A reactive force field for large scale simulations of metal halide perovskites

Pols, Mike C.W.M.

Award date:
2021

A Reactive Force Field for Large Scale Simulations of Metal Halide Perovskites

Eindhoven University of Technology

Department of Applied Physics
Materials Simulation & Modelling (MSM)

Department of Chemical Engineering and Chemistry
Inorganic Materials & Catalysis (IMC)

Master thesis submitted by
Mike C.W.M. Pols (0897735)

handed in on
February 8, 2021

Supervised by
dr. Shuxia Tao
dr. ir. Ivo A.W. Filot

Report ID:
R-0001-M

A Reactive Force Field for Large Scale Simulations of Metal Halide Perovskites

Mike C.W.M. Pols

Abstract: In the past decade, metal halide perovskites have emerged as a promising material class to be used in solar cell technology, thanks to their extraordinary optoelectronic properties and facile low-cost production methods. As a result of extensive research efforts, it took only about 10 years before perovskite solar cells reached an efficiency of 25% in 2020. However, metal halide perovskites suffer from various stability problems, hindering the commercialization of perovskite solar cells. Despite much progress in the characterization of the degradation of metal halide perovskites, a comprehensive understanding of the decomposition mechanisms still lacks, mainly due to the limitations in the spatiotemporal resolution of common experiments.

Computational modelling can be used to obtain atomic-scale insights of these degradation mechanisms. However, up to now, the majority of the computational studies were done using computationally demanding methods based on quantum mechanics, severely limiting the length and time scales of the systems that can be investigated. Molecular dynamics simulations have been successfully applied to simulate larger material systems at longer time scales using classical force fields. Nonetheless, the pre-defined connectivity in these force fields makes them unsuited for the simulation of chemical reactions. Molecular dynamics simulations that use reactive force fields (ReaxFF) are a promising alternative to overcome such limitations and are therefore ideal for a detailed description of the degradation reactions in perovskites.

In this thesis, we present the first efforts towards a description of metal halide perovskites with ReaxFF. For this, we create a set of ReaxFF parameters for the inorganic metal halide perovskite, $CsPbI_3$, which is benchmarked against quantum mechanical calculations and experiments. Using ReaxFF, we carry out reactive molecular dynamics simulations to study the intrinsic dynamical processes to analyze their roles in the phase stability of the perovskites. We find that the anharmonic character of the perovskite lattice and thermal entropy are responsible for the phase transitions between the different perovskite phases. Further analysis reveals that the anharmonic character of the metal halide framework gives rise to preferential positions of Cs cations, the movement away from which is likely the cause for the phase instability of $CsPbI_3$, i.e. a phase transition of the perovskite phase towards the non-perovskite phase. Interestingly, at elevated temperatures, we find that the dynamical disorder of the metal halide framework and the dynamics of the Cs cations stabilize the perovskite structure. Importantly, we also apply our ReaxFF force field study reactive processes, including ion migration and defect-mediated degradation reactions. Our results demonstrate that although both iodine interstitials and iodine vacancies play an important role in the overall migration of ions, the vacancies are mainly responsible for the decomposition of the perovskite. Combining our simulations with existing experimental observations, we suggest a detailed atomistic mechanism for the decomposition reactions that are responsible for the overall degradation of $CsPbI_3$ perovskite to PbI_2 clusters.

Contents

1. Introduction

1.1. Perovskite solar cells

Necessitated by a globally ever increasing demand for energy and the negative environmental impact of conventional fossil fuel energy sources, one of the main challenges of present-day society is searching for a sustainable, low-cost energy source. Solar cells that harvest sunlight and convert it to electricity are one of the technologies that fulfil such requirements and have consequently become a popular source of renewable energy. Following a few decades of extensive research and engineering efforts, crystalline silicon solar cells have seen a large decrease of their production costs and a substantial increase in power conversion efficiency (PCE), which has resulted in their widespread adoption as a renewable energy source [1–3].

In the past decade, metal halide perovskites have emerged as a promising alternative to silicon solar cell technology. Owing to the exceptional optoelectronic properties and the low fabrication costs associated with metal halide perovskite absorber layers [4, 5], perovskite solar cells (PSCs) have been a very active topic of research. From the first perovskite solar cell that only showed a PCE of 3.8% [6], scientific research has managed to push the efficiency of PSCs to over 25% in 2020 [7]. However, despite the attractiveness of their high efficiencies, the long-term stability issues of perovskite solar cells impede their commercialization [8–11].

1.2. Stability of metal halide perovskites

The metal halide perovskites that are used in PSCs have a three-dimensional structure with the AMX_3 chemical formula, where A is a monovalent inorganic or organic cation (Cs^+; methylammonium MA^+ or formamidimium FA^+), M is a divalent metal cation (Pb^{2+} or Sn^{2+}) and X is a monovalent halide anion (I^-; Br^- or Cl^-). Examples of the perovskite structure are shown in Figure 1.1. The metal and halide ions form a network of corner-sharing MX_6 octahedra, where the centre of the cuboids formed by these octahedra is occupied by a large monovalent A cation. The atoms in perovskites are held together through a mix of ionic and relatively weak covalent bonds, as a result of which this class of materials has a soft and dynamical crystal lattice [12–15].

Most of the stability issues of perovskite solar cells can be traced back to the intrinsic instability of the metal halide perovskite absorber layers in PSCs [8–11]. Such instability issues include a

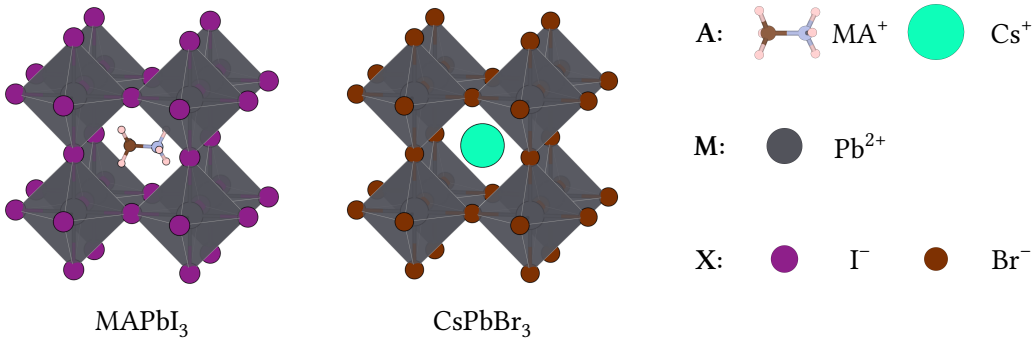

Figure 1.1. A demonstration of the chemical diversity of metal halide perovskites with the AMX_3 structure. (Left): inorganic-organic hybrid $MAPbI_3$. (Right): inorganic $CsPbBr_3$. The same colour scheme for the atomic species is used throughout the entire thesis.

thermal instability, a phase instability and an ionic migration induced instability. To elaborate, some metal halide perovskites exhibit a phase instability where the perovskite structure of the perovskite transforms to a more stable non-perovskite phase, which results in a decreased performance of the PSC due to the worse optoelectronic properties of this non-perovskite phase [16]. Moreover, the relatively facile fabrication of metal halide perovskites using spin coating, typically results in a large number of defects in the perovskite films ($10^{16} - 10^{18}$ cm^{-3}) [17, 18]. Although the easily formed defects are found to have a minimal impact on the optoelectronic properties of the metal halide perovskites [19, 20], the migration and accumulation of the defects, and thus ionic species, is suggested to have a major impact on the long-term stability of the perovskites [21–23]. Besides, the defect induced degradation of halide perovskites is often accelerated by external stimuli, such as a high temperature [24], moisture [25] and oxygen species combined with ultraviolet light [26].

Although much progress has been made to characterize the degradation of halide perovskites and its consequences for the performance of PSCs, a comprehensive understanding of the decomposition mechanisms is still lacking [27]. All above-mentioned physical and chemical processes that cause instability issues are closely linked to the soft and therefore dynamical crystal lattice of the perovskites. The dynamical processes, which include ion diffusion, lattice vibrations and degradation reactions, typically occur at time scales of a couple of femtoseconds up to the picosecond time scale [28]. These processes are often difficult to study in experiments since common experimental techniques are tailored to investigate the macroscopic properties of the materials and devices. The limitations of both the spatial and temporal resolution in experiments pose significant challenges in the fundamental understanding of the relevant degradation mechanisms.

1.3. State-of-the-art computational research and our approach

Computational modelling of materials is complementary to experiments and can be used to obtain atomic-scale insights into the degradation mechanisms of metal halide perovskites. Many

computational investigations of metal halide perovskites have used first-principles methods based on quantum mechanics to investigate metal halide perovskites and their degradation mechanisms [29–31]. However, the large computational expense of these methods only allows for the simulation of short time scales and small system sizes, thus limiting the insights obtained with such methods. Molecular dynamics simulations based on classical force fields have been successfully employed to study metal halide perovskites in large systems and for long durations [32–34]. Nevertheless, the pre-defined connectivity in simulations with classical force fields does not allow for chemical reactions to occur during the simulation, thus making them ill-suited to study the decomposition reactions of the perovskites.

A reactive force field (ReaxFF) is an extension to classical force fields that employs a dynamical bond order based on the interatomic distance of particles to describe the creation and breaking of bonds [35, 36]. ReaxFF simulations can thus describe chemical reactions at length and time scales that are inaccessible to methods based on quantum mechanics. As a result of this, ReaxFF simulations are a valuable tool to understand the degradation mechanisms of metal halide perovskites. However, up to now, no ReaxFF simulations of metal halide perovskites are found because of the unavailability of ReaxFF parameters for this class of materials, mainly because it is still a relatively new class of materials in the scientific community.

This work is the first effort towards a description of metal halide perovskites with ReaxFF. We start by developing a ReaxFF description for one of the most studied and most straightforward metal halide perovskites: $CsPbI_3$. A set of accurate reference data from quantum mechanical calculations is employed to train the ReaxFF parameters required for this description. The ReaxFF force field is then employed in dynamical simulations to investigate the impact of a range of processes on the stability of the perovskite. These simulations will help us answer the central research question of this work:

What roles do the dynamical and reactive processes play in the stability of metal halide perovskites?

1.4. Outline of the thesis

The thesis is structured in the following manner. We start with an explanation of all the computational methods that were used to investigate $CsPbI_3$ with ReaxFF in chapter 2. In this chapter, we first cover the quantum mechanical calculations used to generate accurate reference data and then detail the principles of reactive molecular dynamics simulations with ReaxFF. Next, we outline the basics of the parameterization of a force field and subsequently elaborate on how these basics were applied to create a representative ReaxFF force field for $CsPbI_3$ in chapter 3. Then, in chapter 4, we analyze from reactive molecular dynamics simulations with our ReaxFF force field what processes impact the stability of metal halide perovskites, and in particular $CsPbI_3$. Finally,

we conclude the thesis with a summary of the main findings and provide an outlook on how the work in this thesis can be progressed further in chapter 5.

2. Computational methods

Several computational methods are used to run reactive molecular dynamics simulations for metal halide perovskites in this work. This chapter explains the basics of these computational methods. First, we cover the quantum mechanical calculations, specifically density functional theory, that we used to generate accurate reference data for the ReaxFF force field training. Secondly, the basics of molecular dynamics, the simulation technique that allows us to investigate the dynamical behaviour of the perovskites, are treated. Finally, the details of ReaxFF, the potential energy expression that allows for reactive events to take place, are explained.

2.1. Quantum mechanical calculations

2.1.1. Electronic structure calculations

Electronic structure calculations aim to find a numerical solution to the non-relativistic time-independent Schrödinger equation[1] [37]. Since the material systems investigated with these methods consist of many electrons and nuclei, finding such a solution is a many-body problem. In the case of a material system that consists of N electrons and M nuclei, the Schrödinger equation of the system is of the following form

$$\hat{H}\Psi\left(\vec{r}_i;\vec{R}_I\right) = E\Psi\left(\vec{r}_i;\vec{R}_I\right), \tag{2.1}$$

where $\hat{H}$ is the Hamiltonian, Ψ the wavefunction and E is the energy of the many-body system. Notice that Ψ depends on the coordinates of all electrons and nuclei which are described by the position vectors $\vec{r}_i$ and $\vec{R}_I$, respectively[2]. In a system with positively charged nuclei and negatively charged electrons, the non-relativistic Hamiltonian can be expressed in atomic units as

$$\hat{H} = \hat{T}_\mathrm{n} + \hat{T}_\mathrm{e} + \hat{V}_\mathrm{n,n} + \hat{V}_\mathrm{e,e} + \hat{V}_\mathrm{n,e}, \tag{2.2}$$

[1] For brevity the *non-relativistic time-independent Schrödinger equation*, is referred to as just the *Schrödinger equation* from this point onwards

[2] Here we used $\vec{r}_i$ as a shorter notation of $\vec{r}_1, \ldots, \vec{r}_N$ and $\vec{R}_i$ as a shorter notation of $\vec{R}_1, \ldots, \vec{R}_M$

where

$$\hat{T}_{\mathrm{n}} = -\frac{1}{2}\sum_{I=1}^{M}\frac{1}{M_{\mathrm{I}}}\nabla_I^2; \qquad \hat{T}_{\mathrm{e}} = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2$$

$$\hat{V}_{\mathrm{n,n}} = \sum_{I=1}^{M}\sum_{J>I}^{M}\frac{Z_I Z_J}{\left|\vec{R}_I - \vec{R}_J\right|}; \qquad \hat{V}_{\mathrm{e,e}} = \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{|\vec{r}_i - \vec{r}_j|}$$

$$\hat{V}_{\mathrm{n,e}} = -\sum_{i=1}^{N}\sum_{I=1}^{M}\frac{Z_I}{\left|\vec{r}_i - \vec{R}_I\right|},$$

in which we use lower and upper case indices to iterate over the N electrons and M nuclei, respectively [37]. ∇^2 is the Laplacian in the respective nuclear or electronic coordinates, M the ratio between the nuclear mass and the electronic mass and Z the nuclear charge. In the above expression, the first two terms are the kinetic energy of the nuclei ($\hat{T}_{\mathrm{n}}$) and electrons ($\hat{T}_{\mathrm{e}}$), respectively. The third and fourth terms are the Coulomb repulsion between the nuclei ($\hat{V}_{\mathrm{n,n}}$) and electrons ($\hat{V}_{\mathrm{e,e}}$), respectively. Finally, the fifth term is the Coulomb attraction between the electrons and nuclei ($\hat{V}_{\mathrm{n,e}}$).

An important approximation that is made to simplify the Schrödinger equation in almost all electronic structure calculations is the Born-Oppenheimer approximation [38]. In the approximation, the system's electronic and nuclear motion is separated based on the large mass disparity between the electrons and nuclei. Due to the large mass of the nuclei relative to the electrons, the nuclei are assumed to be stationary. The electrons are fast-moving through the electrostatic potential determined by these fixed nuclei. The kinetic energy of the nuclei $\hat{T}_{\mathrm{n}}$ can thus be neglected. Furthermore, for a particular configuration of the nuclei, the Coulomb repulsion between these nuclei $\hat{V}_{\mathrm{n,n}}$ is a constant that does not involve the electrons. Altogether, this allows us to write the following expression for the electronic Hamiltonian $\hat{H}_{\mathrm{e}}$

$$\hat{H}_{\mathrm{e}} = \hat{T}_e + \hat{V}_{\mathrm{e,e}} + \hat{V}_{\mathrm{ext}} = \hat{T}_e + \hat{V}_{\mathrm{e,e}} + \sum_{i=1}^{N} v_{\mathrm{ext}}\left(\vec{r}_i\right), \tag{2.3}$$

in which we rewrite the Coulomb attraction between the nuclei and electrons $\hat{V}_{\mathrm{n,e}}$ into a more general external potential $\hat{V}_{\mathrm{ext}}$ that also allows for additional external fields or potentials apart from interaction with the nuclei[3]. Based on this electronic Hamiltonian, we rewrite the Schrödinger equation to an electronic form

$$\hat{H}_{\mathrm{e}}\psi_{\vec{R}_I}\left(\vec{r}_i\right) = E_{\mathrm{e}}\psi_{\vec{R}_I}\left(\vec{r}_i\right), \tag{2.4}$$

where ψ is the many-body electronic wavefunction[4] that explicitly depends on the coordinates

[3] The external potential is split into separate contributions for each atom, the use of which becomes apparent in the construction of the energy functional.

[4] The parameteric dependence on the nuclear coordinates is indicated with the $\vec{R}_I$-subscript. This subscript is dropped in the other places the electronic many-body wavefunction is used.

of the electrons $\vec{r}_i$, but only parameterically on the nuclear coordinates $\vec{R}_I$ and E_e is the energy of the many-body electron system. The solution to this electronic Schrödinger equation gives us insights into the electronic structure of the material and is thus the central topic of the rest of this section. Essentially, the use of the Born-Oppenheimer approximation has reduced the complexity of electronic structure calculations by reducing the degrees of freedom in the Schrödinger equation from $3M + 3N$ to $3N$ degrees of freedom by going from the total many-body Schrödinger equation to the electronic Schrödinger equation. Although some approaches have been devised to solve this many-electron Schrödinger equation, which include Hartree-Fock and more accurate post-Hartree-Fock methods [37], these methods have proven to be too computationally demanding to use in large systems.

2.1.2. Density functional theory

Density functional theory (DFT) solves the electronic Schrödinger equation, but instead of using the many-electron wavefunction as its central quantity, it makes use of the electron density. This electron density $\rho(\vec{r})$ can be calculated from the many-electron wavefunction in a system of N electrons as

$$\rho(\vec{r}) = N \int \cdots \int |\psi(\vec{r}, \vec{r}_2, \ldots, \vec{r}_N)|^2 \, d^3\vec{r}_2 \cdots d^3\vec{r}_N \,, \tag{2.5}$$

where we made use of the fact that the electrons are indistinguishable. It shows that regardless of the number of electrons, the electron density is always three-dimensional. In contrast to the use of the many-electron wavefunction, that for an N electron system depends on $3N$ spatial coordinates, the change to the electron density has significantly decreased the complexity of the problem [39]. As a result of this decrease in complexity, DFT can be used to calculate much larger system sizes, which has caused it to be a widely adopted method for electronic structure calculations.

The theoretical foundation of DFT was laid in 1964 in the work of Hohenberg and Kohn [40]. In their work, they posed two theorems that form the basis of DFT. In the first theorem, they proved a one-to-one correspondence between the electron density and the external potential felt by the electrons. Given a certain electron density $\rho(\vec{r})$ we can thus find a unique external potential $v_\mathrm{ext}(\vec{r})$. Moreover, since the Hamiltonian of the system depends on this external potential, the states and any observables of the many-electrons system, are determined by this electron density. Consequently, we can define the energy functional of the system $E[\rho]$ by making use of the expectation value of the electronic Hamiltonian as

$$\begin{aligned} E[\rho] &= \langle\psi|\hat{H}_\mathrm{e}|\psi\rangle = \langle\psi|\hat{T}_e + \hat{V}_\mathrm{e,e} + \hat{V}_\mathrm{ext}|\psi\rangle \\ &= T_\mathrm{e}[\rho] + V_\mathrm{e,e}[\rho] + \int \rho(\vec{r})\, v_\mathrm{ext}(\vec{r})\, d^3\vec{r} \\ &= F[\rho] + \int \rho(\vec{r})\, v_\mathrm{ext}(\vec{r})\, d^3\vec{r}, \end{aligned} \tag{2.6}$$

where the universal part of the energy functional, $F[\rho]$, that holds for any external potential v_{ext}, is defined as

$$F[\rho] = T_{\text{e}}[\rho] + V_{\text{e,e}}[\rho], \tag{2.7}$$

a summation of the kinetic energy and interaction energy for the electrons. In the second theorem, Hohenberg and Kohn proved that the ground state energy of this many-electron system could be obtained from the energy functional using the variational principle. The minimum energy of this energy functional is the ground state energy, and density that corresponds to this minimum is the exact ground-state electron density. Thus, having found the minimum of the energy functional, we have effectively obtained all information on the ground state of the system. Unfortunately, the two theorems only provide mathematical proof for the existence of—not an expression for—this energy functional, thus rendering the theory inapplicable in practice.

It was not until a year later, in 1965, Kohn and Sham published a paper in which a good approximation for this energy functional $E[\rho]$ was introduced, that method could be applied to real material systems [41]. The central idea behind the approximation is that any system of interacting electrons can be mapped onto a system of non-interacting electrons with a similar electron density. To do so, Kohn and Sham proposed the following decomposition of the universal functional

$$F[\rho] = T_{\text{s}}[\rho] + \frac{1}{2}\int\int \frac{\rho(\vec{r})\,\rho(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,d^3\vec{r}\,d^3\vec{r}\,' + E_{\text{xc}}[\rho], \tag{2.8}$$

where $T_{\text{s}}[\rho]$ is the kinetic energy of a non-interacting system of electrons with an electron density $\rho(\vec{r})$, the second term is the classical Coulomb interaction between the electrons in an electron density $\rho(\vec{r})$ and $E_{\text{xc}}[\rho]$ is the exchange-correlation (XC) energy of a system of electrons. Using this decomposition and subsequently applying the variational principle to a system of interacting electrons and a system of non-interacting electrons, whilst constraining the number of electrons to N, we find that problems are mathematically identical. Provided that the electrons in the non-interacting system experience the following effective potential v_{eff}

$$v_{\text{eff}}(\vec{r}) = v_{\text{ext}}(\vec{r}) + \int \frac{\rho(\vec{r}\,')}{|\vec{r}-\vec{r}\,'|}\,d^3\vec{r}\,' + \frac{\delta E_{\text{xc}}[\rho]}{\delta\rho(\vec{r})}, \tag{2.9}$$

the two systems of electrons have the same ground state electron density. In this effective potential, the first, second and third term represent the external potential, electron Coulomb potential and XC potential, respectively. Thus, to find the the ground state electron density, it suffices to solve for the system of non-interacting electrons in the effective potential v_{eff}, which is done by solving the one-particle Schrödinger equation

$$\left[-\frac{1}{2}\nabla^2 + v_{\text{eff}}(\vec{r})\right]\phi_i(\vec{r}) = \epsilon_i\phi_i(\vec{r}), \tag{2.10}$$

where ϕ_i are the Kohn-Sham orbitals and ϵ_i the Kohn-Sham orbital energies, from which the

electron density can be calculated as

$$\rho\left(\vec{r}\right) = \sum_{i=1}^{N} \left|\phi_i\left(\vec{r}\right)\right|^2. \tag{2.11}$$

Since the effective potential depends on the electron density, these one-particle Schrödinger equations must be solved self-consistently. From an initial guess of the electron density, the set of Kohn-Sham orbitals can be obtained. In turn, these orbitals are used to determine a new electron density. This cycle is repeated until the input and output electron densities are the same, which is when the solutions have converged. Consequently, the total energy of the interacting many-electron system E_{KS} can be determined from the converged electron density $\rho\left(\vec{r}\right)$ and the Kohn-Sham orbital energies ϵ_i as

$$E_{\mathrm{KS}} = \sum_{i=1}^{N} \epsilon_i - \frac{1}{2} \int\int \frac{\rho\left(\vec{r}\right)\rho\left(\vec{r}\,'\right)}{\left|\vec{r} - \vec{r}\,'\right|}\, d^3\vec{r}\, d^3\vec{r}\,' + E_{\mathrm{xc}}\left[\rho\right] - \int \rho\left(\vec{r}\right) \frac{\delta E_{\mathrm{xc}}\left[\rho\right]}{\delta\rho\left(\vec{r}\right)}\, d^3\vec{r}, \tag{2.12}$$

this energy is called the Kohn-Sham energy.

