Cell Absorber*, Appl. Phys. Lett. **104**, 063903 (2014).

[232] G. Liu, Y. Wu, Y. Liu, B. Cai, Y. Dong, S. Zhang, and H. Zeng, *Halide Ion Migration in Lead-Free All-Inorganic Cesium Tin Perovskites*, Appl. Phys. Lett. **119**, 031902 (2021).

[233] U.-G. Jong, C.-J. Yu, G.-C. Ri, A. P. McMahon, N. M. Harrison, P. R. F. Barnes, and A. Walsh, *Influence of Water Intercalation and Hydration on Chemical Decomposition and Ion Transport in Methylammonium Lead Halide Perovskites*, J Mater Chem A **6**, 1067 (2018).

[234] J. Haruyama, K. Sodeyama, L. Han, and Y. Tateyama, *First-Principles Study of Ion*

Diffusion in Perovskite Solar Cell Sensitizers, J. Am. Chem. Soc. **137**, 10048 (2015).

[235] C. C. Battaile, *The Kinetic Monte Carlo Method: Foundation, Implementation, and Application*, Recent Adv. Comput. Study Nanostructures **197**, 3386 (2008).

[236] T. Zhang, C. Hu, and S. Yang, *Ion Migration: A "Double-Edged Sword" for Halide-Perovskite-Based Electronic Devices*, Small Methods **4**, 1900552 (2020).

[237] M. H. Futscher, J. M. Lee, L. McGovern, L. A. Muscarella, T. Wang, M. I. Haider, A. Fakharuddin, L. Schmidt-Mende, and B. Ehrler, *Quantification of Ion Migration in CH3NH3PbI3 Perovskite Solar Cells by Transient Capacitance Measurements*, Mater Horiz **6**, 1497 (2019).

[238] Y. Xiao, C. Xue, X. Wang, Y. Liu, Z. Yang, and S. Liu, *Bulk Heterostructure BA2PbI4/MAPbI3 Perovskites for Suppressed Ion Migration To Achieve Sensitive X-Ray Detection Performance*, ACS Appl. Mater. Interfaces **14**, 54867 (2022).

[239] S. Meloni, T. Moehl, W. Tress, et al., *Ionic Polarization-Induced Current–Voltage Hysteresis in CH3NH3PbX3 Perovskite Solar Cells*, Nat. Commun. **7**, 10334 (2016).

[240] T. Zhang, H. Chen, Y. Bai, S. Xiao, L. Zhu, C. Hu, Q. Xue, and S. Yang, *Understanding the Relationship between Ion Migration and the Anomalous Hysteresis in High-Efficiency Perovskite Solar Cells: A Fresh Perspective from Halide Substitution*, Nano Energy **26**, 620 (2016).

[241] A. A. Cheenady and K. Rajan, *Strain-Induced Anisotropic Ion Migration in Single-Crystal Cesium Lead Halide Perovskites*, J. Appl. Phys. **133**, 205105 (2023).

[242] A. Pertsinidis and X. S. Ling, *Diffusion of Point Defects in Two-Dimensional Colloidal Crystals*, Nature **413**, 147 (2001).

[243] O. C. Ibe, *Two-Dimensional Random Walk.*, in *Elements of Random Walk and Diffusion Processes* (John Wiley & Sons, Ltd, 2013), pp. 103–128.

[244] Y. Rakita, I. Lubomirsky, and D. Cahen, *When Defects Become 'Dynamic': Halide Perovskites: A New Window on Materials?*, Mater Horiz **6**, 1297 (2019).

[245] G. Pólya, *Über Eine Aufgabe Der Wahrscheinlichkeitsrechnung Betreffend Die Irrfahrt Im Straßennetz*, Math. Ann. **84**, 149 (1921).

[246] P. Singh, Y. Soffer, D. R. Ceratti, M. Elbaum, D. Oron, G. Hodes, and D. Cahen, *A-Site Cation Dependence of Self-Healing in Polycrystalline APbI3 Perovskite Films*, ACS Energy Lett. **8**, 2447 (2023).

[247] W. Tress, *Metal Halide Perovskites as Mixed Electronic–Ionic Conductors: Challenges and Opportunities—From Hysteresis to Memristivity*, J. Phys. Chem. Lett. **8**, 3106 (2017).

[248] K. Kang, W. Hu, and X. Tang, *Halide Perovskites for Resistive Switching Memory*,

J. Phys. Chem. Lett. **12**, 11673 (2021).

[249] H. Yang, Y. Zhou, Y. Yang, D. Yi, T. Ye, T. D. Lam, D. Golberg, B. Bao, J. Yao, and X. Wang, *Crystal Facet Engineering Induced Anisotropic Transport of Charge Carriers in a Perovskite*, J Mater Chem C **6**, 11707 (2018).

[250] D. Kim, J.-H. Yun, M. Lyu, J. Kim, S. Lim, J. S. Yun, L. Wang, and J. Seidel, *Probing Facet-Dependent Surface Defects in MAPbI3 Perovskite Single Crystals*, J. Phys. Chem. C **123**, 14144 (2019).

[251] S. P. Ong, W. D. Richards, A. Jain, G. Hautier, M. Kocher, S. Cholia, D. Gunter, V. L. Chevrier, K. A. Persson, and G. Ceder, *Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis*, Comput. Mater. Sci. **68**, 314 (2013).

[252] L. Zhang, J. Han, H. Wang, R. Car, and E. Weinan, *Deep Potential Molecular Dynamics: A Scalable Model with the Accuracy of Quantum Mechanics*, Phys Rev Lett **120**, 143001 (2018).