Provided that we have a functional form of the exchange-correlation energy $E_{\mathrm{xc}}\left[\rho\right]$, the entire DFT theory is exact. Unfortunately, we do not know this functional form, and thus a range of approximations for this exchange-correlation energy have been developed over the years [39]. Upon the addition of additional exact constraints in the approximations, they were found to become more accurate. Perdew and Schmidt arranged these approximations based on their accuracy into the so-called Jacob's ladder of DFT [42, 43]. The simplest approximation, corresponding to the ladder's lowest rung, is the local density approximation (LDA). In LDA the exchange-correlation energy of the system is assumed to solely depend on the electron density $\rho\left(\vec{r}\right)$. For generalized gradient approximations (GGA), the next rung of the ladder, the exchange-correlation energy also depends on the gradient of the charge density $\boldsymbol{\nabla}\rho\left(\vec{r}\right)$, as a result of which the approximation also takes into account inhomogeneities, making it semi-local. In the third rung of the ladder, the meta-GGA approximation, where the exchange-correlation energy also depends on the kinetic energy density of the Kohn-Sham orbitals $\tau\left(\vec{r}\right)$. Higher up the ladder are the hybrid functionals that include the exact exchange energy that follows from Hartree-Fock calculations into the exchange-correlation energy from approximations of the lower three rungs. However, due to the great computational demands of Hartree-Fock calculations and thus hybrid functionals, the work in this thesis is done using the lower three rungs of Jacob's ladder of DFT. The DFT data produced with these functionals, including LDA, GGA, and meta-GGA, is compared to experiments in Appendix B. A GGA functional is used in the rest of this study since it results in the best trade-off between accuracy and computational cost.

2.2. Molecular dynamics

2.2.1. Basic principles

The aim of molecular dynamics (MD) simulations is to obtain a representative time evolution of a model system of particles. To do so, the atoms in the model system are allowed to interact with each other, the resulting force of which is combined with Newton's equations of motion to evolve the system over time [44, 45]. For a model system of N particles, this results in the following set of equations of motion

$$\vec{f}_i = m_i \ddot{\vec{r}}_i = m_i \vec{a}_i, \tag{2.13}$$

where $\vec{f}_i$ is the force acting on i^{th} particle, m_i is the mass of i^{th} particle and $\ddot{\vec{r}}_i$ is the second time derivative to the position $\vec{r}_i$, or acceleration $\vec{a}_i$, of the i^{th} particle. Nevertheless, we still lack any information on the forces that result from the interaction of the particles in the system to evolve the system over time. The covalent and non-bonding interactions[5] between the atoms in a MD simulation are captured through a variety of functional forms by a potential energy function $E\left(\vec{r}^N\right)$ that is dependent on the position of all N particles $\vec{r}^N$. Examples of such potential energy functions, or force fields, are CHARMM [46] and AMBER [47] both used in the simulation of biomolecules. In general, these force fields cannot describe chemical reactions, since they use pre-defined connectivity. ReaxFF is a potential energy function that makes use of a dynamic bond order and is thus able to describe chemical reactions [35] (see section 2.3). Despite the underlying differences in the functional forms used in each potential, the forces acting on the particles in the model system can be determined from the potential energy as

$$\vec{f}_i = -\frac{\partial}{\partial \vec{r}_i} E. \tag{2.14}$$

Whilst the expression used for the evaluation of the forces looks rather simple, in practice, the evaluation of the forces turns out to be the most computationally demanding step. The large system sizes in MD simulations require the iteration over many particles, making the step time-consuming and resource intensive [44].

Due to the many-body nature of our potential energy function, finding an analytical solution to Newton's equations of motion is a very difficult task[6]. Therefore, we employ numerical techniques to evolve our system over time. Although many numerical algorithms to solve differential equations have been devised over the years, such as the Euler, Verlet and corrector-predictor schemes, not all schemes are suited to be used in MD simulations. Here we focus on the most widely used numerical scheme in MD simulations: the velocity Verlet scheme[7] [48]. This scheme

[5] The covalent interactions in a system are usually, but not limited to: bond stretching, valence angle strain and dihedral strain. The non-bonding interactions describe the dispersive and electrostatic interactions in the system.

[6] That is assuming an analytical solution does exist, if not this becomes an impossible task.

[7] The AMS2020 software package used for the ReaxFF simulations in this thesis also makes use of this scheme.

is time-reversible, low order in time and requires just a single evaluation of the force per timestep, making it a suitable numerical scheme for MD simulations [45]. The scheme solves the position $\vec{r}_i$ and velocity $\vec{v}_i$ of each of the N particles at the next timestep, $t + \Delta t$, starting from the positions, velocities and forces of the particles in the previous timestep, t, with the following scheme

$$\begin{aligned} \vec{r}_i\left(t+\Delta t\right) &= \vec{r}_i\left(t\right) + \vec{v}_i\left(t\right)\Delta t + \frac{1}{2}\frac{\vec{f}_i\left(t\right)}{m_i}\Delta t^2 + \mathcal{O}\big(\Delta t^4\big) \\ \vec{v}_i\left(t+\Delta t\right) &= \vec{v}_i\left(t\right) + \frac{\vec{f}_i\left(t\right) + \vec{f}_i\left(t+\Delta t\right)}{2m_i}\Delta t + \mathcal{O}\big(\Delta t^2\big), \end{aligned} \tag{2.15}$$

where the truncation errors of the numerical scheme are denoted with the 'big O'-notation ($\mathcal{O}$). The scheme first determines the new positions of the particles $\vec{r}_i\left(t+\Delta t\right)$, it then evaluates the interatomic forces at the new positions $\vec{f}_i\left(t+\Delta t\right)$ and finally calculates the velocities of the particles at the next timestep $\vec{v}_i\left(t+\Delta t\right)$. Starting from a definition of the initial positions and velocities of the particles, the use of this scheme over longer periods of time results in the atom trajectories for the model system. Whenever the timestep of integration is small enough, the truncation errors of the scheme also become negligible, which means that the obtained trajectories are an accurate description of the time evolution of the model system.

Periodic boundary conditions (PBCs) are employed when bulk properties are to be determined using MD simulations. Using PBCs, the simulation cell or unit cell is arranged in a pseudo-infinite lattice. The unit cell is replicated in each direction, creating an array of simulation cells in which the particles are located in equivalent positions. The minimum-image convention is used to calculate the interactions between the particles. This convention allows for interactions between the particles through the boundaries of the simulation box, which means that the finite model system is mapped onto an infinite system. This allows us to study macroscopic material properties using a model system of particles. The size of the simulation cell should be chosen with care when used together with PBCs. Whenever a simulation cell dimension is of similar size as the interaction lengths in the simulation, correlational artefacts may arise, resulting in unphysical macroscopic behaviour during the MD simulations [44].

2.2.2. Sampling from simulations

The ergodic hypothesis is an integral part of molecular dynamics. The hypothesis states that in an ergodic system the ensemble average of a quantity can be obtained from the time average of that quantity over a long enough time interval [49], which is mathematically expressed as

$$\langle A \rangle = \lim_{t\to\infty} \frac{1}{t} \int A\left(\vec{r}^{\,N}\left(t\right), \vec{v}^{\,N}\left(t\right)\right) dt\,, \tag{2.16}$$

where $\langle \cdots \rangle$ denotes the ensemble average, and A is the quantity of interest that is calculated from the atomic positions $\vec{r}^{\,N}$ and velocities $\vec{v}^{\,N}$. The hypothesis allows us to relate the time evo-

lution of our system to the ensemble-averaged quantities that are measured in experiments. Thus, mapping the quantities we can obtain from our MD simulations to the macroscopic quantities observed in experiments. Finally, it should be highlighted that a long enough time interval, in this case, means that the averaging should take place over times longer than the largest relaxation time in the system.

To complete the mapping between our MD simulations and actual experiments, we should explore the types of systems that we can simulate in some more detail. As a result of the conservation laws and the sole use of forces that stem from a potential energy function, MD simulations by default sample an isolated system, or microcanonical ensemble [44]. In such an isolated system the number of particles N, system volume V and total energy E are all constant, as a result of this, the ensemble is often referred to as the *NVE*-ensemble [44]. However, experiments are rarely carried out in a system with constant energy, typically make use of a constant temperature, a constant pressure or both. This requires the use of a different type of ensemble during the modelling of such systems. Constant temperature simulations are done in the *NVT*-ensemble, also known as the canonical ensemble [44]. Because the ensemble allows for the exchange of energy with a heat bath that is kept at a fixed temperature, the energy of the system fluctuates, but the temperature of the system is kept constant[8]. The canonical ensemble, therefore, has a constant number of particles N, system volume V and temperature T. Simulations in which the temperature and pressure are kept constant make use of the *NpT*-ensemble, also named the isothermal-isobaric ensemble [44]. In addition to a heat bath found in the canonical ensemble, this ensemble also employs a pressure bath. As a result of this pressure bath, the coupling to this pressure bath allows the system volume to change so that the system maintains a constant pressure[9]. In the *NpT*-ensemble, the number of particles N, system pressure p and temperature of the system T are all constant. The exact implementation of the heat and pressure baths is explained in the next subsection, in which the principles of thermostats and barostats are explained.

2.2.3. Simulation conditions

The regulation of the temperature and pressure is important for MD simulations. The use of a thermostat or barostat allows for the control over the temperature and pressure, respectively. In this work, we employ the Berendsen thermostat and barostat for, respectively, temperature and pressure control during the equilibration of our systems, and make use of the more refined NHC-thermostat and MTK-barostat to control the temperature and pressure during the production runs. What follows is a short explanation of the used thermostats and barostats. Other temperature and pressure control principles can be found in literature [44, 50, 51]

[8] Note that constant temperature, in this case, means that it fluctuates around a constant average value, it is not a single-valued.

[9] Again, constant means fluctuating around a constant value of the pressure, it is not necessarily single-valued.

Thermostats

The temperature in MD simulations can be calculated based on the equipartition theorem [44]. In this theorem, which holds under conditions of thermal equilibrium, each degree of freedom that enters the Hamiltonian quadratically is assigned the same average energy. As a result of this, we can associate a temperature T with the average kinetic energy, which has the following definition for a system of N particles

$$T = \frac{2}{N_\mathrm{f} k_\mathrm{B}} \sum_{i=1}^{N} \frac{m_i \vec{v}_i{}^2}{2} \tag{2.17}$$

where N_f is the number of degrees of freedom which for a three-dimensional simulation with N particles is $3N$ and k_B is the Boltzmann constant.

The temperature in an MD simulation can thus be controlled by changing the velocities $\vec{v}_i$ of the particles in the system at every simulation step. The Berendsen thermostat is one method to rescale the particle velocities and thus regulate the temperature in the model system [52]. The method makes use of a weak coupling of the model system to an external heat bath. This heat bath is kept at a constant target temperature T_0 and the tightness of the coupling is determined by a damping constant τ_T. The velocities are rescaled at every timestep with the velocity rescaling factor λ

$$\lambda = \left[1 + \frac{\Delta t}{\tau_\mathrm{T}} \left(\frac{T_0}{T} - 1\right)\right]^{\frac{1}{2}}, \tag{2.18}$$

in which Δt is the timestep of the simulation. The rescaled velocities $\vec{v}_{i,\mathrm{Ber}}$ are obtained for each particle after every integration step through a multiplication of the particle velocity with the scaling factor as

$$\vec{v}_{i,\mathrm{Ber}} = \lambda \vec{v}_i. \tag{2.19}$$

The thermostat scheme demonstrates that the coupling between the heat bath and the system can be adjusted through the damping constant value.

However, there are two disadvantages to the use of a Berendsen thermostat. First, it dampens out the kinetic energy fluctuations, as a result of which simulations with the Berendsen thermostat do not properly sample an *NVT*-ensemble [50]. Moreover, the thermostat has the potential to result in simulation artefacts such as the 'flying-ice cube'-problem[10] that are typical for velocity rescaling thermostat algorithms [53, 54]. Despite its shortcomings, the Berendsen thermostat can still be of use during the equilibration stage of the simulations, since it is very effective in equilibrating a system to a target temperature [50].

Another method to control the temperature during MD simulations is the Nosé-Hoover thermostat [55, 56]. Instead of coupling to a heat bath, a Nosé-Hoover thermostat extends the real

[10] The 'flying-ice cube'-problem is a simulation artefact in which the kinetic energy of the internal degrees of freedom is transferred to the linear momentum of the system, as a result of which a 'frozen' system (read: without any internal movements) starts to 'fly' through the simulation space, hence its name.

simulated system with an additional degree of freedom representing a fictitious heat bath. A subsequent derivation of the equations of motions for this extended system results in the following set

$$\begin{aligned} \frac{d\vec{v}_i\left(t\right)}{dt} &= \frac{\vec{f}_i\left(t\right)}{m_i} - \zeta\vec{v}_i\left(t\right) \\ \frac{d\zeta\left(t\right)}{dt} &= \frac{1}{Q}\left[\sum_{i=1}^{N} m_i\vec{v}_i\left(t\right)^2 - (N_\mathrm{f}+1)k_\mathrm{B}T\right], \end{aligned} \tag{2.20}$$

where ζ is a thermodynamic friction coefficient and Q is the coupling constant that determines the timescales at which the heat is exchanged between the system and the heat bath. As is demonstrated by these equations of motion, the addition of this additional degree of freedom results in a rescaling of the velocities and thus a control over the system temperature. Although the Nosé-Hoover thermostat is already a better approximation of the *NVT*-ensemble than the Berendsen thermostat is, it was shown that for small systems the method deviates from the canonical ensemble [57]. Martyna *et al.* demonstrated that this deficiency for small systems could be overcome by coupling a series of Nosé-Hoover thermostats to each other, which did result in the proper sampling of a canonical ensemble (*NVT*-ensemble) [57]. As a result of this, the Nosé-Hoover chains (NHC) thermostat is preferred over the Berendsen and Nosé-Hoover thermostat whenever production runs are done during which the temperature is controlled.

Barostats

Analogous to the temperature T, we can also calculate the pressure p during MD simulations. The pressure can be calculated using the virial expression for the pressure [44], as

$$p = \frac{Nk_\mathrm{B}T}{V} + \frac{1}{d\,V}\sum_{i=1}^{N}\vec{r}_i\cdot\vec{f}_i \tag{2.21}$$

where V is the system volume and d is the dimensionality of the simulation, which for a three-dimensional simulation is equal to 3.

From the above expression, we observe that the pressure in the simulation can be controlled by regulating the volume V of the system. A Berendsen barostat is one such method with which the pressure of the system can be controlled [52]. The idea behind the barostat is very similar to that of the Berendsen thermostat. Instead, the system is now weakly coupled to a pressure bath at target pressure p_0. Alternatively, the volume of the system and positions of the particles are changed, not the velocities of the particles which was done by a Berendsen thermostat. Again, a damping constant τ_p is defined, which results in a spatial scaling factor μ

$$\mu = \left[1 - \frac{\Delta t}{\tau_\mathrm{p}}\left(p_0 - p\right)\right]^{\frac{1}{3}}, \tag{2.22}$$

where Δt is the timestep of the simulation and p_0 is the target pressure. This scaling factor is then used to scale the particle positions $\vec{r}_i$ and the dimensions of the simulation box $\vec{R}_i$, after every integration step, as

$$\begin{aligned} \vec{r}_{i,\text{Ber}} &= \mu \vec{r}_i \\ \vec{R}_{\text{Ber}} &= \mu \vec{R}, \end{aligned} \tag{2.23}$$

with $\vec{r}_{i,\text{Ber}}$ and $\vec{R}_{i,\text{Ber}}$ as the rescaled particle positions and simulation cell, respectively. The scaling relations demonstrate that analogous to the thermostat, the damping constant can be used to control the coupling between the pressure bath and the system.

Unfortunately, the Berendsen barostat has a similar flaw as the Berendsen thermostat: it does not allow for realistic fluctuations. As a result of this, when combined with a thermostat, using a Berendsen barostat does not result in a proper sampling of a *NpT*-ensemble, since the pressure fluctuations are dampened too heavily [50]. However, similar to the Berendsen thermostat, the Berendsen barostat still sees some use in the equilibration of the system to a target pressure.

Analogous to the thermostats, alternative methods for barostats have been devised. One of these methods is the Martyna-Tobias-Klein (MTK) barostat [58]. The method closely resembles that of the NHC-thermostat, but instead of a chain of heat baths, we now use a chain of pressure baths and include that as additional degrees of freedom in the system. Again the equations of motion can be derived for this extended system. In these equations, an overview of which can be found in the paper by Martyna *et al.* [58], the system size, particle velocities and particle positions are all scaled to allow for control over the system pressure. This MTK-barostat results in a representative sampling of the *NpT*-ensemble making it suitable control mechanism for thep ressure during simulations.

2.3. Reactive force field (ReaxFF) potential

As is the case for all classes of interatomic potentials, a ReaxFF potential is built up from various energy contributions. The ReaxFF potential energy expression E_{system} contains the following energy contributions

$$E_{\text{system}} = \underbrace{E_{\text{bond}} + E_{\text{val}} + E_{\text{tors}} + E_{\text{over}} + E_{\text{under}} + E_{\text{lp}}}_{\text{Covalent}} + \underbrace{E_{\text{vdW}} + E_{\text{Coul}}}_{\text{Non-bonding}} + E_{\text{spec}}, \tag{2.24}$$

where the energy terms were divided into a covalent and non-bonding group. E_{bond} describes the energy associated with the bonds between the atoms. E_{val} and E_{tors} are the three- and four-body terms associated with valence angle strain and torsional strain, respectively. E_{over} and E_{under} are the energy penalties associated with the over- and undercoordination of the atoms in the system and E_{lp} is the energy penalty assigned to the break up of a lone pair of electrons. E_{vdW} and E_{Coul} represent the electrostatic and dispersive interactions between the atoms. E_{spec}

contains system-specific interactions such as hydrogen bonds or conjugation effects, but these are not included in the ReaxFF potential form by default. An overview of the scheme with which all contributions to the ReaxFF potential are calculated is shown in Figure 2.1.

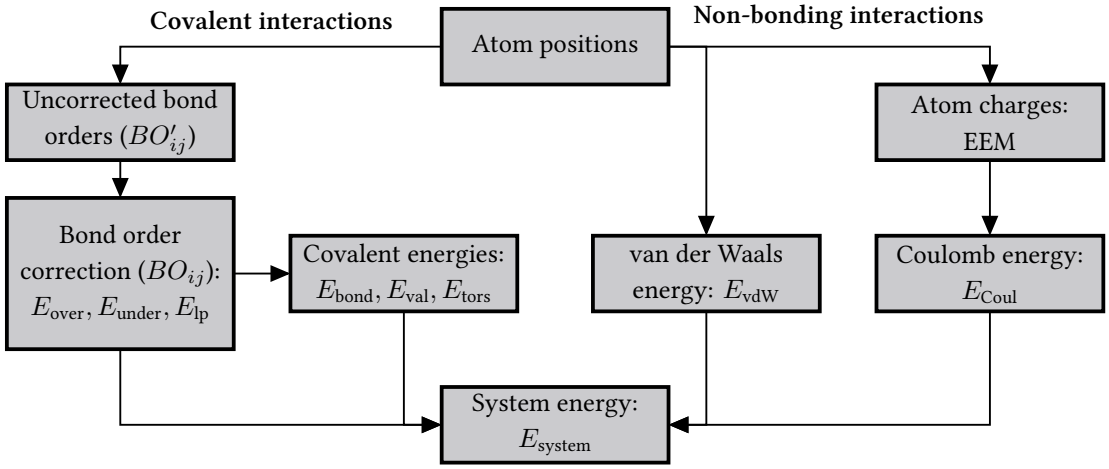

Figure 2.1. A schematic overview of the scheme with which the separate contributions to the ReaxFF potential energy function are calculated. The covalent interactions (left) are only calculated for the bonded atoms, whereas the non-bonding interactions (right) are calculated for all atom pairs in the system. The scheme is adapted from [59].

All the covalent interactions in ReaxFF depend on the connectivity, i.e. bonding between the atoms. In ReaxFF, the so-called bond order defines whether two atoms in the system are bonded. A fundamental assumption in ReaxFF is that this bond order, which is partitioned into contributions of the σ-, π- and $\pi\pi$-bond, can be directly obtained from the interatomic distance of atoms. The uncorrected bond order BO'_{ij} between an atom i and j can therefore be determined with the following empirical relation

$$\begin{aligned} BO'_{ij} = BO'^{\sigma}_{ij} + BO'^{\pi}_{ij} + BO'^{\pi\pi}_{ij} &= \exp\left(p_{bo1} \cdot \left(\frac{r_{ij}}{r_0^{\sigma}}\right)^{p_{bo2}}\right) \\ &+ \exp\left(p_{bo3} \cdot \left(\frac{r_{ij}}{r_0^{\pi}}\right)^{p_{bo4}}\right) + \exp\left(p_{bo5} \cdot \left(\frac{r_{ij}}{r_0^{\pi\pi}}\right)^{p_{bo6}}\right), \end{aligned} \tag{2.25}$$

where r_{ij} is the interatomic distance between the atom i and j, r_0^x ($x = \sigma, \pi, \pi\pi$) are the equilibrium bond distances and $p_{bo,y}$ ($y = 1, \ldots, 6$) is a set of empirical bond order parameters. For each bond type, where the typing is purely based on the types of atoms involved in the bond, a ReaxFF force field has unique values for the parameters (r_0^x and $p_{bo,y}$) found in the above relation. This allows for a unique definition of each type of bond in the system and thus an accurate description of the overall bonding. As an example, a hydrogen-hydrogen (H-H) bond description by ReaxFF only makes use of the σ-bond that has a very short equilibrium distance[11], whereas it

[11] To not use a certain bonding contribution in ReaxFF, the equilibrium bond distance of the contribution is typically set to a very small value. The functional form of the uncorrected bond order then ensures the bond order that exists for that contribution is negligible.

uses longer-reaching σ-, π- and $\pi\pi$-bonds to represent a carbon-carbon (C-C) [35].

Although Equation 2.25 is a good starting point for describing the overall chemical bonding in the system, the rather long-ranged nature of the uncorrected bond order tends to overcoordinate the atoms in the system [35, 59]. Therefore, ReaxFF employs a bond order correction scheme based on a pre-defined atom valency and some over- or undercoordination parameters. This scheme changes the uncorrected bond order BO'_{ij} to a corrected bond order BO_{ij} that results in a more realistic coordination of the atoms, and thus overall bonding in the system. Even though the bonding in the system is corrected based on the valencies of the types of atoms, over- and undercoordination can still occur for the atoms in the system, potentially resulting in a break up of the lone pairs of electrons at that atom. The discrepancies in the coordination of the atoms with respect to their valency are resolved by adding E_{over}, E_{under} and E_{lp} penalty terms.

As an example of the functional dependence of the covalent interactions on the bond order E_{bond} is calculated with the following expression

$$E_{\text{bond}} = -D_e^{\sigma} \cdot BO_{ij}^{\sigma} \cdot \exp\left(p_{be1} \cdot (1-(\,BO_{ij}^{\sigma}\,)^{p_{be2}})\right) - D_e^{\pi} \cdot BO_{ij}^{\pi} - D_e^{\pi\pi} \cdot BO_{ij}^{\pi\pi}, \quad (2.26)$$

where D_e^x ($x = \sigma, \pi, \pi\pi$) are the bond dissociation energies of the bonding contributions and p_{be1} and p_{be2} are empirical bond energy parameters. These parameters are unique to the type of bond that is considered. The functional dependence of the bond order ensures that whenever two atoms are not bonded to each other, i.e. a negligible bond order between the atom pair, the bond energy for that pair becomes negligible as well. The functional forms of E_{val} and E_{tors} are set up in a similar way so that they only contribute whenever bonds are present between the atom pairs. A detailed description of the functional forms of the other covalent interactions is beyond the scope of this thesis, their functional forms can be found in literature [35, 60, 61].

The non-bonding interactions, the electrostatic and dispersive interactions, are calculated for all atom pairs in the system, regardless of the interatomic bonding. ReaxFF calculates the atomic charges, required for the electrostatic interactions, with the electronegativity-equalization method (EEM) proposed by Mortier *et al.* [62]. The dispersive (van der Waals) interactions are determined from these parameters using a shielded Morse potential, and shielded Coulomb potentials are used to calculate the electrostatic interactions in the system. Both types of interactions use a shielded form, which is done to prevent any excessive attractive or repulsive interactions from occurring at short distances. As is the case for the other interactions, these non-bonding interactions are captured by a set of empirical parameters, their complete functional for can be found in literature [35, 61].

To sum up, the heavy dependence of the ReaxFF potential on the dynamic bond order requires the determination of this bond order at every timestep during a ReaxFF simulation [59]. This step makes ReaxFF simulations an order of magnitude more demanding than classical force field methods that use pre-defined connectivities. However, this scheme also enables MD simulations with ReaxFF to simulate the creation and breaking of bonds and thus describe chemical reactions.

Finally, it should be noted that the energy contributions in the ReaxFF potential are heavily parameterized and empirical. To obtain values for the ReaxFF parameters that result in an accurate description of the chemical system, the parameters are trained against a large set of reference data. The training procedure is outlined in the next chapter (chapter 3).

3. Force field parameterization

In this chapter, we explain the principles behind creating a ReaxFF parameter set, a procedure that is typically called *'force field parameterization'*. The chapter starts with an explanation of the basic concepts of the parameterization of the force field. After that, a brief overview of the available methods is given, with the optimization method of choice, a Monte Carlo optimization algorithm, explained in more detail. We then show how we applied this parameterization procedure to obtain a ReaxFF parameter set for our material system of interest: $CsPbI_3$. Here we describe the creation of the training set, the set-up of the parameterization runs and a short validation of the ReaxFF parameter set that best captures the material behaviour.

3.1. Basics of force field parameterization

An error function is used to find an optimal set of parameters for the material system of interest. This error function quantifies how much the ReaxFF force field deviates from the reference data in the training set. In short: the smaller the value of the error function, the better the ReaxFF parameter set describes the material. As is conventional for parameterizations of ReaxFF force fields, this error function is a sum of squares [35], that is defined as

$$\text{Error}\left(\{p_j\}\right) = \sum_{i=1}^{N} \left[\frac{x_{i,\text{ref}} - x_{i,\text{calc}}\left(\{p_j\}\right)}{\sigma_i} \right]^2, \tag{3.1}$$

where $x_{i,\text{ref}}$ is the reference value of the i^{th} property as found in the training set, $x_{i,\text{calc}}\left(\{p_j\}\right)$ is the value of that same property calculated with the ReaxFF parameter set $\{p_j\}$ and σ_i is the weight of that property in the training set. The weight enables the scaling of the values of the different properties in the training set based on the relative importance and accuracy of the data. Moreover, since many different types of quantities[12] can be contained in a training set, this weighting factor ensures each property is dimensionless, which allows for the properties to be summed together.

Thus, force field parameterization aims to find the set of parameters that corresponds to the lowest value for the error function. Since ReaxFF force fields can contain approximately 100

[12] The types of quantities that can be included in a ReaxFF training set include energies, atomic charges, bond lengths, valence angles and unit cell parameters.

parameters for every type of atom [63], it is evident that finding the minimum of the error function is a hyper-dimensional global optimization problem. To illustrate the complexity of such optimization problems, a hypothetical error function and its dependence on two arbitrary parameters is plot in Figure 3.1. Even though the surface in this figure does not resemble an actual ReaxFF error surface, it does contain typical features that make the minimization of a ReaxFF error function difficult. The features are the many irregularities in the energy surface and the abundance of local minima [63, 64].

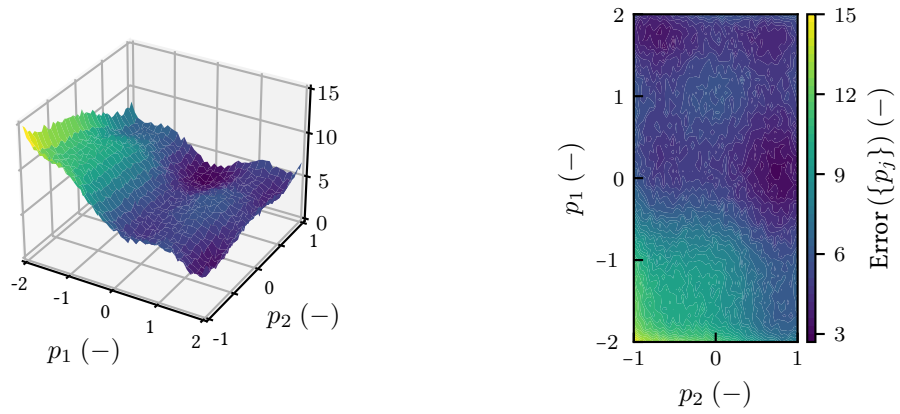

Figure 3.1. A visualization of the complexity of the error function by plotting the hypothetical dependence of the error function on two parameters from the parameter space $\{p_j\}$. (Left): surface plot. (Right): contour plot.

3.2. A Monte Carlo-based force field optimizer

The majority of the ReaxFF parameter sets in literature were optimized with the successive one-parameter search algorithm [65]. This method optimizes the parameter space on a parameter-by-parameter basis under the assumption that the error function is related to each of the parameters in a parabolic fashion[13]. Although many ReaxFF force fields have been published with this single-parameter parabolic-search, such as a hydrocarbon [35] or silica force field [60], the method of optimization is not very robust. The method is a local optimization method that can only find local minima, thus requiring an already good guess for the initial parameters to end up with a reasonable ReaxFF force field [63]. Moreover, the error surface is not necessarily parabolic, making it questionable whether the one-parameter search method can locate a minimum in every ReaxFF parametrization problem [63, 64].

A recent development in ReaxFF force field optimization is to make use of global optimization techniques. Due to their stochastic nature, these methods are more robust in finding the

[13] Since all parameters in a ReaxFF force field are related to each other, the optimization of one parameter shifts the optimal value of another parameter. Therefore, the successive one-parameter search algorithm is repeated until the parameters are found to have converged: when the value of the error function does not change anymore.

global minimum and thus a more accurate ReaxFF description of the material system. Some global optimization algorithms that have been used in conjunction with ReaxFF are Monte Carlo methods [63], genetic algorithms [64] and the covariance matrix adaptive evolutionary strategy [66], with each of the named global optimization algorithms outperforming the successive one-parameter search algorithm.

In comparing the global optimization techniques for ReaxFF force fields, none of the algorithms was found to outperform the other global optimization techniques [67]. Therefore, we opted to parameterize our ReaxFF force field with the most established global ReaxFF force field optimizer: the Monte Carlo-based optimizer (MCFF) [63]. The optimization method closely follows the Monte Carlo simulated annealing method as proposed by Kirkpatrick *et al.* [68], which makes use of the Metropolis Monte Carlo algorithm to sample the parameter space appropriately [69]. As a result of this, the optimization procedure in MCFF can be mapped to a physically analogous system that cools down over time, an additional benefit that makes the interpretation of the optimization progress easier. We made use of the implementation of MCFF in AMS2020 [70, 71]. The optimization procedure of MCFF, of which a schematic overview is shown in Figure 3.2, is detailed in the rest of this section.

A ReaxFF force field parameterization requires reference data, active parameters, optimization bounds and a starting point. The reference data is a set of data, experimental or computational, representing the material system of interest. Through the active parameters and optimization bounds, the user defines which and to what extent parameters are allowed to change to end up with a ReaxFF force field that accurately replicates the training set. The starting point for the optimization is handed to the optimizer through an initial set of ReaxFF parameter $\{p_j\}_{\text{init}}$. To aid the parameter optimization, these parameters should already roughly describe the chemical system at hand. The algorithm iteratively improves the initial guess for the ReaxFF parameters to end up with an accurate description of the training set and thus the chemical system of interest.

At every iteration of the MCFF optimization algorithm, a new set of parameters $\{p_j\}_{\text{new}}$ is generated from the current best parameter set $\{p_j\}_{\text{best}}$. In the first iteration of the algorithm, this best parameter set is the initial ReaxFF parameter set. The new parameters sets are obtained by changing the values of a fraction of the active parameters in the current best force field within the pre-defined bounds. With both parameter sets, the new and the current best, established for this iteration, the force field optimizer evaluates the quality of the two parameter sets with the error function as defined in Equation 3.1. To compare their relative quality, the optimizer calculates the difference in the error functions ΔError which is defined as

$$\Delta\text{Error} = \text{Error}\left(\{p_j\}_{\text{new}}\right) - \text{Error}\left(\{p_j\}_{\text{best}}\right). \tag{3.2}$$

A negative or positive difference, respectively, indicates that the new parameter set is, either of better or of worse quality than the current best parameter set.

Having established the relative quality of the two parameter sets during this iteration, the algo-

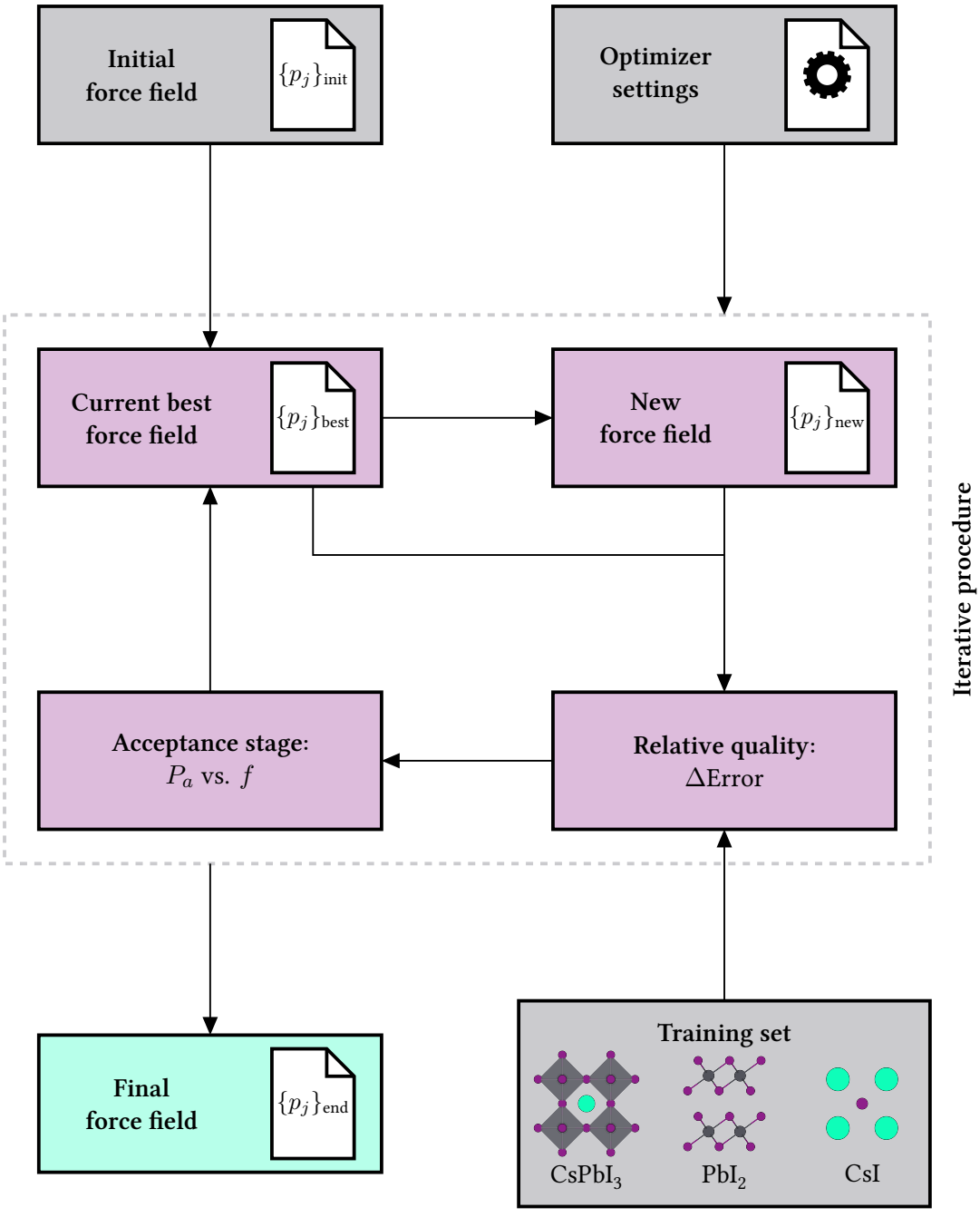

Figure 3.2. A schematic overview of the Monte Carlo-based reactive force field optimizer [63]. The fields are colour-coded to represent the type of data they represent: input (grey), production (purple) and output (cyan).

rithm shifts to the *'acceptance stage'*. In this stage, the algorithm either accepts or rejects the new parameter set based on the acceptance probability. The probability P_a used by MCFF optimizer is the same as used in the Metropolis Monte Carlo algorithm [63, 69], and is defined as

$$P_a = \min\left[1, \exp(-\beta \Delta \text{Error})\right], \tag{3.3}$$

where $\beta = \frac{1}{k_B T}$ is a measure of the temperature of our optimization problem. A random number f is then sampled on the interval $0 \leq f < 1$. Only if $P_a > f$ the new parameter set becomes the best parameter set for the next iteration, else the current best parameter set remains the best parameter set for the next iteration. To sum up, this means that for $\Delta\text{Error} < 0$ ($P_a = 1$) the new parameter set always becomes the current best parameter set in the next iteration, and whenever $\Delta\text{Error} > 0$ ($P_a = \exp(-\beta\Delta\text{Error})$) there is a chance, which decreases with increasing ΔError, that the new parameter set is accepted as the current best parameter set.

The outlined procedure is then repeated over many cycles, with each successive cycle having a slightly larger value of β, through which the gradual cooling of the system is mimicked. As a result of this cooling, the parameter set has an increasingly smaller probability for uphill motion in the error landscape as the optimization progresses. After a certain number of optimization cycles, the number of which is preset by the user, the ReaxFF parameters are found not to change any further, and the parameters have converged. If this final parameter set $\{p_j\}_{\text{end}}$ is close enough to the global minimum, the final set of parameters is a good description of the reference data in the training set and thus the material system.

To control the Monte Carlo-based optimization procedure, a variety of settings are supplied to MCFF. First, we define the initial value and rate of change for β, through which we define the cooling process the parameter set is subjected to during the optimization with simulated annealing. In addition to the temperature, we set the fraction of parameters and the step size of those parameters that are allowed to vary during the generation of a new set of ReaxFF parameters in a MCFF iteration. Finally, it should be highlighted that we can set a target acceptance rate for the optimizer. The optimizer scales the step size of the active parameters to attain this acceptance rate for new ReaxFF parameter sets. Since we only covered some basic optimizer options, it is advised to consult the documentation of the optimizer for a more in-depth coverage of the options [70, 71].

In Figure 3.3 the progression of the error function during a ReaxFF force field optimization with the MCFF algorithm is shown. During the initial iterations of the optimization algorithm, the error function shows a rather large decrease in value, which is associated with the movement of the parameter set to a minimum. Near the end of the optimization, the error function shows a small decrease of the error function, which is related to the movement of the parameter set close to a minimum. The optimization shows occasional upward spikes in the error function because of which the optimization algorithm is less prone to getting stuck in a local minimum of the error function. Nevertheless, the overall progression of the error function remains a downward one, demonstrating that the parameter set optimizes into a minimum through simulated annealing.

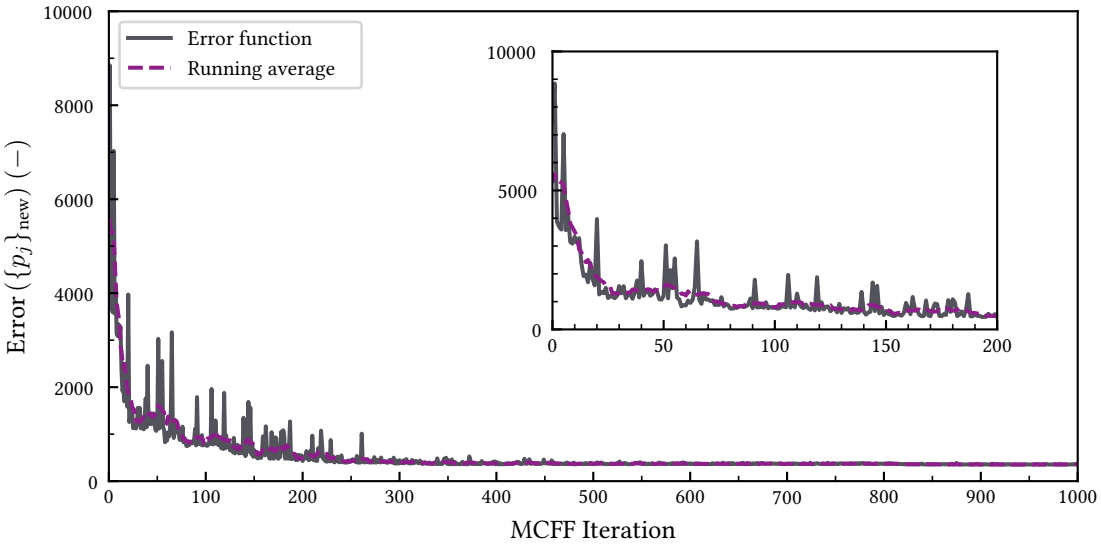

Figure 3.3. Visualization of the error function evolution during a parameterization procedure with the MCFF optimizer. The inset is a zoomed-in picture of the initial 200 MCFF iterations.

3.3. Parameterization of $CsPbI_3$ ReaxFF parameters

In this section, we explain the parameterization of the ReaxFF parameters for $CsPbI_3$. The section is split into three subsections that cover the training set, parameterization procedure and validation of the ReaxFF parameter set.

3.3.1. Training set

Since the ReaxFF parameter set aims to simulate the dynamical behaviour of different perovskite phases, equations of state of the orthorhombic, tetragonal and cubic phase of $CsPbI_3$ were included in our training set. A full description of all phases of $CsPbI_3$ is found in Appendix A. In addition to the $CsPbI_3$ equations of state, we also included some equations of state for the precursors, CsI and different polytypes of PbI_2. The reason for this was to make our final set of ReaxFF parameters more transferable across different phases. As was noted earlier, metal halide perovskites are defect-rich. Therefore a realistic description of $CsPbI_3$ should also cover those defects. For that reason, we included both the formation energies and migration barriers, of a variety of vacancies in monolayer PbI_2 and bulk $CsPbI_3$ in our training set. Finally, since metal halide perovskites are mixed covalent-ionic compounds, we also included atomic charges in our training set to enable the ReaxFF force field to capture the system's ionicity accurately. An overview of the types of data included in the training set is shown in Figure 3.4.

The reference data in our training set was generated using the VASP software package [72–75]. All types of reference data: equations of state, atomic charges, defect formation energies and

defect migration barriers, were calculated using the PBE XC-functional [76, 77], with the long-range dispersive interactions being accounted for by the DFT-D3(BJ) dispersion correction [78] (see Appendix B for the motivation). We treated the outermost electrons of Cs [$5s^2 5p^6 6s^1$]; Pb [$5d^{10} 6s^2 6p^2$] and I [$5s^2 5p^5$] as valence electrons, the electron-ion interaction was modeled with the projector-augmented wave (PAW) method [79, 80]. Furthermore, we used the plane-wave basis set to an energy cutoff of 500 eV and made use of Monkhorst-Pack meshes for the Brillouin zone integration [81]. None of the calculations included spin-orbit coupling, since it is known to have minimal effects on the ground state geometry and relative stability between the different phases [82–84].

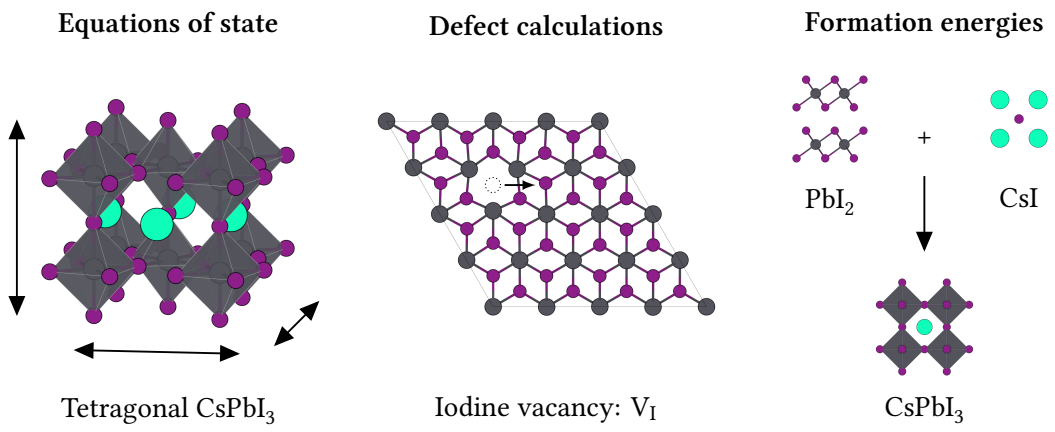

Figure 3.4. An overview of the types of data that were included in the training set for the parameterization of the $CsPbI_3$ ReaxFF force field.