[253] G. Bussi, D. Donadio, and M. Parrinello, *Canonical Sampling through Velocity Rescaling*, J. Chem. Phys. **126**, 014101 (2007).

[254] A. Stukowski, *Visualization and Analysis of Atomistic Simulation Data with OVITO–the Open Visualization Tool*, Model. Simul. Mater. Sci. Eng. **18**, 015012 (2009).

[255] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein, *MDAnalysis: A Toolkit for the Analysis of Molecular Dynamics Simulations*, J. Comput. Chem. **32**, 2319 (2011).

[256] Y. Hu, T.-M. Li, L. Anderson, J. Ragan-Kelley, and F. Durand, *Taichi: A Language for High-Performance Computation on Spatially Sparse Data Structures*, ACM Trans Graph **38**, 1 (2019).

[257] E. H. Lieb and B. Simon, *The Thomas-Fermi Theory of Atoms, Molecules and Solids*, Adv. Math. **23**, 22 (1977).

[258] E. Prodan and W. Kohn, *Nearsightedness of Electronic Matter*, Proc. Natl. Acad. Sci. **102**, 11635 (2005).

[259] A. Grisafi, A. Fabrizio, B. Meyer, D. M. Wilkins, C. Corminboeuf, and M. Ceriotti, *Transferable Machine-Learning Model of the Electron Density*, ACS Cent. Sci. **5**, 57 (2019).

[260] A. Fabrizio, A. Grisafi, B. Meyer, M. Ceriotti, and C. Corminboeuf, *Electron Density Learning of Non-Covalent Systems*, Chem Sci **10**, 9424 (2019).

[261] A. M. Lewis, A. Grisafi, M. Ceriotti, and M. Rossi, *Learning Electron Densities in the Condensed Phase*, J. Chem. Theory Comput. **17**, 7203 (2021).

[262] A. J. Lee, J. A. Rackers, and W. P. Bricker, *Predicting Accurate Ab Initio DNA Electron Densities with Equivariant Neural Networks*, Biophys. J. **121**, 3883 (2022).

[263] K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer, *Unifying Machine Learning and Quantum Chemistry with a Deep Neural Network for Molecular Wavefunctions*, Nat. Commun. **10**, 5024 (2019).

[264] J. A. Rackers, L. Tecot, M. Geiger, and T. E. Smidt, *A Recipe for Cracking the Quantum Scaling Limit with Machine Learned Electron Densities*, Mach. Learn. Sci. Technol. **4**, 015027 (2023).

[265] F. Brockherde, L. Vogt, L. Li, M. E. Tuckerman, K. Burke, and K.-R. Müller, *Bypassing the Kohn-Sham Equations with Machine Learning*, Nat. Commun. **8**, 872 (2017).

[266] A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, *Solving the Electronic Structure Problem with Machine Learning*, Npj Comput. Mater. **5**, 22 (2019).

[267] S. Gong, T. Xie, T. Zhu, S. Wang, E. R. Fadel, Y. Li, and J. C. Grossman, *Predicting Charge Density Distribution of Materials Using a Local-Environment-Based Graph Convolutional Network*, Phys. Rev. B **100**, 184103 (2019).

[268] L. Zepeda-Núñez, Y. Chen, J. Zhang, W. Jia, L. Zhang, and L. Lin, *Deep Density: Circumventing the Kohn-Sham Equations via Symmetry Preserving Neural Networks*, J. Comput. Phys. **443**, 110523 (2021).

[269] D. Kamal, A. Chandrasekaran, R. Batra, and R. Ramprasad, *A Charge Density Prediction Model for Hydrocarbons Using Deep Neural Networks*, Mach. Learn. Sci. Technol. **1**, 025003 (2020).

[270] P. B. Jørgensen and A. Bhowmik, *Equivariant Graph Neural Networks for Fast Electron Density Estimation of Molecules, Liquids, and Solids*, Npj Comput. Mater. **8**, 183 (2022).

[271] A. G. Eguiluz, *Self-Consistent Static-Density-Response Function of a Metal Surface in Density-Functional Theory*, Phys. Rev. B **31**, 3303 (1985).

[272] T. M. Rice, *The Effects of Electron-Electron Interaction on the Properties of Metals*, Ann. Phys. **31**, 100 (1965).

[273] M. Zaheer, S. Kottur, S. Ravanbakhsh, B. Poczos, R. R. Salakhutdinov, and A. J. Smola, *Deep Sets*, in *Advances in Neural Information Processing Systems*, edited by I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Vol. 30 (Curran Associates, Inc., 2017).

[274] J. Sun, A. Ruzsinszky, and J. P. Perdew, *Strongly Constrained and Appropriately Normed Semilocal Density Functional*, Phys. Rev. Lett. **115**, 036402 (2015).

[276] W. Jiang, Y. Zhang, L. Zhang, and H. Wang, *Accurate Deep Potential Model for the Al–Cu–Mg Alloy in the Full Concentration Space**, Chin. Phys. B **30**, 050706 (2021).

[277] C. Zhang, F. Tang, M. Chen, J. Xu, L. Zhang, D. Y. Qiu, J. P. Perdew, M. L. Klein, and X. Wu, *Modeling Liquid Water by Climbing up Jacob's Ladder in Density Functional Theory Facilitated by Using Deep Neural Network Potentials*, J. Phys. Chem. B **125**, 11444 (2021).

[278] D. Zhang, X. Liu, X. Zhang, et al., *DPA-2: Towards a Universal Large Atomic Model for Molecular and Material Simulation*.