The equilibrium geometry of all materials was obtained from structural relaxations. During these relaxations we allowed the ionic positions, cell shape and cell volume to change until the energy and force converged to within $1 \cdot 10^{-3}$ meV and 10 meV Å^{-1}, respectively. Here we made use of the following k-space grids, which resulted in an energy convergence to within 1 meV/atom: PbI_2: $11 \times 11 \times 7$; CsI: $12 \times 12 \times 12$; cubic $CsPbI_3$: $10 \times 10 \times 10$; tetragonal $CsPbI_3$: $7 \times 7 \times 10$; orthorhombic $CsPbI_3$: $7 \times 7 \times 5$. The atomic charges were calculated for these equilibrium geometries with the Bader charge analysis method [85–88]. Whenever a monolayer material was modeled (e.g. PbI_2 monolayer), we employed a vacuum layer of at least 15 Å to prevent interactions between the periodic images of the monolayer. The equations of state were generated by straining the lattice vectors of the equilibrium geometries, and subsequently allowing the ionic positions to relax to the above-mentioned energy and force convergence criteria.

Defect calculations were done in both PbI_2 monolayers and the orthorhombic phase of $CsPbI_3$. To limit the interactions between the periodic images of the defects, these calculations employed supercell geometries. For PbI_2 and $CsPbI_3$ we, respectively, used a $4 \times 4 \times 1$ and $2 \times 2 \times 1$ supercell with the k-points scaled to $3 \times 3 \times 1$ and $2 \times 2 \times 3$. The computational cost of these defect calculations was reduced by making the convergence criteria less strict, for PbI_2 this resulted in an energy and force convergence criterion of $1 \cdot 10^{-1}$ meV and 30 meV Å^{-1}, whereas for $CsPbI_3$ it was set to $1 \cdot 10^{-2}$ meV and 50 meV Å^{-1}. The defect formation energies were determined from

the difference in energy between the defective and corresponding pristine structures. The defect migration barriers were determined from transition state calculations with five intermediate geometries, using the Climbing Image Nudged Elastic Band (CI-NEB) method [89, 90].

3.3.2. Parameterization

We used the I and Cs ReaxFF parameters of the electrolyte/water force field parameters published by Fedkin *et al.* [91] as a starting point for the parameter optimization. Without a ReaxFF description for lead in literature, the starting point for the Pb-parameters was based on appropriately modified parameters[14] from the chemically similar element platinum as published by Fantauzzi *et al.* [92].

Following an analysis of the nature of the bonding in metal halide perovskites, our force field was set up so that the atomic species (I, Pb and Cs) would only interact through σ-bonds [93]. To limit the complexity of the force field optimization, we kept the number of interactions in the ReaxFF description of $CsPbI_3$ to a minimum. Because the chemical bonds in $CsPbI_3$ and its precursors are relatively simple, it was not necessary to include torsional and hydrogen bonding terms. That gave us a total of 4 bonds (I-I, I-Pb, I-Cs, Pb-Pb) and 7 valence angles (I-I-Pb, I-Pb-Pb, I-Pb-I, Pb-I-Pb, I-Cs-I, Cs-I-Cs, Cs-I-Pb).

During the parameterization, we optimized the atomic parameters of each element and interactions between the elements. For all elements, I, Pb and Cs, the charge (EEM) and van der Waals parameters were optimized. Since the lead parameters were not based on existing ReaxFF parameters for atomic lead, we also allowed the σ- and van der Waals bond radii of Pb to vary during the parameterization. To make our parameterization efforts more robust to the stochastic nature of the MCFF optimizer, we carried out multiple (typically 20) parameterizations in parallel. This allowed us to select the overall best parameter set from the different runs, making our parameterization procedure less prone to be trapped in local minima. The initial temperature was set through an inverse temperature of $\beta = 0.001$. The system was cooled by increasing the value of β in each consecutive cycle by dividing it by 0.999 and subsequently adding 0.001 to it. In each iteration, we allowed a maximum of 10% of the active parameters to vary when generating the new force field. A total of 1000 Monte Carlo iterations, with a target acceptance rate of 30%, was found to be enough to result in parameter convergence.

The parameterization procedure resulted in a final best ReaxFF parameter set; the force field file is presented in Appendix C. A comparison between the best ReaxFF parameter set and the reference data is shown in Figure 3.5. To show the different data types in the training set, we divided the entries into several groups: equations of state (E_{EOS}), formation energies (E_{f}) and migration barriers (E_{m}). The figure illustrates an overall good agreement of the force field with reference data entries in the training set, verifying that our parameterization procedure was

[14] The valency was kept at 2 since the most common oxidation state of Pb is $+2$. However, we adjusted the atomic mass in the parameters to match that of Pb: 207.2 u.

effective in producing a representative parameter set. The most notable discrepancy between the force field and the reference data is found for the defect formation energies (E_f^{def}). The force field is shown to overestimate those for the lead vacancies and underestimate that of the iodine vacancies. Nevertheless, the relative magnitude of the defect formation energies is still properly captured.

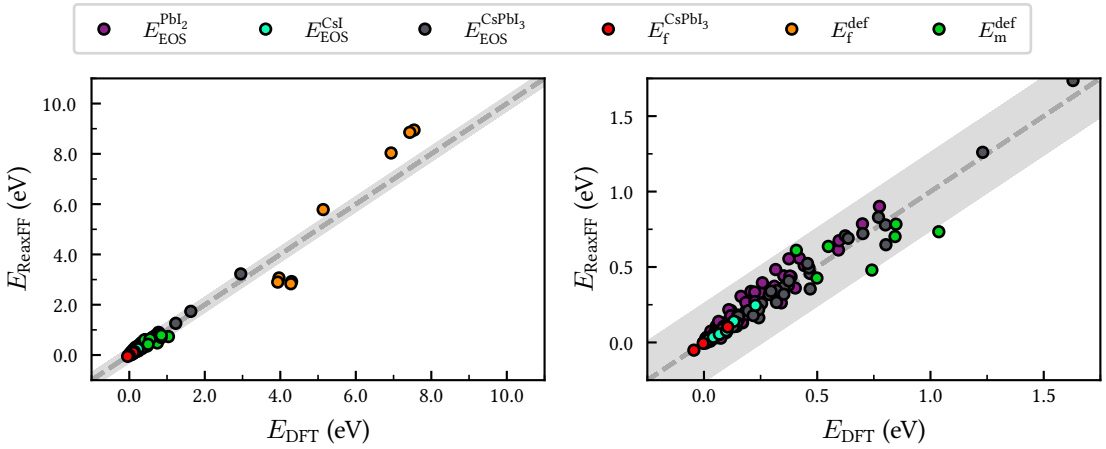

Figure 3.5. An overview of agreement between the ReaxFF parameter set and the DFT reference data from the training set. The figures show (left) an overview of the entire training set and (right) a zoomed-in look into the data located near the origin.

3.3.3. Validation

In the previous subsection, it was demonstrated that we were able to obtain a set of ReaxFF parameters that agrees well with the training set, here we turn towards the validation of the parameter set. This validation aims to show that our ReaxFF parameter set not only replicates the data from the training set well but also that the parameter set also has predictive power, i.e. correctly capturing processes that were not included in the training set. We show in total three of validations.

As a first validation of the parameter set, we took a closer look at the equations of state for the perovskite phases of $CsPbI_3$ against which the ReaxFF force field was trained. A comparison of the DFT reference data from the training set and the results from the ReaxFF parameter set is shown in Figure 3.6. Due to their inclusion in the training procedure, the DFT- and ReaxFF-calculated equations of state resemble each other well. We fitted the Birch-Murnaghan equation of state [94, 95] to the α-$CsPbI_3$ equation of state, which resulted in a value for the bulk modulus of $B_0 = 16.87$ GPa and 16.92 GPa, for ReaxFF and DFT, respectively. Both of which are very close to the experimentally reported values of the bulk modulus of cubic $CsPbI_3$ that range from 14.38 GPa - 17 GPa [34, 96–98]. This demonstrates that our ReaxFF parameter set captures the proper curvature of the potential energy surface of $CsPbI_3$ near the equilibrium.

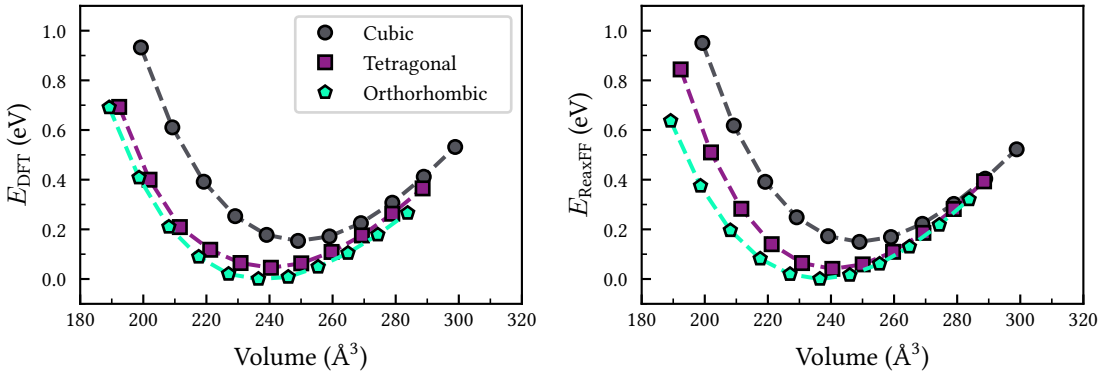

Figure 3.6. A comparison of different equations of state of $CsPbI_3$ with (left) the DFT reference data from the training set and (right) the results of the best ReaxFF parameter set. To ease the comparison of the different phases, the unit cell volumes were reported per unit of $CsPbI_3$.

Since the prediction of the correct equilibrium structures is an important requirement for a good force field, we next carried out geometry optimization of the three phases of $CsPbI_3$. We started from the DFT-optimized structures and allowed the ionic positions, cell shape and cell volume to change during the optimizations. The results of these calculations are shown together with benchmark results from DFT calculations and experimental X-ray diffraction measurements [99] in Table 3.1. From the comparison of the unit cell geometries, we learn that the ReaxFF parameter set replicates both experiments and DFT calculations well (within 2%), thus indicating that our force field performs very well in geometry optimizations of $CsPbI_3$. The ability of the force field to result in accurate geometry optimization confirms that our force field captures the details of the potential energy surface very well.

Table 3.1. Equilibrium geometries of the perovskite phases of $CsPbI_3$ as calculated with ReaxFF or DFT calculations, with experimental X-ray diffraction measurements as reference [99].

Structure	Type	a (Å)	b (Å)	c (Å)	V (Å^3)
Cubic $CsPbI_3$	ReaxFF	6.29	6.29	6.29	248.86
	DFT	6.29	6.29	6.29	249.05
	Exp.	6.30	6.30	6.30	249.64
Tetragonal $CsPbI_3$	ReaxFF	8.67	8.67	6.41	481.31
	DFT	8.66	8.66	6.41	480.88
	Exp.	8.83	8.83	6.30	490.78
Orthorhombic $CsPbI_3$	ReaxFF	8.59	8.95	12.39	952.56
	DFT	8.43	8.99	12.48	945.94
	Exp.	8.62	8.85	12.50	953.85

Finally, we validated our ReaxFF parameter set based on some energy barriers not explicitly included in the training set. To this end, we compare the energy paths that result from CI-NEB

calculations with our force field in AMS2020 to those calculated with DFT in VASP. The calculations made use of five intermediate images and an energy and force convergence criterion of $1 \cdot 10^{-1}$ meV and 30 meV Å^{-1}. In the tests we focus on two types of barriers, both shown in Figure 3.7: the transition between the 1H-phase (ground state) and 1T-phase of a PbI_2 monolayer, and the inversion of the octahedral tilting in β-$CsPbI_3$. From the figure, we can see that our ReaxFF parameter set is a satisfactory describing the relative stability of 1H- and 1T-phases of monolayer PbI_2. Whereas the force field properly captures the overall reaction pathway of the phase transition, it overpredicted the activation barrier by about 0.21 eV, being 0.56 eV compared with the 0.35 eV from DFT. Such discrepancy is very likely caused by a combination of the bonding, Coulomb and undercoordination contributions to the ReaxFF potential energy function and which merits further refinement, especially when focusing on the properties of PbI_2 monolayer. Nevertheless, the ReaxFF calculations perform noticeably better for similar processes in $CsPbI_3$, in this case, the prediction of the inversion barrier for the octahedral tilting in β-$CsPbI_3$. The calculated barrier of 0.14 eV agrees well with our DFT result of 0.17 eV. With this, we have confirmed that even processes that we did not explicitly train against are described with reasonable accuracy, thus paving the way for further exploration of new structures and processes.

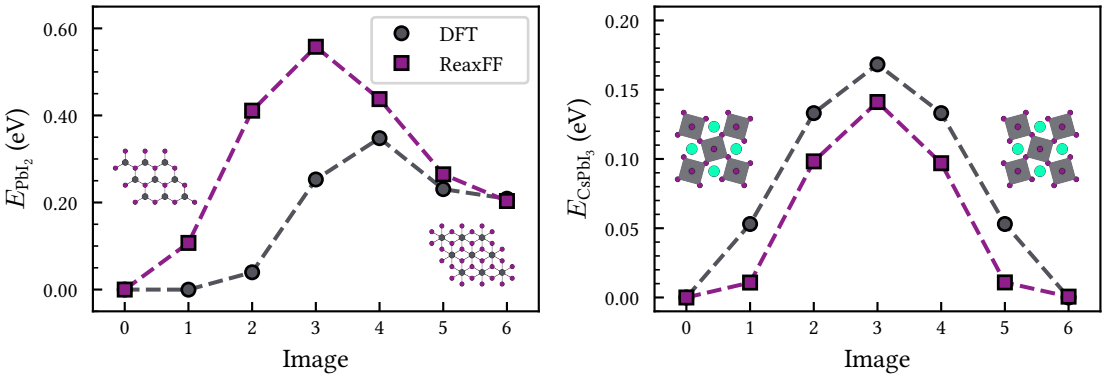

Figure 3.7. Comparison of the energy barrier as predicted with DFT and ReaxFF calculations with (left) the phase transition of monolayer PbI_2 from the 1H- to 1T-phase and (right) the inversion of the octahedral tilting in tetragonal $CsPbI_3$. The reported energies are per formula unit.

4. Reactive molecular dynamics simulations

With a set of ReaxFF parameters for $CsPbI_3$ and its precursors, PbI_2 and CsI, created in the previous chapter, we are now equipped to do ReaxFF simulations for metal halide perovskites. Here we cover the results that we obtained from reactive molecular dynamics (MD) simulations with our best ReaxFF parameter set in `AMS2020` [70]. The chapter is divided into two sections, each aimed at answering part of the research question we postulated in chapter 1. In section 4.1 we focus on the dynamical processes in perovskites and their role in the evolution of various phases. To do so, we took a closer look at the different perovskite phases of $CsPbI_3$ by constructing a phase diagram as a function of temperature and closely examine the lattice and cation dynamics as well as their internal structure. The results in section 4.2 are centred around reactive processes in perovskites and their impact on material stability. To this end, we explored the diffusion of iodine point defects in $CsPbI_3$ and examined their impact on the degradation reactions in $CsPbI_3$ under thermal stress.

4.1. Dynamical processes in $CsPbI_3$

4.1.1. Phase transitions in metal halide perovskites

Our first investigation with ReaxFF into the dynamical properties of metal halide perovskites was focused on the phase diagram of $CsPbI_3$ in which we monitored the evolution of the lattice vector as a function of the temperature. The reactive molecular dynamics simulations were carried out in an *NpT*-ensemble with a simulation timestep of $\Delta t = 0.25\,\text{fs}$ and at atmospheric pressure. Before the production run was started, we allowed each system to equilibrate to the initial conditions for a duration of $t_{\text{eq}} = 50\,\text{ps}$. During this equilibration run we used a Berendsen thermostat and barostat [52] with damping constants of $\tau_{\text{T}} = 100\,\text{fs}$ and $\tau_{\text{p}} = 2500\,\text{fs}$, respectively. The production run was then started from the atom positions and velocities of the final equilibration frame. In the production run, we made use of a NHC-thermostat [57] of chain length 10 and MTK-barostat [58], with a similar coupling as during the equilibration run ($\tau_{\text{T}} = 100\,\text{fs}$; $\tau_{\text{p}} = 2500\,\text{fs}$).

To generate the phase diagram, we carried out simulations in an orthorhombic $4\times4\times4$-supercell of $CsPbI_3$ with 256 stoichiometric units at a range of constant temperatures, from $T = 100\,\text{K}$ to 700 K. A temperature spacing of 50 K was used, except for the region where phase transitions

were expected to occur, in that temperature range we decreased the spacing to 25 K. The total simulation duration for all temperatures was $t_{\text{prod}} = 0.5$ ns. The unit cell dimensions were extracted by fitting Gaussian distributions to the lattice vectors during the simulation, the exact method of which is outlined in Appendix D. These lattice vectors were converted to the pseudo-cubic lattice vectors, of which the definition can be found in Appendix A, to allow for a comparison between the different perovskite phases. An overview of the pseudo-cubic lattice vectors as a function of temperature is shown in Figure 4.1. We named this approach the *discrete*-approach to the phase diagram.

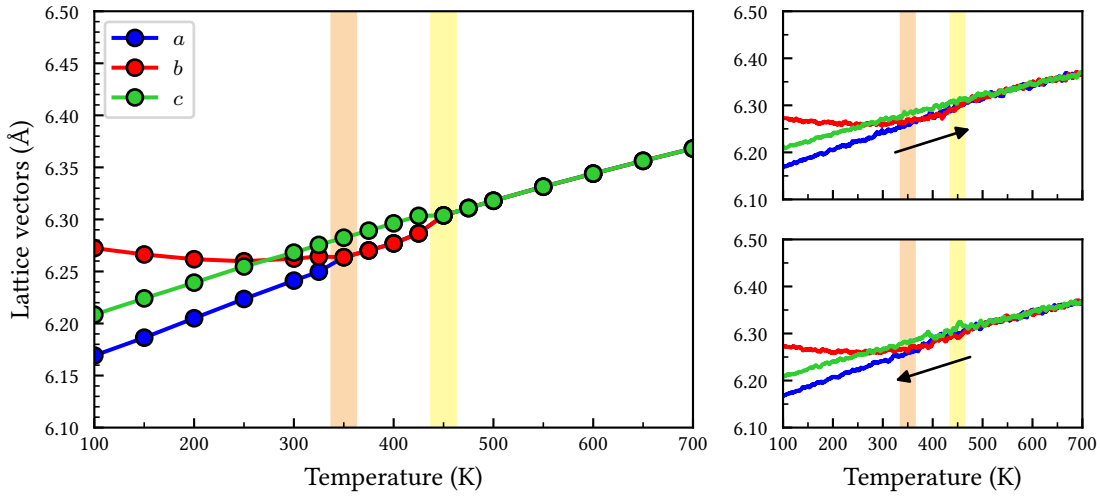

Figure 4.1. The phase diagram of $CsPbI_3$ using ReaxFF simulations. The pseudo-cubic lattice constants are plot for (left) the *discrete*-approach, (top right) a heating *temperature ramp*-approach and (bottom right) a cooling *temperature ramp*-approach. The orange and yellow bars indicate the phase transitions from orthorhombic to tetragonal and tetragonal to cubic phase transition of $CsPbI_3$, respectively.

First of all, based on the evolution of the lattice vectors[15], we conclude that the $CsPbI_3$ perovskite adopts an orthorhombic, tetragonal and cubic phase from low to progressively higher temperatures, which is supported by experimental investigations into the phase evolution of $CsPbI_3$ [99, 100]. The phase transition temperatures that we observe for our model system are $T_{\text{T}\to\text{C}} = (450 \pm 10)$ K for the cubic to tetragonal phase transition and $T_{\text{O}\to\text{T}} = (350 \pm 10)$ K for the phase transition between the tetragonal and orthorhombic phase. In comparison to experiments, that predict $T_{\text{T}\to\text{C}} = 554$ K and $T_{\text{O}\to\text{T}} = 457$ K, respectively [99], our ReaxFF simulation underestimate the phase transition temperature with about 100 K. We attribute this underestimation of the phase transition temperature to the slight overprediction of the lattice parameters and thus unit cell volume during our simulations (1%) in comparison to experimental results [99]. This is in part the result of training the ReaxFF parameters against reference data that accurately describes the unit cell geometries at a finite temperature.

[15] Note that here we made use of the relative size of the pseudo-cubic lattice vectors to distinguish between the different perovskite phases. These are: $(a \neq b \neq c)$, $(a = b \neq c)$ and $(a = b = c)$, for the orthorhombic, tetragonal and cubic phase of $CsPbI_3$, respectively.

Furthermore, we probed the thermal behaviour of our ReaxFF simulations through the volumetric expansion coefficient. The value we find from our ReaxFF simulations is $12.9 \cdot 10^{-5}\,\mathrm{K}^{-1}$. This value is found to agree well with values obtained from X-ray diffraction experiments: $11.8 \cdot 10^{-5}\,\mathrm{K}^{-1}$ [101] and $15.3 \cdot 10^{-5}\,\mathrm{K}^{-1}$ [99].

To check the reversibility of the phase evolution of $CsPbI_3$, we generated phase diagrams in which the material system was subjected to a continuously changing temperature, i.e. gradually heating up or cooling down the system. We named this method the *temperature ramp*-approach to the phase diagram. In the *temperature ramp*-approach, we used a $6 \times 6 \times 6$-supercell of orthorhombic $CsPbI_3$ with 864 stoichiometric units. Similar to the *discrete*-approach, temperatures between $T = 100\,\mathrm{K}$ and $700\,\mathrm{K}$ were investigated, and the heating or cooling of the system was carried out by using a linear temperature control over the NHC-thermostat during the production run. The temperate ramp rate of the thermostat was set to $\frac{dT}{dt} = 5.0 \cdot 10^{-4}\,\mathrm{K\,fs}^{-1}$, which resulted in a total simulation time of $t_{\mathrm{prod}} = 1.2\,\mathrm{ns}$ for a full heating or cooling run. The data shown in Figure 4.1 is smoothed with a running average over 10 ps since it was somewhat noisy as a result of the stochastic nature of the MD simulations. Again we made use of the pseudo-cubic lattice vectors as the data in this phase diagram. An analysis of the sensitivity of the phase diagrams to the MD simulation parameters, which include the type of thermostat and barostat as well as the temperature ramp rate, showed the same phase evolution for $CsPbI_3$ at similar phase transition temperatures demonstrating the phase evolution is robust against changes in the MD parameters. Moreover, the analysis revealed that the data smoothing did not influence the phase transition temperatures. The full sensitivity analysis can be found in Appendix E.

A comparison of the phase diagram as generated with the *discrete*- and *temperature ramp*-approach demonstrates that both approaches to the phase diagram result in a similar phase evolution for $CsPbI_3$. For both approaches, the same perovskite phases and phase transition temperatures are observed. A comparison of the two phase diagrams that make use of the continuous heating or cooling of the system reveals that our force field also captures the reversibility of the perovskite phase transitions. During both the gradual heating and cooling of the perovskite system, in the continuous phase diagram of Figure 4.1, we observe that the systems nicely transform from orthorhombic, to tetragonal to cubic upon heating and the reverse order upon cooling at the same transition temperatures. It should be noted that our ReaxFF simulations did not always show complete reversibility of the lattice vectors; an example of such a cooling run is found in Appendix E. We can attribute this incomplete reversibility of the lattice vectors to the formation of an orthorhombic structure that consists of multiple differently oriented domains. As a result of fluctuations in the lattice during the cooling process, domains of the simulation cell can get stuck in domains with a different orthorhombic orientation, which they are unable to overcome due to insufficient thermal energy. A similar formation of domains in the orthorhombic phase of $CsPbI_3$ has also been observed in experiments [102, 103].

4.1.2. The internal structure of metal halide perovskites

Lattice anharmonicity in $CsPbI_3$

After having investigated the overall phase behaviour of $CsPbI_3$, we now turn our focus to the lattice and ion dynamics. Here, we look into the anharmonic nature of the $CsPbI_3$ perovskite lattice, which was recently investigated by Liu *et al.* via neutron scattering experiments [104]. Since their investigation of the anharmonicity of the lattice was limited to the cubic phase, we employ our ReaxFF description of $CsPbI_3$ to obtain insights into the anharmonicity and its implications on the different perovskite phases and their stability. In this analysis, we follow the method outlined by Carignano *et al.* [105]. The method uses the probability distribution $P(\delta)$ of the geometrical parameter δ that serves as an indicator for the lattice anharmonicity. This parameter is defined as the shortest distance of an I atom from the straight line that interconnects the two neighbouring Pb atoms (see Figure 4.2). A subsequent comparison of this probability distribution from simulations against the one derived from the harmonic approximation (see Appendix F) allows for a qualitative description of the anharmonicity of the metal halide framework and thus the perovskite lattice.

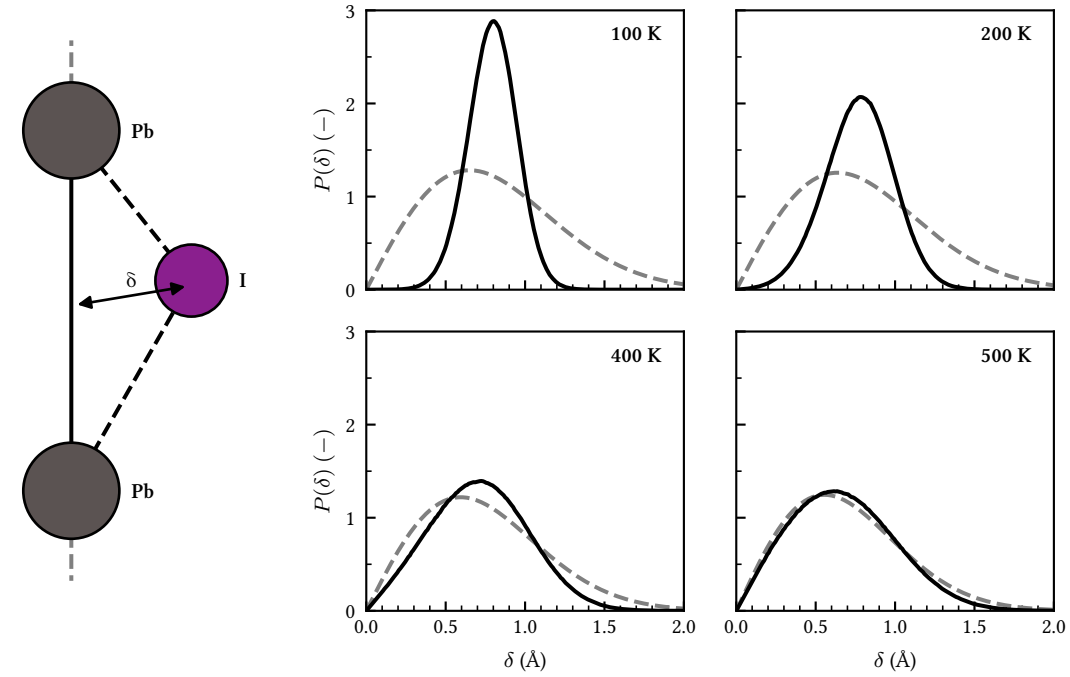

Figure 4.2. (Left): An illustration of the geometrical parameter δ used to characterize the anharmonicity of the lattice $CsPbI_3$. (Right): An overview of the probability distributions for the displacement δ at different temperatures, with the data from ReaxFF simulations (solid lines) and the best fits for a harmonic approximation (dashed lines) both included.

The set of simulations carried out for this analysis used the same settings in the equilibration and production run as during the *discrete*-approach to the phase diagram shown in Figure 4.1.

The simulated temperatures were 100 K, 200 K, 400 K and 500 K, allowing us to explore the anharmonicity in all three phases, i.e. orthorhombic, tetragonal and cubic phase of $CsPbI_3$. All simulations were started from a $4 \times 4 \times 4$-supercell of orthorhombic $CsPbI_3$ with 256 stoichiometric units. Each system was equilibrated to its target temperature for $t_{eq} = 50$ ps. The subsequent production runs were $t_{prod} = 2$ ns, which resulted in sufficiently good statistics for the construction of a probability distribution.

Figure 4.2 shows the comparison between the probability distribution obtained from ReaxFF simulations and the best fit of the harmonic approximation. The comparison shows that the harmonic approximation to the probability distribution breaks down at each simulated temperature. This is evidenced by the fact that the distribution of the geometrical parameter δ is shifted to larger values than in the harmonic approximation. This demonstrates that the ReaxFF simulations predict that the perovskite structure shows characteristics of anharmonicity for all three phases. The presence of this anharmonicity can be rationalized through Goldschmidt's principle of maximum anion contact [106, 107]. The tilting of the octahedra combined with a shift of the cation position allows for better contact between the Cs ion and the surrounding iodide ions, stabilizing the perovskite structure through distortions in the metal halide framework in comparison to the non-distorted structure. Interestingly, such distortions are highly sensitive to the temperature. A comparison of the four temperatures (100 K - 500 K) shows that the highest degree of anharmonicity is found at the lowest temperature and it gradually decreases as the temperature approaches 500 K. We note that the decrease in the anharmonicity as a function of temperature is in line with the increase of the system symmetry when going from the orthorhombic to tetragonal and eventually cubic phase of $CsPbI_3$.

Next, we zoom in on the atomistic implications of this anharmonicity of the $CsPbI_3$ lattice on the different perovskite structures. To this end, we compare instantaneous snapshots from the simulation of the perovskite against its 2 ns time-averaged structure to find out what effect this anharmonicity has on the final perovskite structure. We use the structures from ReaxFF simulations performed at 200 K and 500 K, an overview of which is shown in Figure 4.3. At each of the temperature, we compare the instantaneous structure with a time-averaged structure. At any frame of the simulation, the instantaneous structure is a structure that contains many local distortions. These distortions result from the lattice anharmonicity, which, as rationalized earlier, brings a slight stabilization of the perovskite structure.

At low temperature, i.e. 200 K, both structures are relatively ordered and resemble closely with each other. The main characteristics of both structures are a single octahedra tilting pattern, resulting from stabilization effect from the anharmonic metal halide framework. In contrast, at high temperature 500 K, the instantaneous structure is a structure that contains many local distortions, whereas the time average structure is a highly symmetric perovskite phase, namely, the experimentally reported cubic structure of $CsPbI_3$ ($Pm\bar{3}m$ space group) [99, 101, 104]. The distinct features of the two can be understood by considering the relatively large thermal energy at the elevated temperatures. At high temperatures the thermal energy is high enough to over-

come the stabilization effect from the tilting of the octahedra, thus allowing the $CsPbI_3$ perovskite structure to fluctuate between many locally distorted structures. As a result of the rapid fluctuations between the many distorted structures, the time-averaged structure becomes cubic. To summarize, the formation of the cubic phase of $CsPbI_3$ at high temperature is a combined result of lattice anharmonicity (the tilting of the octahedra to accommodate the internal strain caused by ion size mismatch) and entropy-driven dynamical disorder (thermally activated fluctuations between locally distorted structures).

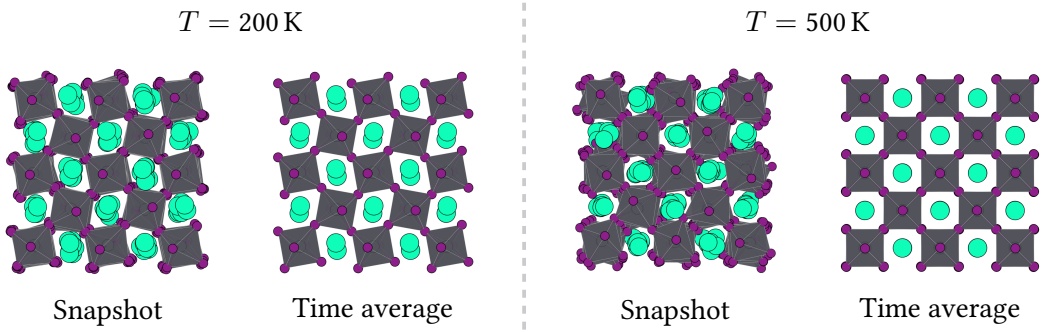

Figure 4.3. A comparison of the perovskite structure in the final frame of the simulation against the 2 ns time-averaged structure of $CsPbI_3$ as obtained from ReaxFF simulations at (left) 200 K and (right) 500 K.

Radial distribution function

To obtain a more detailed understanding of the internal structure of the perovskite at a low and high temperature, we next carry out a more refined structural analysis of the basis of radial distribution functions (RDFs). The RDF captures the probability for atoms of one type to be a certain distance r from the same or other types of atoms. Different material phases result in different structural footprints, which allows us to distinguish between the different structures. The RDF between atoms of type a and those of type b, g_{ab} is defined as

$$g_{ab}\left(r\right) = \frac{1}{N_a N_b} \sum_{i=1}^{N_a} \sum_{j=1}^{N_b} \langle \delta\left(|\vec{r}_i - \vec{r}_j| - r\right) \rangle, \tag{4.1}$$

where we sum over Dirac delta functions $\delta\left(\cdots\right)$ that are offset by the interatomic distance between the i^{th} atom of type a located at $\vec{r}_i$ and the j^{th} atom of type b located at $\vec{r}_j$, with $\langle \cdots \rangle$ indicating an ensemble average [108]. The summation covers all possible atom pairs by running over all N_a and N_b atoms of, respectively, type a and b.

We applied the RDF analysis to perovskite systems at two different temperatures: 200 K and 500 K, which based on our previous findings correspond to the orthorhombic and cubic structure of $CsPbI_3$, respectively. To benchmark our simulations against experimental results, we compared the results against the peak locations obtained for $CsPbI_3$ unit cells that were determined from X-ray diffraction measurements by Marronnier *et al.* [99]. Since the phase diagram revealed an

underestimation of the phase transitions temperatures by our ReaxFF parameter set with respect to experiments, the comparison is shown in Figure 4.4, compares the simulations at 200 K and 500 K against the experimental structures at 325 K ($Pnma$ space group) and 645 K ($Pm\bar{3}m$ space group), respectively.

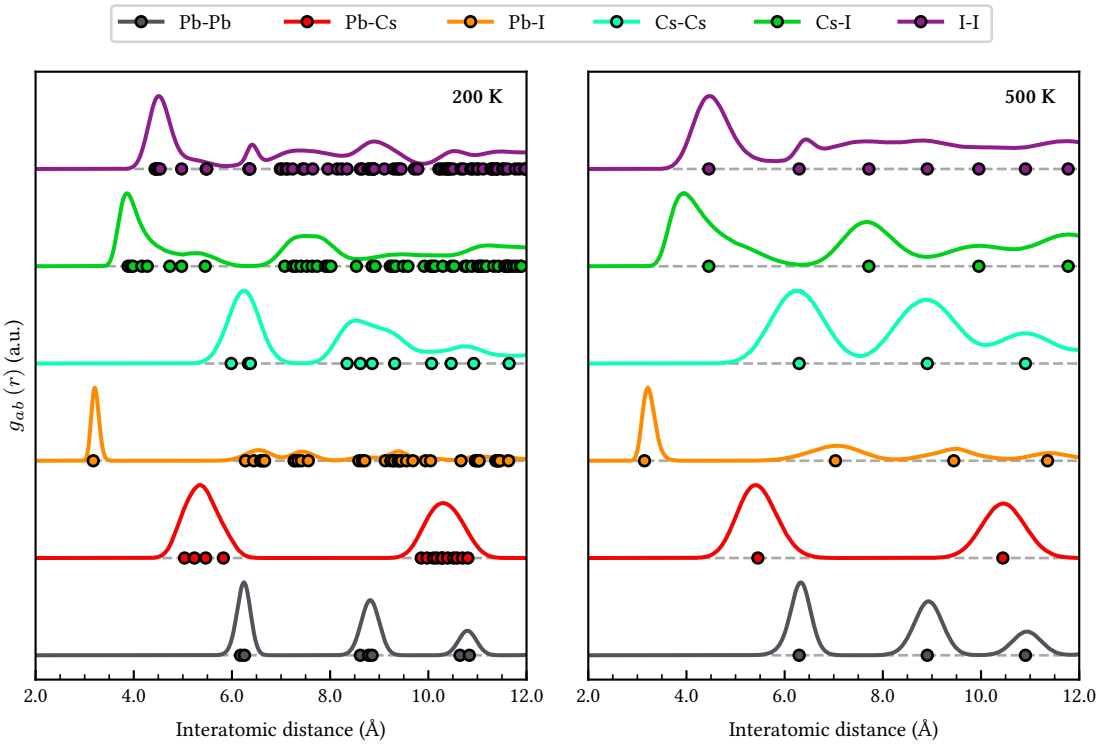

Figure 4.4. The radial distribution function (RDF) as obtained from ReaxFF simulations of $CsPbI_3$ at 200 K and 500 K (solid lines) together with the location of the respective peaks as found in XRD experiments [99] (circles).

Upon inspection of the RDFs, it shows that the peak positions of the simulations coincide well with those obtained from experiments. This is clearly demonstrated by the splitting or broadening of the secondary peaks[16] in the RDF of Pb-I, Cs-Cs and Cs-I, which is seen in both the simulations and experiments when going from high (500 K) to low (200 K) temperatures. However, there is a discrepancy between ReaxFF simulated RDF at 500 K and the experimental results. The first peak in the RDF for Cs-I from the simulations is downshifted approximately 0.6 Å. This mismatch can be explained by the fact that the diffraction experiments only capture the time-averaged structures. In contrast, our simulations also contain information on the instantaneous locally distorted structures of the cubic perovskite phase. As explained earlier, these distortions stabilize the structure by bringing the I anions closer to the Cs cations, thus resulting in a downshift of the first peak in the RDF for Cs-I. All in all, the comparison of the RDFs confirms that our ReaxFF simulations match with experimental reports [99, 100], properly dis-

[16] The secondary peaks we refer to are seen in the RDF as generated at 500 K at an interatomic distance of approximately 7.0 Å for Pb-I, 8.4 Å for Cs-Cs and 7.6 Å for Cs-I.

tinguishing between the high- and low-temperature phases of black phase $CsPbI_3$ by predicting an orthorhombic ($Pnma$) and cubic ($Pm\bar{3}m$) structure at 200 K and 500 K, respectively.

Inorganic cation dynamics

Finally, we explore the internal dynamics of the Cs cation in $CsPbI_3$. To do so, we created spatial heatmaps of the Cs cations at a variety of temperatures. Although a $4 \times 4 \times 4$-supercell of orthorhombic $CsPbI_3$ contains a total of 256 Cs cations, not all were used in the construction of the heatmap. The orthorhombic unit cell contains four distinct sites for Cs, each with a different orientation of the metal halide framework around it. We therefore only used the equivalent Cs cations that were located in one type of those cation sites, and thus the heatmaps with a total of 64 cations. The time-averaged position for each cation was set as the origin so that the data for the cations belonging to the same group could be added together. The normalized xy-projected heatmaps from the full 2 ns ReaxFF simulations are shown in Figure 4.5.

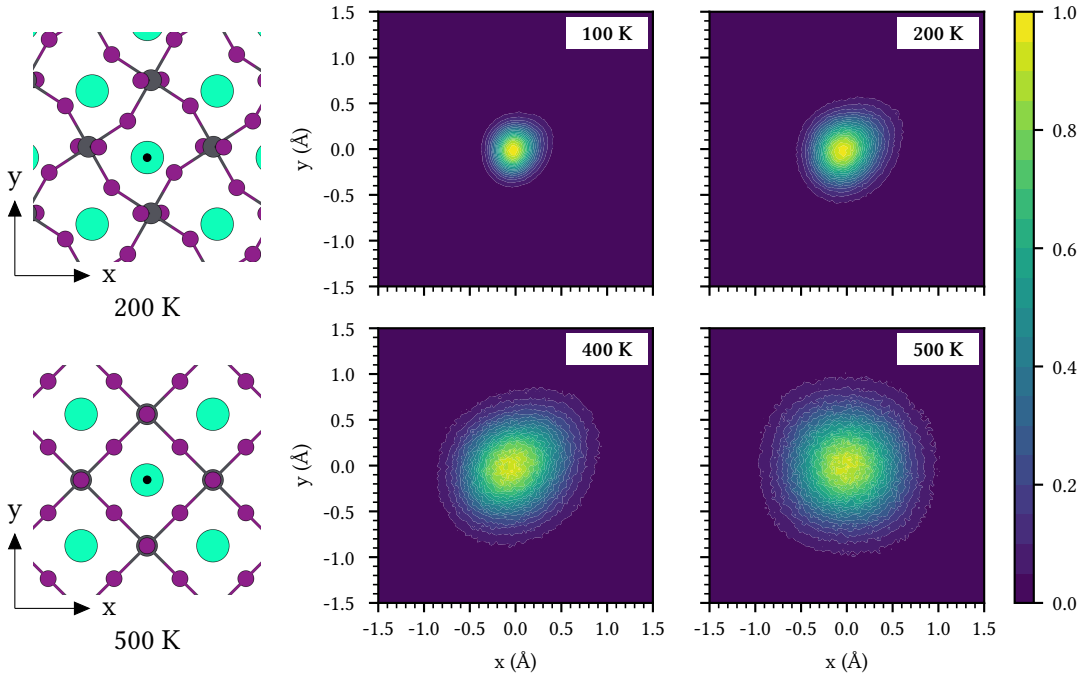

Figure 4.5. (Left): A schematic overview of the cation in the perovskite structures at 200 K and 500 K, the average position of the cation is indicated by a black dot. (Right): An overview of the normalized heatmaps for the Cs cation in $CsPbI_3$ at different temperatures.

An inspection of the cation heatmaps for the $CsPbI_3$ reveals that the spatial distribution of the Cs cations in the perovskite is a rather complex function of the temperature. At low temperatures (100 K - 400 K) we observe a very specific directionality in the dynamics of the equivalent Cs cations with preferential movement in the positive xy-direction. The directional cation dynamics are attributed to the anharmonicity of the perovskite crystal lattice that we established earlier

in this thesis. The anharmonicity of the perovskite lattice causes the metal halide cages to be 'frozen' in a distorted configuration at low temperatures (see Figure 4.5), to which the motion of the cations is confined. Earlier we rationalized that this distortion, combined with a shift of the equilibrium position of the Cs cation, results in better contacts between the iodide anions and the Cs cations, causing a stabilization of the perovskite structure. Despite this stabilization, the low temperature phases of $CsPbI_3$ are known to be metastable, showing a phase instability in which the perovskite structure of $CsPbI_3$ converts to the non-perovskite yellow phase [99, 100, 102, 109]. Straus *et al.* showed from single-crystal X-ray diffraction measurements that the movement away from the equilibrium position of the Cs cation, resulting in worse contacts between the iodide anions and the Cs cation, destabilizing the perovskite structure. On that basis, we suggest that the observed directional motion of the cation in the 'frozen' distorted perovskite structures during our simulations is likely the cause of the phase instability of the low-temperature perovskite phases of $CsPbI_3$. The progression through the temperatures shows that this destabilizing motion of the cation is increasingly more present as the temperature is raised (from 100 K to 400 K), which can be explained based on the thermal energy available to the Cs cations for movement.

However, a closer look at the heatmap of the high-temperature simulation at 500 K, where ReaxFF predicts the cubic phase of $CsPbI_3$, demonstrates that the dynamics of the Cs cations have lost any preferential direction, resulting in an isotropic heatmap. This observation can readily be explained by the lack of any long-time local structure due to the rapid fluctuations of the metal halide framework. On average, these fluctuations result in a cubic cage geometry (see Figure 4.5). At these elevated temperatures, the cubic phase of $CsPbI_3$ is found to be stabilized [99, 101]. We hypothesize that the rapid thermal fluctuations of the metal halide framework and the lack of any long-time local ordering make it so that the motion of the Cs cations does not result in a destabilization of the structure. Instead, we suggest that the rapid thermal fluctuations of both the Cs cations and iodide anions result in good contacts between the two, resulting in a stabilization of the perovskite phase at high temperatures.

4.2. Reactive processes in $CsPbI_3$

4.2.1. Defect-assisted ion migration

The investigation of the reactive processes with ReaxFF is centred around the migration of ions, which, as stated before, impacts the stability of perovskites. A considerable number of experimental and computational results in the literature demonstrate that the ion migration in metal halide perovskite is mainly due to the migration of halide ions [110–113]. Investigations with DFT demonstrated that two types of mechanisms, a vacancy-assisted and interstitial-assisted, have similar energy barriers for iodine migration in $MAPbI_3$ [114, 115]. Moreover, using MD

simulations, it was established that the diffusion of iodine in $MAPbI_3$ occurs through both migration mechanisms [116]. Since detailed investigations of these two mechanisms have been limited to $MAPbI_3$, we employed our ReaxFF parameter set to study the competition of the two mechanisms of iodine migration in $CsPbI_3$. We investigated the importance of both iodine migration mechanisms, of which a schematic overview is shown in Figure 4.6, by determining the temperature evolution of the self-diffusion coefficients of iodine point defects.

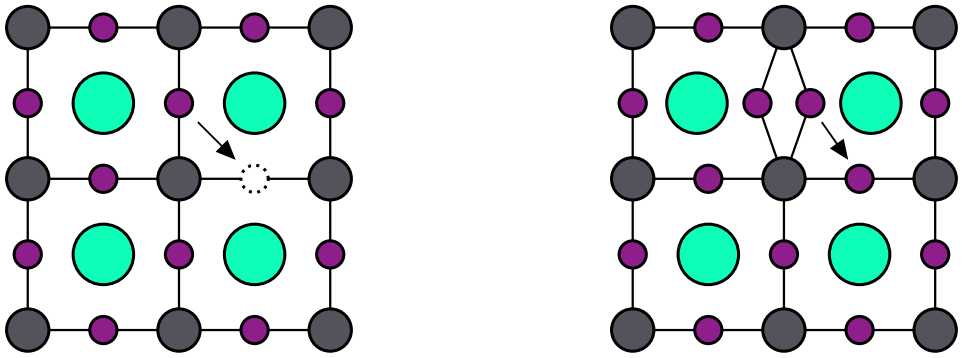

Figure 4.6. An illustration of the different defect-assisted migration pathways for iodine in $CsPbI_3$. (Left): Vacancy-assisted migration. (Right): Interstitial-assisted migration.

The model perovskite system used in these simulations was an orthorhombic $4 \times 4 \times 4$-supercell of $CsPbI_3$ which contained 256 formal units. Point defects were created in the supercell by either removing or inserting iodine atoms, respectively, creating vacancies (V_I) or interstitials (I_I) in the model system. For both, the vacancies and interstitials, two defects were used to study the dynamical behaviour of the iodine point defects in $CsPbI_3$. To ensure a homogeneous spread of the defects throughout the model system, the defects were created at least 20 Å apart. For each model system, the geometry was optimized and then subjected to a two-step equilibration process, which allowed the model systems to attain their target temperature and pressure in an *NpT*-ensemble. The first step, aimed at a quick equilibration, made use of a Berendsen thermostat and Berendsen barostat [52], whereas the second equilibration step, used to realize a realistic starting point for the production runs, was carried out with an NHC-thermostat [57] and MTK-barostat [58]. Equilibration times of $t_{eq} = 50\,ps$ and $100\,ps$ were used in each equilibration step for the systems containing vacancies and interstitials, respectively. Once equilibrated, the production run of each system was started from the final frame of the equilibration process and carried out in a *NVT*-ensemble in which the temperature was controlled by an NHC-thermostat[17]. In the production runs a total time of simulation of $t_{prod} = 2.0\,ns$ and $2.5\,ns$ was used for the systems containing vacancies and interstitials, respectively. The timestep was set to $\Delta t = 0.25\,fs$ in each simulation. A similar coupling with the thermostats and barostats was used as seen earlier: $\tau_T = 100\,fs$ and $\tau_p = 2500\,fs$, where for the NHC-thermostat we again used a chain

[17] No barostat was included in the production runs because it is found that the volume rescaling used by a barostat influences the self-diffusion coefficients obtained from simulations [117]. In contrast, an NHC-thermostat is found to a negligible impact on the self-diffusion coefficients [118] and is therefore still used in the production runs.

length of 10.

The self-diffusion coefficients of the point defects were obtained from the atom trajectories using the Einstein method[18] to the diffusion coefficient [119]. In this method, the self-diffusion coefficient D of a species is computed using the following relation

$$D = \frac{1}{2d} \lim_{t\to\infty} \frac{d}{dt} \mathrm{MSD}\,(t)\,, \tag{4.2}$$

in which d is the dimensionality of the system and $\mathrm{MSD}\,(t)$ the mean square displacement of that species over time. The MSD can be obtained from the atom trajectories using

$$\mathrm{MSD}\,(t) = \left\langle \frac{1}{N} \sum_{i=1}^{N} \left|\vec{r}_i\,(t) - \vec{r}_i\,(0)\right|^2 \right\rangle_{t_0}, \tag{4.3}$$

where $\vec{r}_i\,(t)$ and $\vec{r}_i\,(0)$ are the time evolved and reference position of the i^{th} instance of the species of interest, respectively. It should be stressed that in order to reduce the fluctuations, thus enhancing the accuracy with which the self-diffusion coefficients are determined, both a time- ($\langle\cdots\rangle_{t_0}$) and ensemble-average ($\frac{1}{N}\sum_{i=1}^{N}$) are used [117]. The temperature evolution of the self-diffusion coefficients was investigated on the basis of an Arrhenius relation [120]. For diffusion the Arrhenius relation gives the following temperature dependent form of the diffusion coefficient

$$D = D_0 \cdot \exp\left(-\frac{E_\mathrm{a}}{k_\mathrm{B}T}\right), \tag{4.4}$$

in which D_0 is the prefactor, E_a is the energy barrier for migration and k_B the Boltzmann constant.

Table 4.1. Diffusion constants of two iodine interstitials (I_I) and two iodine vacancies (V_I) as obtained from ReaxFF simulations for $CsPbI_3$ with the average values in brackets.

T (K)	$D_{\mathrm{I_I}}$ ($\cdot$ 10^{-9} cm^2 s^{-1})	$D_{\mathrm{V_I}}$ ($\cdot$ 10^{-9} cm^2 s^{-1})
450	1.76, 1.88, 2.48; **[2.04]**	2.31, 2.62; **[2.47]**
500	3.72, 5.15, 5.35; **[4.74]**	3.13, 3.66, 4.02, 4.88 5.02, 5.65; **[4.39]**
550	7.66, 9.32, 10.92; **[9.30]**	4.03, 5.73, 7.84; **[5.87]**
600	13.47, 13.85, 14.57; **[13.97]**	7.95, 9.77; **[8.86]**
650	18.38, 18.97, 20.26; **[19.21]**	-
700	26.69, 27.87, 30.45; **[28.34]**	-

Multiple defect simulations were carried out at each temperature in the range of 450 K to 700 K at atmospheric pressure. In this temperature range, we were able to investigate the temperature evolution of the diffusion coefficients in the cubic phase of $CsPbI_3$, without needing to take into

[18] Another method that can be used for the determination of the self-diffusion coefficients is the Green-Kubo method, which makes use of velocity autocorrelation functions [117]. However, the Green-Kubo method is generally regarded as a less robust method. Therefore we employ the Einstein method to analyze the simulation results.

account the effects of a phase transition in the metal halide perovskite. Using simulations snapshots, it was confirmed that the defect geometries predicted by our ReaxFF parameter set match well with those seen in previous static DFT calculations on iodine point defects in $CsPbI_3$ [121–123]. Moreover, a comparison of simulations in a defective supercell of $CsPbI_3$ with those in a perfectly crystalline lattice demonstrated that iodine migration only occurred in the presence of the iodine defects. The temperature dependence of the average diffusion coefficients from the simulations, as well as some snapshots of the defects at 500 K, are shown in Figure 4.7. A complete overview of the diffusion coefficients is found in,Table 4.1 and an explanation of the analysis method is given in Appendix G. It should be noted that during some high-temperature simulations (600 K and higher) of $CsPbI_3$ with vacancies, degradation of the perovskite structure was observed. These simulations required an in-depth analysis that can be found in the next section.

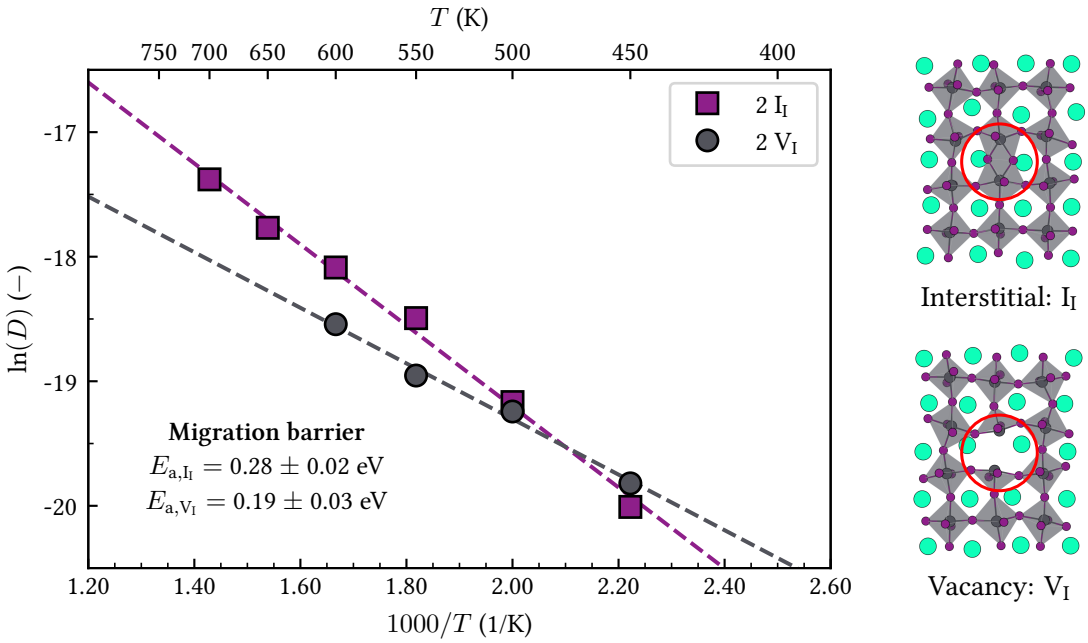

Figure 4.7. (Left): An Arrhenius plot of the temperature dependence of the self-diffusion coefficients of two point defects in $CsPbI_3$. (Right): Structural snapshots of an iodine interstitial (I_I) and an iodine vacancy (V_I) are shown for ReaxFF simulations at 500 K.

The temperature evolution of the self-diffusion coefficients of both types of iodine point defects was well-described by a single Arrhenius relation. As a result of this, we can associate the migration of both point defects with a single activation energy over the entire investigated temperature range. The migration barrier for the iodine interstitials was determined at $E_{a,I_I} = (0.28 \pm 0.02)$ eV and for the iodine vacancies at $E_{a,V_I} = (0.19 \pm 0.03)$ eV. The relatively low energy barrier for the migration of iodine vacancies from ReaxFF (0.19 eV) agrees well with the low vacancy migration barriers found with DFT calculations in literature (0.17 - 0.36 eV) [111, 124]. Because of a lack of experimental measurements for the migration barrier of iodine defects in $CsPbI_3$, we compared the activation energy for the migration of iodine vacancies against the activation energies for vacancy migration in other inorganic metal halide perovskites as measured

by Mizusaki *et al.* using impedance spectroscopy [125]. Our value for iodine is slightly smaller than those reported for $CsPbBr_3$ 0.29 eV and $CsPbCl_3$ 0.25 eV.

From the Arrhenius analysis, we also obtained the prefactors for the diffusion of iodine interstitials is $D_{0,\mathrm{I_I}} = 3.1 \cdot 10^{-6}\,\mathrm{cm^2\,s^{-1}}$ and for iodine vacancies is $D_{0,\mathrm{V_I}} = 3.6 \cdot 10^{-7}\,\mathrm{cm^2\,s^{-1}}$. The larger prefactor for iodine interstitials can be explained by the larger configurational entropy gain associated with the migration of an iodine interstitial compared to the migration of an iodine vacancy. When an iodine interstitial is present, two iodide anions occupy a site that a single iodide anion would usually occupy. The dumbbell configuration of the two iodides is thus rather rigid (see Figure 4.7) [116]. The subsequent migration of one of the iodide anions in the interstitial defect is thus associated with a significant gain in the configurational entropy, therefore, based on transition state theory, resulting in a relatively large prefactor [126].

A comparison of the self-diffusion coefficients for the two types of defects reveals that both types of point defects in the investigated temperature range exhibit similar diffusivities. Near the low end of the temperature range ($T = 450\,\mathrm{K}$ to $500\,\mathrm{K}$), similar diffusion coefficients were found. For the higher temperatures ($T > 600\,\mathrm{K}$) the interstitial diffusion would become twice as diffusive. Nevertheless, the similarity of the diffusion coefficients indicates that at these temperatures, both mechanisms (see Figure 4.6) are relevant and important in facilitating the migration of iodine in the cubic phase of $CsPbI_3$.

4.2.2. Defect-accelerated perovskite decomposition

As mentioned in the previous section, the high-temperature simulations of $CsPbI_3$ containing vacancies tended to result in the decomposition of the perovskite structure. Therefore, we here focus on the dynamical mechanism of the degradation of perovskites triggered by such defects.

To do so, we analyzed the atom trajectories from a ReaxFF simulation that exhibited this such perovskite decomposition. Snapshots of the atom trajectories of a simulation at 600 K are shown in Figure 4.8. These snapshots give a general overview of the different stages during the decomposition of the perovskite structures. In this figure, the first snapshot is 125 ps into the simulation, with the second and third snapshots, respectively, being 0.3 ns and 0.9 ns later. The time axis of the figures is a shifted time axis of which the origin is set to the first snapshot. At the beginning of the simulation, the two iodine vacancies are dispersed throughout the perovskite phase of $CsPbI_3$, with one of them highlighted in the left panel of Figure 4.8. After some time, we observe that a lead species close to the iodine vacancy moves away from its position in the lattice, shown in the middle panel in Figure 4.8. As a result of the formation of such a lead interstitial, the perovskite structure now has a locally Pb/I-rich domain. Further evolution of the system causes the Pb/I-rich region to grow in size, resulting in the formation of a Pb/I cluster, highlighted in the right panel of Figure 4.8.

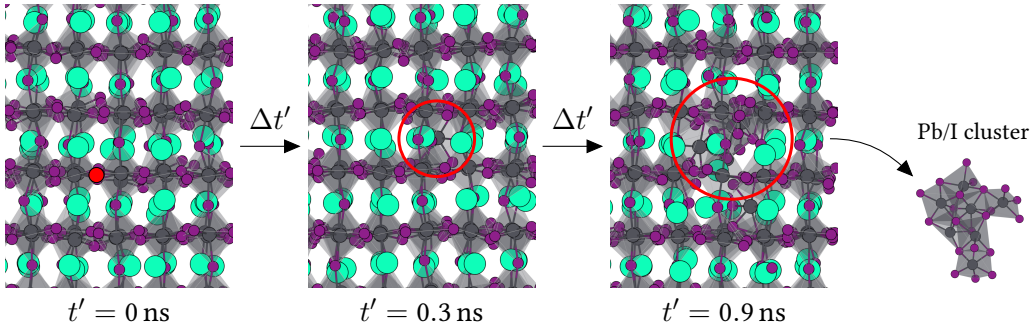

Figure 4.8. Snapshots of a simulation of $CsPbI_3$ containing two iodine vacancies (V_I) at 600 K. (Left): The location of the iodine vacancy at the beginning of the simulation indicated with the red dot. (Middle): Formation of a lead interstitial near the iodide vacancy, forming a Pb/I-rich domain. (Right): Decomposition of the perovskite structure through the growth of a Pb/I cluster.

To elucidate the details of the degradation mechanism of the perovskite, we zoom into the initial stages of the perovskite decomposition in Figure 4.9. This figure contains the snapshot taken 125 ps into the simulation, with two additional snapshots taken 52.5 ps and 100 ps later, again a shifted time axis is used. For a straightforward identification of the atomic movements, a two-dimensional schematic illustration is also provided for each of the snapshots, for clarity the Cs cations were not included in this schematic.

The snapshots clearly show that the degradation process starts with the formation of an iodine Frenkel defect near the existing iodine vacancy; one iodine atom leaves its original position to create a vacancy and at the same time form one iodine interstitial site. As a result of this, two octahedra are connected through the newly formed interstitial site, forming a complex of edge-sharing metal halide octahedra. Two iodine vacancies are present close to this complex of edge-sharing octahedra, causing the octahedra in this complex to only be weakly bound to the rest of the perovskite lattice. Consequently, the lead species of either of these octahedra easily move away from their position in the lattice, forming a face-sharing Pb_xI_y complex. As discussed earlier in Figure 4.8, this Pb_xI_y complex eventually grows to a larger Pb/I cluster.

We speculate that such face-sharing Pb_xI_y complexes serve as a nucleation centre for the decomposition of the perovskite structure, eventually leading to the decomposition of the metal halide perovskite into PbI_2. Our finding corroborates with recent observations from transmission electron microscopy experiments by Manekkathodi *et al.*, where PbI_2 nanoparticles are detected in the vicinity of lattice defects such as grain boundaries [127]. We note that although our analysis is based on a simulation done at 600 K, similar reaction pathways were seen during other high-temperature simulations of 600 K or higher. Moreover, we suggest that the investigation of the effect of grain boundaries merits its own separate study.

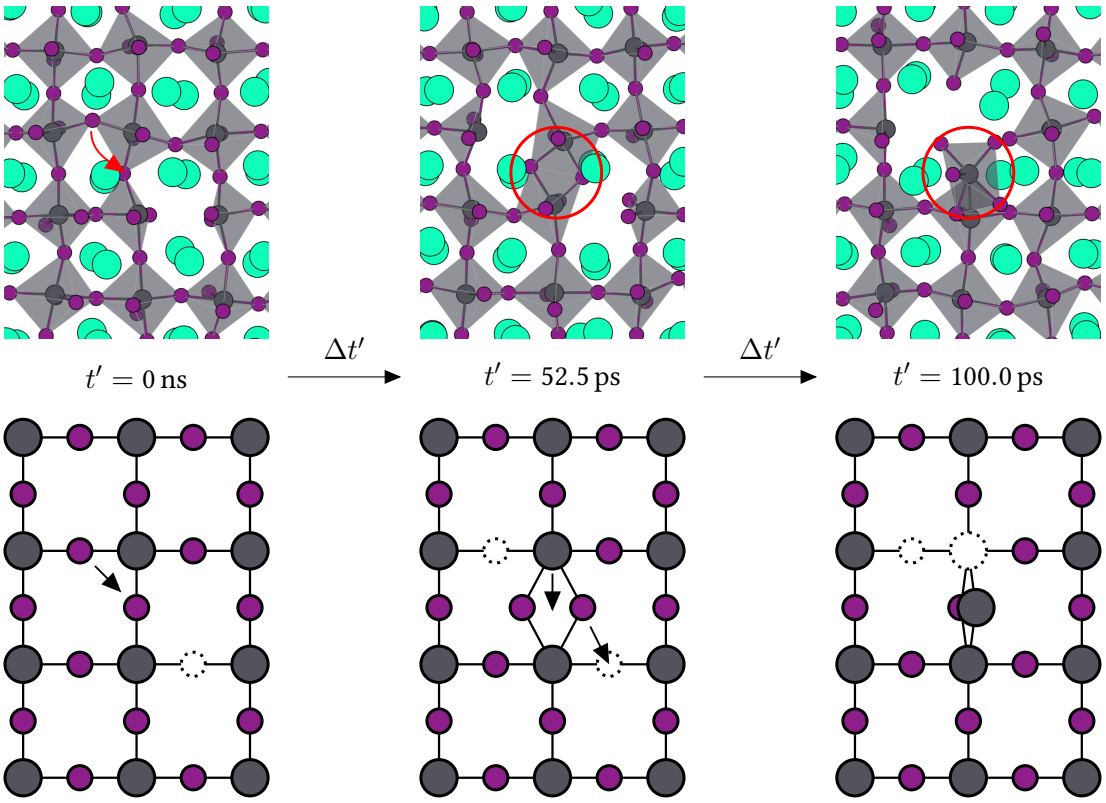

Figure 4.9. Detailed snapshots of the formation of a lead interstitial defect during a 600 K simulation of $CsPbI_3$ with an additional schematic drawing of the degradation mechanism. For clarity the Cs species were not drawn in this schematic. (Left): An iodine vacancy in $CsPbI_3$. (Middle): Formation of an iodine Frenkel defect, resulting in a small complex of edge-sharing metal halide octahedra that is only weakly bound to the rest of the perovskite lattice. (Right): Formation of an interstitial lead atom causing a Pb_xI_y complex to form.

5. Conclusion and outlook

In this thesis, we present the first set of ReaxFF parameters for inorganic metal halide perovskite $CsPbI_3$ for reactive molecular dynamics simulations. Using this ReaxFF description, we study the dynamical and reactive processes that impact the stability of metal halide perovskites. Here we briefly state the most important findings of this work and provide suggestions for how this work can be expanded upon on future work.

Using a Monte Carlo-based parameterization algorithm, we train a ReaxFF force field for $CsPbI_3$ against a set of accurate quantum mechanical reference data from DFT calculations. Several validation tests, including the bulk modulus, structural optimizations and activation barriers for phase transitions, confirm that our ReaxFF force field correctly captures the material's behaviour and has a reasonably accurate predictive power.

With the help of molecular dynamics simulations with our ReaxFF parameter set, we study a variety of dynamical processes relevant to the stability of metal halide perovskites, particularly $CsPbI_3$. To this end, we show the phase evolution of the perovskite as a function of temperature. The transition temperatures from orthorhombic to tetragonal and tetragonal to cubic phase are predicted with reasonable accuracy compared to the experiments. Combining several analyses, such as RDFs, the lattice dynamics of the metal halide framework and a heatmap of the Cs cation, we show that the anharmonicity of the crystal lattice of metal halide perovskites combined with thermal entropy are responsible for the phase transitions between the different perovskite phases.

Additionally, we explain the phase instability of$CsPbI_3$ by linking the anharmonicity of the metal halide framework to the dynamical behaviour of the Cs cations. We suggest that the Cs cations prefer to locate at positions for better contact with the metal halide framework. However, due to the dynamical nature, Cs cations tend to move away from this preferential position, resulting in an instability that causes the conversion of the perovskite phase to the non-perovskite phase. At elevated temperatures, we observe that thermal entropy starts to dominate the system, causing rapid fluctuations in both the metal halide framework and the cation dynamics. As a result of this, we propose that the destabilizing effect of the cation dynamics is replaced by a stabilizing effect, which causes the perovskite phase of $CsPbI_3$ to be stabilized at elevated temperatures, which is in line with experiments.

Finally, we employ ReaxFF simulations to investigate the reactive processes that impact the stability of metal halide perovskites, which include the migration of ions and defect-assisted degra-

dation reactions. We show that both iodine interstitials and iodine vacancies play a large role in the migration of ions in $CsPbI_3$. Furthermore, we find that particularly the iodine vacancies are detrimental to the stability of metal halide perovskites. By combining the results from our simulations with experimental observations, we propose a degradation mechanism at the atomic scale for the $CsPbI_3$ perovskite to PbI_2 clusters. We find that Pb/I-rich domains are readily formed in the vicinity of iodine vacancies in the perovskite lattice through iodine Frenkel defects. Further growth of these domains causes the formation of a Pb/I cluster, which results in the decomposition of the perovskite lattice.

Outlook

From the above, using $CsPbI_3$ as an example, we demonstrate that molecular dynamics simulations using ReaxFF are a valuable tool for studying the stability of halide perovskites. We expect that the set of ReaxFF parameters from this work, focusing on the inorganic perovskite $CsPbI_3$, can readily be expanded to a broader range of metal halide perovskites or combined with other materials that are found in perovskite solar cells.

Although we have carried out various simulations, our efforts with the current ReaxFF parameter set are in no way exhaustive. Most simulations in this thesis were done for single crystalline bulk systems of metal halide perovskites. However, experimental absorber layers in perovskite solar cells are of often polycrystalline nature with many grain boundaries [128]. The long-term stability of the perovskite solar cells is largely impacted by these grain boundaries, as they are shown to be sensitive to thermal stress [129, 130] as well as the diffusion of and reaction with environmental species such as water [130, 131]. DFT calculations have been employed to study grain boundaries in metal halide perovskites before. However these calculations have always been limited to small model systems and short timescales [132–135]. Here, ReaxFF simulations would be ideal to complement DFT in studying dynamical processes at the grain boundaries, since they can access much longer timescales and model larger, more realistic systems. This would also allow us to access the complex interplay of several defects on the grain boundaries and their exchange and reactions with those in the bulk.

Furthermore, by expanding the ReaxFF parameters of $CsPbI_3$ to other perovskite compositions, we would be able to simulate state-of-the-art experimental compositions. Nowadays, the best performing perovskite solar cells, those with an efficiency of higher than 20% and reasonable long-term stability, use a perovskite composition with multiple different cations, multiple types of halide ions, or both [136–138]. It is found that by using these mixed compositions, the perovskite crystallinity, stability and charge carrier dynamics in the perovskite absorber layer are improved substantially, resulting in an overall better performance [139]. An expansion of our ReaxFF parameter set would therefore be the inclusion of a description of the organic cations and the other halide ions encountered in experimental recipes. Such an extended parameter set would allow us to understand why some compositions are more stable than the others and thus

potentially aid the synthesis of more stable perovskite films that can result in long-term stable perovskite solar cells.

Whereas our current focus is on the absorber layer, the operational efficiency and stability of a perovskite solar cell are determined by more than just the perovskite layer. Indeed, the interfaces between different layers in the perovskite solar cell are another decisive factor that greatly influences the operational characteristics, i.e. efficiency and long-term stability of perovskite solar cells [140]. Thus, an extension of the ReaxFF parameter set towards the materials frequently contacted with the absorber layer (the electron and hole transport layer) is also deemed important. ReaxFF simulations of the interface between the perovskite surface and a transport layer should result in an increased understanding of the interplay between the two materials and the effect this has on the operational stability. The most straightforward expansion would be the interface of perovskites with the electron transport layer TiO_2 because of the following reasons. First of all, some TiO_2 ReaxFF parameters are readily available in literature as published by Kim *et al.* [141], therefore slightly reducing the required training effort. Besides, ample existing DFT and experimental data are available for the training and validation of the ReaxFF parameters. Finally, several research questions can be readily identified, all relevant to the reactive processes involving halide species and oxygen vacancies [142, 143]. Such questions include: What is the role of ion migration/exchange at the interface? What type of degradation reactions are at play at the interface? Can redox reactions (the reaction of oxygen species with halogen species) occur?

6. Acknowledgements

The work in this thesis is the result of 11 months of *hard* work. If you had asked me at the start of this project to predict how these months would have panned out, surely I would not have said my entire Master's thesis would be completed from home. Then again, no one would make that prediction a year ago. Although the past year was spent behind the same desk each day, the time was a rather joyful one, which I would like to thank a couple of people for.

Shuxia, thanks for the general time and support you had for me during this project. Whenever I felt something needed to be discussed, you were available to discuss matters. Although I always tended to stretch the meetings a bit longer than originally planned, you were patient enough to let me ramble on about the simulations I was doing and gave me plenty of advice and suggestions to work with. This really helped to shape the project to its current form. **Ivo**, thanks for your willingness to join in on the project and supervise me from the chemistry side. Metal halide perovskites as a material are rather distant from the materials you deal with daily; however, your expertise and numerous suggestions related to the computational methods helped me a lot during the project.

Adri, your help and expertise have been crucial in creating the metal halide perovskite ReaxFF parameter set. Although we could not meet in person, the countless suggestions made by you through video calls and emails were more than enough to get some nice results with ReaxFF simulations, thanks for that. **Vicent**, as the project progressed, I became increasingly dependent on your knowledge of molecular dynamics simulations. You were readily available whenever I had questions about what setting to use or what tests to run to create sensible results with my ReaxFF simulations. Without your help, the simulations results would not have been what they are now, thanks.

Furthermore, I would like to thank the people I have interacted with a lot over the past year. These are the members of the **MSM** and **IMC STudent Meeting**. Whereas everyone was confined to their home, the ability to chat with others through a weekly get-together, albeit online, is refreshing and a nice distraction from the daily routine of working on this thesis.

Finally, I would like to thank all of my graduation committee members for their time and willingness to be involved in the final stages of my Master's education. **Shuxia**, **Ivo**, **Emiel**, **Sofía** and **Kees** I hope all of you were able to find interesting topics in this thesis, in which I tried my best to give a broad but focused overview of the application of ReaxFF to metal halide perovskites.

References

[1] H. J. Snaith. "Perovskites: The Emergence of a New Era for Low-Cost, High-Efficiency Solar Cells". In: *The Journal of Physical Chemistry Letters* 4.21 (2013), pp. 3623–3630. DOI: 10.1021/jz4020162.

[2] C. Battaglia, A. Cuevas, and S. D. Wolf. "High-efficiency crystalline silicon solar cells: status and perspectives". In: *Energy & Environmental Science* 9.5 (2016), pp. 1552–1576. DOI: 10.1039/C5EE03380B.

[3] J. Liu, Y. Yao, S. Xiao, and X. Gu. "Review of status developments of high-efficiency crystalline silicon solar cells". In: *Journal of Physics D: Applied Physics* 51.12 (2018), p. 123001. DOI: 10.1088/1361-6463/aaac6d.

[4] M. A. Green, A. Ho-Baillie, and H. J. Snaith. "The emergence of perovskite solar cells". In: *Nature Photonics* 8.7 (2014), pp. 506–514. DOI: 10.1038/nphoton.2014.134.

[5] H. J. Snaith. "Present status and future prospects of perovskite photovoltaics". In: *Nature Materials* 17.5 (2018), pp. 372–376. DOI: 10.1038/s41563-018-0071-z.

[6] A. Kojima, K. Teshima, Y. Shirai, and T. Miyasaka. "Organometal Halide Perovskites as Visible-Light Sensitizers for Photovoltaic Cells". In: *Journal of the American Chemical Society* 131.17 (2009), pp. 6050–6051. DOI: 10.1021/ja809598r.

[7] National Renewable Energy Laboratory. *Best Research-Cell Efficiencies*. Best Research-Cell Efficiency Chart. 2020. URL: https://www.nrel.gov/pv/cell-efficiency.html (visited on 01/26/2021).

[8] G. Niu, X. Guo, and L. Wang. "Review of recent progress in chemical stability of perovskite solar cells". In: *Journal of Materials Chemistry A* 3.17 (2015), pp. 8970–8980. DOI: 10.1039/C4TA04994B.

[9] D. Wang, M. Wright, N. K. Elumalai, and A. Uddin. "Stability of perovskite solar cells". In: *Solar Energy Materials and Solar Cells* 147 (2016), pp. 255–275. DOI: 10.1016/j.solmat.2015.12.025.

[10] J.-P. Correa-Baena, M. Saliba, T. Buonassisi, M. Grätzel, A. Abate, W. Tress, and A. Hagfeldt. "Promises and challenges of perovskite solar cells". In: *Science* 358.6364 (2017), pp. 739–744. DOI: 10.1126/science.aam6323.

[11] B.-w. Park and S. I. Seok. "Intrinsic Instability of Inorganic–Organic Hybrid Halide Perovskite Materials". In: *Advanced Materials* 31.20 (2019), p. 1805337. DOI: 10.1002/adma.201805337.

[12] A. Poglitsch and D. Weber. "Dynamic disorder in methylammoniumtrihalogenoplumbates (II) observed by millimeter-wave spectroscopy". In: *The Journal of Chemical Physics* 87.11 (1987), pp. 6373–6378. DOI: 10.1063/1.453467.

[13] H. Mashiyama, Y. Kurihara, and T. Azetsu. "Disordered Cubic Perovskite Structure of $CH_3NH_3PbX_3$ (X = Cl, Br, I)." In: *Journal of the Korean Physical Society* 32 (Supplement Pt.1 1998), pp. 156–158.

[14] J. Feng. "Mechanical properties of hybrid organic-inorganic $CH_3NH_3BX_3$ (B = Sn, Pb; X = Br, I) perovskites for solar cell absorbers". In: *APL Materials* 2.8 (2014), p. 081801. DOI: 10.1063/1.4885256.

[15] S. Sun, Y. Fang, G. Kieslich, T. J. White, and A. K. Cheetham. "Mechanical properties of organic–inorganic halide perovskites, $CH_3NH_3PbX_3$ (X = I, Br and Cl), by nanoindentation". In: *Journal of Materials Chemistry A* 3.36 (2015), pp. 18450–18455. DOI: 10.1039/C5TA03331D.

[16] Z. Qiu, N. Li, Z. Huang, Q. Chen, and H. Zhou. "Recent Advances in Improving Phase Stability of Perovskite Solar Cells". In: *Small Methods* 4.5 (2020), p. 1900877. DOI: 10.1002/smtd.201900877.

[17] S. D. Stranks, V. M. Burlakov, T. Leijtens, J. M. Ball, A. Goriely, and H. J. Snaith. "Recombination Kinetics in Organic-Inorganic Perovskites: Excitons, Free Charge, and Subgap States". In: *Physical Review Applied* 2.3 (2014), p. 034007. DOI: 10.1103/PhysRevApplied.2.034007.

[18] S. Draguta, S. Thakur, Y. V. Morozov, Y. Wang, J. S. Manser, P. V. Kamat, and M. Kuno. "Spatially Non-uniform Trap State Densities in Solution-Processed Hybrid Perovskite Thin Films". In: *The Journal of Physical Chemistry Letters* 7.4 (2016), pp. 715–721. DOI: 10.1021/acs.jpclett.5b02888.

[19] J. Kang and L.-W. Wang. "High Defect Tolerance in Lead Halide Perovskite $CsPbBr_3$". In: *The Journal of Physical Chemistry Letters* 8.2 (2017), pp. 489–493. DOI: 10.1021/acs.jpclett.6b02800.

[20] D. Meggiolaro, S. G. Motti, E. Mosconi, A. J. Barker, J. Ball, C. A. R. Perini, F. Deschler, A. Petrozza, and F. D. Angelis. "Iodine chemistry determines the defect tolerance of lead-halide perovskites". In: *Energy & Environmental Science* 11.3 (2018), pp. 702–713. DOI: 10.1039/C8EE00124C.

[21] J. Carrillo, A. Guerrero, S. Rahimnejad, O. Almora, I. Zarazua, E. Mas-Marza, J. Bisquert, and G. Garcia-Belmonte. "Ionic Reactivity at Contacts and Aging of Methylammonium Lead Triiodide Perovskite Solar Cells". In: *Advanced Energy Materials* 6.9 (2016), p. 1502246. DOI: https://doi.org/10.1002/aenm.201502246.

[22] J. Li, Q. Dong, N. Li, and L. Wang. "Direct Evidence of Ion Diffusion for the Silver-Electrode-Induced Thermal Degradation of Inverted Perovskite Solar Cells". In: *Advanced Energy Materials* 7.14 (2017), p. 1602922. DOI: https://doi.org/10.1002/aenm.201602922.

[23] D. D. Girolamo et al. “Ion Migration-Induced Amorphization and Phase Segregation as a Degradation Mechanism in Planar Perovskite Solar Cells”. In: *Advanced Energy Materials* 10.25 (2020), p. 2000310. DOI: 10.1002/aenm.202000310.

[24] N.-K. Kim et al. “Investigation of Thermally Induced Degradation in $CH_3NH_3PbI_3$ Perovskite Solar Cells using In-situ Synchrotron Radiation Analysis”. In: *Scientific Reports* 7.1 (2017), p. 4645. DOI: 10.1038/s41598-017-04690-w.

[25] M. Salado, L. Contreras-Bernal, L. Caliò, A. Todinova, C. López-Santos, S. Ahmad, A. Borras, J. Idígoras, and J. A. Anta. “Impact of moisture on efficiency-determining electronic processes in perovskite solar cells”. In: *Journal of Materials Chemistry A* 5.22 (2017), pp. 10917–10927. DOI: 10.1039/C7TA02264F.

[26] G. Abdelmageed, L. Jewell, K. Hellier, L. Seymour, B. Luo, F. Bridges, J. Z. Zhang, and S. Carter. “Mechanisms for light induced degradation in $MAPbI_3$ perovskite thin films and solar cells”. In: *Applied Physics Letters* 109.23 (2016), p. 233905. DOI: 10.1063/1.4967840.

[27] Q. Fu, X. Tang, B. Huang, T. Hu, L. Tan, L. Chen, and Y. Chen. “Recent Progress on the Long-Term Stability of Perovskite Solar Cells”. In: *Advanced Science* 5.5 (2018), p. 1700387. DOI: 10.1002/advs.201700387.

[28] J. M. Frost and A. Walsh. “What Is Moving in Hybrid Halide Perovskite Solar Cells?” In: *Accounts of Chemical Research* 49.3 (2016), pp. 528–535. DOI: 10.1021/acs.accounts.5b00431.

[29] E. Mosconi, J. M. Azpiroz, and F. De Angelis. “Ab Initio Molecular Dynamics Simulations of Methylammonium Lead Iodide Perovskite Degradation by Water”. In: *Chemistry of Materials* 27.13 (2015), pp. 4885–4892. DOI: 10.1021/acs.chemmater.5b01991.

[30] L. Zhang and P. H.-L. Sit. “Ab initio static and dynamic study of $CH_3NH_3PbI_3$ degradation in the presence of water, hydroxyl radicals, and hydroxide ions”. In: *RSC Advances* 6.80 (2016), pp. 76938–76947. DOI: 10.1039/C6RA12781A.

[31] C. Zheng and O. Rubel. “Unraveling the Water Degradation Mechanism of $CH_3NH_3PbI_3$”. In: *The Journal of Physical Chemistry C* 123.32 (2019), pp. 19385–19394. DOI: 10.1021/acs.jpcc.9b05516.

[32] A. Mattoni, A. Filippetti, M. I. Saba, and P. Delugas. “Methylammonium Rotational Dynamics in Lead Halide Perovskite by Classical Molecular Dynamics: The Role of Temperature”. In: *The Journal of Physical Chemistry C* 119.30 (2015), pp. 17421–17428. DOI: 10.1021/acs.jpcc.5b04283.

[33] C. Caddeo, M. I. Saba, S. Meloni, A. Filippetti, and A. Mattoni. “Collective Molecular Mechanisms in the $CH_3NH_3PbI_3$ Dissolution by Liquid Water”. In: *ACS Nano* 11.9 (2017), pp. 9183–9190. DOI: 10.1021/acsnano.7b04116.

[34] 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". In: *Journal of Materials Chemistry A* 8.23 (2020), pp. 11824–11836. DOI: 10.1039/D0TA03200J.

[35] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard. "ReaxFF: A Reactive Force Field for Hydrocarbons". In: *The Journal of Physical Chemistry A* 105.41 (2001), pp. 9396–9409. DOI: 10.1021/jp004368u.

[36] T. P. Senftle et al. "The ReaxFF reactive force-field: development, applications and future directions". In: *npj Computational Materials* 2.1 (2016), pp. 1–14. DOI: 10.1038/npjcompumats.2015.11.

[37] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Revised ed. edition. Mineola, N.Y: Dover Publications, 1996. 480 pp.

[38] M. Born and R. Oppenheimer. "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389.20 (1927), pp. 457–484. DOI: 10.1002/andp.19273892002.

[39] R. O. Jones. "Density functional theory: Its origins, rise to prominence, and future". In: *Reviews of Modern Physics* 87.3 (2015), pp. 897–923. DOI: 10.1103/RevModPhys.87.897.

[40] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Physical Review* 136.3 (1964), B864–B871. DOI: 10.1103/PhysRev.136.B864.

[41] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Physical Review* 140.4 (1965), A1133–A1138. DOI: 10.1103/PhysRev.140.A1133.

[42] J. P. Perdew and K. Schmidt. "Jacob's ladder of density functional approximations for the exchange-correlation energy". In: *AIP Conference Proceedings* 577.1 (2001), pp. 1–20. DOI: 10.1063/1.1390175.

[43] J. P. Perdew. "Climbing the ladder of density functional approximations". In: *MRS Bulletin* 38.9 (2013), pp. 743–750. DOI: 10.1557/mrs.2013.178.

[44] D. Frenkel and B. Smit. *Understanding Molecular Simulation*. Elsevier, 2002. DOI: 10.1016/B978-0-12-267351-1.X5000-7.

[45] M. P. Allen. "Introduction to Molecular Dynamics Simulation". In: *Computational Soft Matter: From Synthetic Polymers to Proteins*. Vol. 23. NIC Series. Jülich: John von Neumann Institute for Computing, 2004, pp. 1–28.

[46] B. R. Brooks, R. E. Bruccoleri, B. D. Olafson, D. J. States, S. Swaminathan, and M. Karplus. "CHARMM: A program for macromolecular energy, minimization, and dynamics calculations". In: *Journal of Computational Chemistry* 4.2 (1983), pp. 187–217. DOI: 10.1002/jcc.540040211.

[47] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman. "A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules". In: *Journal of the American Chemical Society* 117.19 (1995), pp. 5179–5197. DOI: 10.1021/ja00124a002.

[48] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson. "A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters". In: *The Journal of Chemical Physics* 76.1 (1982), pp. 637–649. DOI: 10.1063/1.442716.

[49] D. Chandler. *Introduction to Modern Statistical Mechanics*. 1st edition. New York: Oxford University Press, 1987. 288 pp.

[50] E. Braun, J. Gilmer, H. B. Mayes, D. L. Mobley, J. I. Monroe, S. Prasad, and D. M. Zuckerman. "Best Practices for Foundations in Molecular Simulations [Article v1.0]". In: *Living Journal of Computational Molecular Science* 1.1 (2018), p. 5957. DOI: 10.33011/livecoms.1.1.5957.

[51] P. H. Hünenberger. "Thermostat Algorithms for Molecular Dynamics Simulations". In: *Advanced Computer Simulation: Approaches for Soft Matter Sciences I*. Ed. by C. Dr. Holm and K. Prof. Dr. Kremer. Advances in Polymer Science. Berlin, Heidelberg: Springer, 2005, pp. 105–149. DOI: 10.1007/b99427.

[52] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. "Molecular dynamics with coupling to an external bath". In: *The Journal of Chemical Physics* 81.8 (1984), pp. 3684–3690. DOI: 10.1063/1.448118.

[53] S. C. Harvey, R. K.-Z. Tan, and T. E. Cheatham. "The flying ice cube: Velocity rescaling in molecular dynamics leads to violation of energy equipartition". In: *Journal of Computational Chemistry* 19.7 (1998), pp. 726–740. DOI: 10.1002/(SICI)1096-987X(199805)19:7<726::AID-JCC4>3.0.CO;2-S.

[54] E. Braun, S. M. Moosavi, and B. Smit. "Anomalous Effects of Velocity Rescaling Algorithms: The Flying Ice Cube Effect Revisited". In: *Journal of Chemical Theory and Computation* 14.10 (2018), pp. 5262–5272. DOI: 10.1021/acs.jctc.8b00446.

[55] S. Nosé. "A molecular dynamics method for simulations in the canonical ensemble". In: *Molecular Physics* 52.2 (1984), pp. 255–268. DOI: 10.1080/00268978400101201.

[56] W. G. Hoover. "Canonical dynamics: Equilibrium phase-space distributions". In: *Physical Review A* 31.3 (1985), pp. 1695–1697. DOI: 10.1103/PhysRevA.31.1695.

[57] G. J. Martyna, M. L. Klein, and M. Tuckerman. "Nosé–Hoover chains: The canonical ensemble via continuous dynamics". In: *The Journal of Chemical Physics* 97.4 (1992), pp. 2635–2643. DOI: 10.1063/1.463940.

[58] G. J. Martyna, D. J. Tobias, and M. L. Klein. "Constant pressure molecular dynamics algorithms". In: *The Journal of Chemical Physics* 101.5 (1994), pp. 4177–4189. DOI: 10.1063/1.467468.

[59] M. F. Russo and A. C. T. van Duin. "Atomistic-scale simulations of chemical reactions: Bridging from quantum chemistry to engineering". In: *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*. Computer Simulations of Radiation Effects in Solids 269.14 (2011), pp. 1549–1554. DOI: 10.1016/j.nimb.2010.12.053.

[60] A. C. T. van Duin, A. Strachan, S. Stewman, Q. Zhang, X. Xu, and W. A. Goddard. "ReaxFF-SiO Reactive Force Field for Silicon and Silicon Oxide Systems". In: *The Journal of Physical Chemistry A* 107.19 (2003), pp. 3803–3811. DOI: 10.1021/jp0276303.

[61] K. Chenoweth, A. C. T. van Duin, P. Persson, M.-J. Cheng, J. Oxgaard, and W. A. Goddard. "Development and Application of a ReaxFF Reactive Force Field for Oxidative Dehydrogenation on Vanadium Oxide Catalysts". In: *The Journal of Physical Chemistry C* 112.37 (2008), pp. 14645–14654. DOI: 10.1021/jp802134x.

[62] W. J. Mortier, S. K. Ghosh, and S. Shankar. "Electronegativity-equalization method for the calculation of atomic charges in molecules". In: *Journal of the American Chemical Society* 108.15 (1986), pp. 4315–4320. DOI: 10.1021/ja00275a013.

[63] E. Iype, M. Hütter, A. P. J. Jansen, S. V. Nedea, and C. C. M. Rindt. "Parameterization of a reactive force field using a Monte Carlo algorithm". In: *Journal of Computational Chemistry* 34.13 (2013), pp. 1143–1154. DOI: 10.1002/jcc.23246.

[64] M. Dittner, J. Müller, H. M. Aktulga, and B. Hartke. "Efficient global optimization of reactive force-field parameters". In: *Journal of Computational Chemistry* 36.20 (2015), pp. 1550–1561. DOI: 10.1002/jcc.23966.

[65] A. C. T. v. Duin, J. M. A. Baas, and B. v. d. Graaf. "Delft molecular mechanics: a new approach to hydrocarbon force fields. Inclusion of a geometry-dependent charge calculation". In: *Journal of the Chemical Society, Faraday Transactions* 90.19 (1994), pp. 2881–2895. DOI: 10.1039/FT9949002881.

[66] T. Trnka, I. Tvaroška, and J. Koča. "Automated Training of ReaxFF Reactive Force Fields for Energetics of Enzymatic Reactions". In: *Journal of Chemical Theory and Computation* 14.1 (2018), pp. 291–302. DOI: 10.1021/acs.jctc.7b00870.

[67] G. Shchygol, A. Yakovlev, T. Trnka, A. C. T. van Duin, and T. Verstraelen. "ReaxFF Parameter Optimization with Monte-Carlo and Evolutionary Algorithms: Guidelines and Insights". In: *Journal of Chemical Theory and Computation* 15.12 (2019), pp. 6799–6812. DOI: 10.1021/acs.jctc.9b00769.

[68] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. "Optimization by Simulated Annealing". In: *Science* 220.4598 (1983), pp. 671–680. DOI: 10.1126/science.220.4598.671.

[69] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. "Equation of State Calculations by Fast Computing Machines". In: *The Journal of Chemical Physics* 21.6 (1953), pp. 1087–1092. DOI: 10.1063/1.1699114.

[70] R. Rüger, M. Franchini, T. Trnka, A. Yakovlev, E. van Lenthe, P. Philipsen, T. van Vuren, B. Klumpers, and T. Soini. *AMS 2020*. SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands. URL: http://www.scm.com.

[71] A. C. T. v. Duin, W. A. Goddard, M. M. Islam, H. van Schoot, T. Trnka, and A. Yakovlev. *ReaxFF 2020*. SCM, Theoretical Chemistry, Vrije Universiteit, Amsterdam, The Netherlands. URL: http://www.scm.com.

[72] G. Kresse and J. Hafner. "Ab initio molecular dynamics for liquid metals". In: *Physical Review B* 47.1 (1993), pp. 558–561. DOI: 10.1103/PhysRevB.47.558.

[73] G. Kresse and J. Hafner. "Ab initio molecular-dynamics simulation of the liquid-metal–amorphous-semiconductor transition in germanium". In: *Physical Review B* 49.20 (1994), pp. 14251–14269. DOI: 10.1103/PhysRevB.49.14251.

[74] G. Kresse and J. Furthmüller. "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set". In: *Computational Materials Science* 6.1 (1996), pp. 15–50. DOI: 10.1016/0927-0256(96)00008-0.

[75] G. Kresse and J. Furthmüller. "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set". In: *Physical Review B* 54.16 (1996), pp. 11169–11186. DOI: 10.1103/PhysRevB.54.11169.

[76] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized Gradient Approximation Made Simple". In: *Physical Review Letters* 77.18 (1996), pp. 3865–3868. DOI: 10.1103/PhysRevLett.77.3865.

[77] J. P. Perdew, K. Burke, and M. Ernzerhof. "Generalized Gradient Approximation Made Simple [Phys. Rev. Lett. 77, 3865 (1996)]". In: *Physical Review Letters* 78.7 (1997), pp. 1396–1396. DOI: 10.1103/PhysRevLett.78.1396.

[78] S. Grimme, S. Ehrlich, and L. Goerigk. "Effect of the damping function in dispersion corrected density functional theory". In: *Journal of Computational Chemistry* 32.7 (2011), pp. 1456–1465. DOI: 10.1002/jcc.21759.

[79] P. E. Blöchl. "Projector augmented-wave method". In: *Physical Review B* 50.24 (1994), pp. 17953–17979. DOI: 10.1103/PhysRevB.50.17953.

[80] G. Kresse and D. Joubert. "From ultrasoft pseudopotentials to the projector augmented-wave method". In: *Physical Review B* 59.3 (1999), pp. 1758–1775. DOI: 10.1103/PhysRevB.59.1758.

[81] H. J. Monkhorst and J. D. Pack. "Special points for Brillouin-zone integrations". In: *Physical Review B* 13.12 (1976), pp. 5188–5192. DOI: 10.1103/PhysRevB.13.5188.

[82] A. Amat, E. Mosconi, E. Ronca, C. Quarti, P. Umari, M. K. Nazeeruddin, M. Grätzel, and F. De Angelis. "Cation-Induced Band-Gap Tuning in Organohalide Perovskites: Interplay of Spin–Orbit Coupling and Octahedra Tilting". In: *Nano Letters* 14.6 (2014), pp. 3608–3616. DOI: 10.1021/nl5012992.

[83] C. Shen and G. Wang. "Electronic and optical properties of bilayer PbI_2: a first-principles study". In: *Journal of Physics D: Applied Physics* 51.3 (2017), p. 035301. DOI: 10.1088/1361-6463/aa9cd5.

[84] R. X. Yang and L. Z. Tan. "Understanding size dependence of phase stability and band gap in $CsPbI_3$ perovskite nanocrystals". In: *The Journal of Chemical Physics* 152.3 (2020), p. 034702. DOI: 10.1063/1.5128016.

[85] G. Henkelman, A. Arnaldsson, and H. Jónsson. "A fast and robust algorithm for Bader decomposition of charge density". In: *Computational Materials Science* 36.3 (2006), pp. 354–360. DOI: 10.1016/j.commatsci.2005.04.010.

[86] E. Sanville, S. D. Kenny, R. Smith, and G. Henkelman. "Improved grid-based algorithm for Bader charge allocation". In: *Journal of Computational Chemistry* 28.5 (2007), pp. 899–908. DOI: 10.1002/jcc.20575.

[87] W. Tang, E. Sanville, and G. Henkelman. "A grid-based Bader analysis algorithm without lattice bias". In: *Journal of Physics: Condensed Matter* 21.8 (2009), p. 084204. DOI: 10.1088/0953-8984/21/8/084204.

[88] M. Yu and D. R. Trinkle. "Accurate and efficient algorithm for Bader charge integration". In: *The Journal of Chemical Physics* 134.6 (2011), p. 064111. DOI: 10.1063/1.3553716.

[89] G. Henkelman, B. P. Uberuaga, and H. Jónsson. "A climbing image nudged elastic band method for finding saddle points and minimum energy paths". In: *The Journal of Chemical Physics* 113.22 (2000), pp. 9901–9904. DOI: 10.1063/1.1329672.

[90] G. Henkelman and H. Jónsson. "Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points". In: *The Journal of Chemical Physics* 113.22 (2000), pp. 9978–9985. DOI: 10.1063/1.1323224.

[91] M. V. Fedkin et al. "Development of the ReaxFF Methodology for Electrolyte–Water Systems". In: *The Journal of Physical Chemistry A* 123.10 (2019), pp. 2125–2141. DOI: 10.1021/acs.jpca.8b10453.

[92] D. Fantauzzi, J. E. Mueller, L. Sabo, A. C. T. v. Duin, and T. Jacob. "Surface Buckling and Subsurface Oxygen: Atomistic Insights into the Surface Oxidation of Pt(111)". In: *ChemPhysChem* 16.13 (2015), pp. 2797–2802. DOI: 10.1002/cphc.201500527.

[93] M. G. Goesten and R. Hoffmann. "Mirrors of Bonding in Metal Halide Perovskites". In: *Journal of the American Chemical Society* 140.40 (2018), pp. 12996–13010. DOI: 10.1021/jacs.8b08038.

[94] F. D. Murnaghan. "The Compressibility of Media under Extreme Pressures". In: *Proceedings of the National Academy of Sciences* 30.9 (1944), pp. 244–247. DOI: 10.1073/pnas.30.9.244.

[95] F. Birch. "Finite Elastic Strain of Cubic Crystals". In: *Physical Review* 71.11 (1947), pp. 809–824. DOI: 10.1103/PhysRev.71.809.

[96] M. Roknuzzaman, K. Ostrikov, H. Wang, A. Du, and T. Tesfamichael. "Towards lead-free perovskite photovoltaics and optoelectronics by ab-initio simulations". In: *Scientific Reports* 7.1 (2017), p. 14025. DOI: 10.1038/s41598-017-13172-y.

[97] U.-G. Jong, C.-J. Yu, Y.-H. Kye, Y.-S. Kim, C.-H. Kim, and S.-G. Ri. "A first-principles study on the chemical stability of inorganic perovskite solid solutions $Cs_{1-x}Rb_xPbI_3$ at finite temperature and pressure". In: *Journal of Materials Chemistry A* 6.37 (2018), pp. 17994–18002. DOI: 10.1039/C8TA06553E.

[98] C.-J. Yu, U.-H. Ko, S.-G. Hwang, Y.-S. Kim, U.-G. Jong, Y.-H. Kye, and C.-H. Ri. "First-principles study on material properties and stability of inorganic halide perovskite solid solutions $CsPb(I_{1-x}Br_x)_3$". In: *Physical Review Materials* 4.4 (2020), p. 045402. DOI: 10.1103/PhysRevMaterials.4.045402.

[99] A. Marronnier, G. Roma, S. Boyer-Richard, L. Pedesseau, J.-M. Jancu, Y. Bonnassieux, C. Katan, C. C. Stoumpos, M. G. Kanatzidis, and J. Even. "Anharmonicity and Disorder in the Black Phases of Cesium Lead Iodide Used for Stable Inorganic Perovskite Solar Cells". In: *ACS Nano* 12.4 (2018), pp. 3477–3486. DOI: 10.1021/acsnano.8b00267.

[100] C. C. Stoumpos and M. G. Kanatzidis. "The Renaissance of Halide Perovskites and Their Evolution as Emerging Semiconductors". In: *Accounts of Chemical Research* 48.10 (2015), pp. 2791–2802. DOI: 10.1021/acs.accounts.5b00229.

[101] D. M. Trots and S. V. Myagkota. "High-temperature structural evolution of caesium and rubidium triiodoplumbates". In: *Journal of Physics and Chemistry of Solids* 69.10 (2008), pp. 2520–2526. DOI: 10.1016/j.jpcs.2008.05.007.

[102] C. C. Stoumpos, C. D. Malliakas, and M. G. Kanatzidis. "Semiconducting Tin and Lead Iodide Perovskites with Organic Cations: Phase Transitions, High Mobilities, and Near-Infrared Photoluminescent Properties". In: *Inorganic Chemistry* 52.15 (2013), pp. 9019–9038. DOI: 10.1021/ic401215x.

[103] F. Bertolotti, L. Protesescu, M. V. Kovalenko, S. Yakunin, A. Cervellino, S. J. L. Billinge, M. W. Terban, J. S. Pedersen, N. Masciocchi, and A. Guagliardi. "Coherent Nanotwins and Dynamic Disorder in Cesium Lead Halide Perovskite Nanocrystals". In: *ACS Nano* 11.4 (2017), pp. 3819–3831. DOI: 10.1021/acsnano.7b00017.

[104] J. Liu, A. E. Phillips, D. A. Keen, and M. T. Dove. "Thermal Disorder and Bond Anharmonicity in Cesium Lead Iodide Studied by Neutron Total Scattering and the Reverse Monte Carlo Method". In: *The Journal of Physical Chemistry C* 123.24 (2019), pp. 14934–14940. DOI: 10.1021/acs.jpcc.9b02936.

[105] M. A. Carignano, S. A. Aravindh, I. S. Roqan, J. Even, and C. Katan. "Critical Fluctuations and Anharmonicity in Lead Iodide Perovskites from Molecular Dynamics Supercell Simulations". In: *The Journal of Physical Chemistry C* 121.38 (2017), pp. 20729–20738. DOI: 10.1021/acs.jpcc.7b08220.

[106] V. M. Goldschmidt. "Crystal structure and chemical constitution". In: *Transactions of the Faraday Society* 25.0 (1929), pp. 253–283. DOI: 10.1039/TF9292500253.

[107] D. B. Straus, S. Guo, A. M. Abeykoon, and R. J. Cava. "Understanding the Instability of the Halide Perovskite $CsPbI_3$ through Temperature-Dependent Structural Analysis". In: *Advanced Materials* 32.32 (2020), p. 2001069. DOI: 10.1002/adma.202001069.

[108] S. Billinge. "Local Structure from Total Scattering and Atomic Pair Distribution Function (PDF) Analysis". In: *Powder Diffraction: Theory and Practice*. 2008, pp. 464–493. DOI: 10.1039/9781847558237-00464.

[109] G. E. Eperon, G. M. Paternò, R. J. Sutton, A. Zampetti, A. A. Haghighirad, F. Cacialli, and H. J. Snaith. "Inorganic caesium lead iodide perovskite solar cells". In: *Journal of Materials Chemistry A* 3.39 (2015), pp. 19688–19695. DOI: 10.1039/C5TA06398A.

[110] C. Eames, J. M. Frost, P. R. F. Barnes, B. C. O'Regan, A. Walsh, and M. S. Islam. "Ionic transport in hybrid lead iodide perovskite solar cells". In: *Nature Communications* 6.1 (2015), p. 7497. DOI: 10.1038/ncomms8497.

[111] S. Meloni et al. "Ionic polarization-induced current–voltage hysteresis in $CH_3NH_3PbX_3$ perovskite solar cells". In: *Nature Communications* 7.1 (2016), p. 10334. DOI: 10.1038/ncomms10334.

[112] T.-Y. Yang, G. Gregori, N. Pellet, M. Grätzel, and J. Maier. "The Significance of Ion Conduction in a Hybrid Organic–Inorganic Lead-Iodide-Based Perovskite Photosensitizer". In: *Angewandte Chemie International Edition* 54.27 (2015), pp. 7905–7910. DOI: 10.1002/anie.201500014.

[113] C. Besleaga et al. "Iodine Migration and Degradation of Perovskite Solar Cells Enhanced by Metallic Electrodes". In: *The Journal of Physical Chemistry Letters* 7.24 (2016), pp. 5168–5175. DOI: 10.1021/acs.jpclett.6b02375.

[114] J. M. Azpiroz, E. Mosconi, J. Bisquert, and F. D. Angelis. "Defect migration in methylammonium lead iodide and its role in perovskite solar cell operation". In: *Energy & Environmental Science* 8.7 (2015), pp. 2118–2127. DOI: 10.1039/C5EE01265A.

[115] J.-H. Yang, W.-J. Yin, J.-S. Park, and S.-H. Wei. "Fast self-diffusion of ions in $CH_3NH_3PbI_3$: the interstiticaly mechanism versus vacancy-assisted mechanism". In: *Journal of Materials Chemistry A* 4.34 (2016), pp. 13105–13112. DOI: 10.1039/C6TA03599J.

[116] P. Delugas, C. Caddeo, A. Filippetti, and A. Mattoni. "Thermally Activated Point Defect Diffusion in Methylammonium Lead Trihalide: Anisotropic and Ultrahigh Mobility of Iodine". In: *The Journal of Physical Chemistry Letters* 7.13 (2016), pp. 2356–2361. DOI: 10.1021/acs.jpclett.6b00963.

[117] E. J. Maginn, R. A. Messerly, D. J. Carlson, D. R. Roe, and J. R. Elliott. "Best Practices for Computing Transport Properties 1. Self-Diffusivity and Viscosity from Equilibrium Molecular Dynamics [Article v1.0]". In: *Living Journal of Computational Molecular Science* 1.1 (2018), p. 6324. DOI: 10.33011/livecoms.1.1.6324.

[118] J. E. Basconi and M. R. Shirts. "Effects of Temperature Control Algorithms on Transport Properties and Kinetics in Molecular Dynamics Simulations". In: *Journal of Chemical Theory and Computation* 9.7 (2013), pp. 2887–2899. DOI: 10.1021/ct400109a.

[119] A. Einstein. "Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". In: *Annalen der Physik* 322.8 (1905), pp. 549–560. DOI: 10.1002/andp.19053220806.

[120] S. Arrhenius. "Über die Dissociationswärme und den Einfluss der Temperatur auf den Dissociationsgrad der Elektrolyte". In: *Zeitschrift für Physikalische Chemie* 4U.1 (1889), pp. 96–116. DOI: 10.1515/zpch-1889-0408.

[121] Y.-H. Kye, C.-J. Yu, U.-G. Jong, K.-C. Ri, J.-S. Kim, S.-H. Choe, S.-N. Hong, S. Li, J. N. Wilson, and A. Walsh. "Vacancy-Driven Stabilization of the Cubic Perovskite Polymorph of $CsPbI_3$". In: *The Journal of Physical Chemistry C* 123.15 (2019), pp. 9735–9744. DOI: 10.1021/acs.jpcc.9b01552.

[122] R. A. Evarestov, A. Senocrate, E. A. Kotomin, and J. Maier. "First-principles calculations of iodine-related point defects in $CsPbI_3$". In: *Physical Chemistry Chemical Physics* 21.15 (2019), pp. 7841–7846. DOI: 10.1039/C9CP00414A.

[123] R. A. Evarestov, E. A. Kotomin, A. Senocrate, R. K. Kremer, and J. Maier. "First-principles comparative study of perfect and defective $CsPbX_3$ (X = Br, I) crystals". In: *Physical Chemistry Chemical Physics* 22.7 (2020), pp. 3914–3920. DOI: 10.1039/C9CP06322F.

[124] L. A. Muscarella, E. M. Hutter, F. Wittmann, Y. W. Woo, Y.-K. Jung, L. McGovern, J. Versluis, A. Walsh, H. J. Bakker, and B. Ehrler. "Lattice Compression Increases the Activation Barrier for Phase Segregation in Mixed-Halide Perovskites". In: *ACS Energy Letters* 5.10 (2020), pp. 3152–3158. DOI: 10.1021/acsenergylett.0c01474.

[125] J. Mizusaki, K. Arai, and K. Fueki. "Ionic conduction of the perovskite-type halides". In: *Solid State Ionics* 11.3 (1983), pp. 203–211. DOI: 10.1016/0167-2738(83)90025-5.

[126] K. J. Laidler. "The development of the Arrhenius equation". In: *Journal of Chemical Education* 61.6 (1984), p. 494. DOI: 10.1021/ed061p494.

[127] A. Manekkathodi, A. Marzouk, J. Ponraj, A. Belaidi, and S. Ashhab. "Observation of Structural Phase Transitions and PbI_2 Formation During the Degradation of Triple-Cation Double-Halide Perovskites". In: *ACS Applied Energy Materials* 3.7 (2020), pp. 6302–6309. DOI: 10.1021/acsaem.0c00515.

[128] A.-F. Castro-Méndez, J. Hidalgo, and J.-P. Correa-Baena. "The Role of Grain Boundaries in Perovskite Solar Cells". In: *Advanced Energy Materials* 9.38 (2019), p. 1901489. DOI: 10.1002/aenm.201901489.

[129] B. Conings et al. "Intrinsic Thermal Instability of Methylammonium Lead Trihalide Perovskite". In: *Advanced Energy Materials* 5.15 (2015), p. 1500477. DOI: 10.1002/aenm.201500477.

[130] Q. Wang, B. Chen, Y. Liu, Y. Deng, Y. Bai, Q. Dong, and J. Huang. "Scaling behavior of moisture-induced grain degradation in polycrystalline hybrid perovskite thin films". In: *Energy & Environmental Science* 10.2 (2017), pp. 516–522. DOI: 10.1039/C6EE02941H.

[131] F. Ji, S. Pang, L. Zhang, Y. Zong, G. Cui, N. P. Padture, and Y. Zhou. "Simultaneous Evolution of Uniaxially Oriented Grains and Ultralow-Density Grain-Boundary Network in $CH_3NH_3PbI_3$ Perovskite Thin Films Mediated by Precursor Phase Metastability". In: *ACS Energy Letters* 2.12 (2017), pp. 2727–2733. DOI: 10.1021/acsenergylett.7b00980.

[132] W.-J. Yin, H. Chen, T. Shi, S.-H. Wei, and Y. Yan. "Origin of High Electronic Quality in Structurally Disordered $CH_3NH_3PbI_3$ and the Passivation Effect of Cl and O at Grain Boundaries". In: *Advanced Electronic Materials* 1.6 (2015), p. 1500044. DOI: 10.1002/aelm.201500044.

[133] R. Long, J. Liu, and O. V. Prezhdo. "Unravelling the Effects of Grain Boundary and Chemical Doping on Electron–Hole Recombination in $CH_3NH_3PbI_3$ Perovskite by Time-Domain Atomistic Simulation". In: *Journal of the American Chemical Society* 138.11 (2016), pp. 3884–3890. DOI: 10.1021/jacs.6b00645.

[134] Y. Guo, Q. Wang, and W. A. Saidi. "Structural Stabilities and Electronic Properties of High-Angle Grain Boundaries in Perovskite Cesium Lead Halides". In: *The Journal of Physical Chemistry C* 121.3 (2017), pp. 1715–1722. DOI: 10.1021/acs.jpcc.6b11434.

[135] J.-S. Park, J. Calbo, Y.-K. Jung, L. D. Whalley, and A. Walsh. "Accumulation of Deep Traps at Grain Boundaries in Halide Perovskites". In: *ACS Energy Letters* 4.6 (2019), pp. 1321–1327. DOI: 10.1021/acsenergylett.9b00840.

[136] N. Li et al. "Cation and anion immobilization through chemical bonding enhancement with fluorides for stable halide perovskite solar cells". In: *Nature Energy* 4.5 (2019), pp. 408–415. DOI: 10.1038/s41560-019-0382-6.

[137] E. H. Jung, N. J. Jeon, E. Y. Park, C. S. Moon, T. J. Shin, T.-Y. Yang, J. H. Noh, and J. Seo. "Efficient, stable and scalable perovskite solar cells using poly(3-hexylthiophene)". In: *Nature* 567.7749 (2019), pp. 511–515. DOI: 10.1038/s41586-019-1036-3.

[138] Q. Jiang, Y. Zhao, X. Zhang, X. Yang, Y. Chen, Z. Chu, Q. Ye, X. Li, Z. Yin, and J. You. "Surface passivation of perovskite film for efficient solar cells". In: *Nature Photonics* 13.7 (2019), pp. 460–466. DOI: 10.1038/s41566-019-0398-2.

[139] W. Rehman, D. P. McMeekin, J. B. Patel, R. L. Milot, M. B. Johnston, H. J. Snaith, and L. M. Herz. "Photovoltaic mixed-cation lead mixed-halide perovskites: links between crystallinity, photo-stability and electronic properties". In: *Energy & Environmental Science* 10.1 (2017), pp. 361–369. DOI: 10.1039/C6EE03014A.

[140] S. Shao and M. A. Loi. "The Role of the Interfaces in Perovskite Solar Cells". In: *Advanced Materials Interfaces* 7.1 (2020), p. 1901469. DOI: 10.1002/admi.201901469.

[141] S.-Y. Kim, N. Kumar, P. Persson, J. Sofo, A. C. T. van Duin, and J. D. Kubicki. "Development of a ReaxFF Reactive Force Field for Titanium Dioxide/Water Systems". In: *Langmuir* 29.25 (2013), pp. 7838–7846. DOI: 10.1021/la4006983.

[142] T. Leijtens, G. E. Eperon, S. Pathak, A. Abate, M. M. Lee, and H. J. Snaith. "Overcoming ultraviolet light instability of sensitized TiO_2 with meso-superstructured organometal tri-halide perovskite solar cells". In: *Nature Communications* 4.1 (2013), p. 2885. DOI: 10.1038/ncomms3885.

[143] R. Liu, L. Wang, Y. Fan, Z. Li, and S. Pang. "UV degradation of the interface between perovskites and the electron transport layer". In: *RSC Advances* 10.20 (2020), pp. 11551–11556. DOI: 10.1039/C9RA10960A.

[144] D. M. Ceperley and B. J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Physical Review Letters* 45.7 (1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566.

[145] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke. "Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces". In: *Physical Review Letters* 100.13 (2008), p. 136406. DOI: 10.1103/PhysRevLett.100.136406.

[146] J. P. Perdew, A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke. "Erratum: Restoring the Density-Gradient Expansion for Exchange in Solids and Surfaces [Phys. Rev. Lett. 100, 136406 (2008)]". In: *Physical Review Letters* 102.3 (2009), p. 039902. DOI: 10.1103/PhysRevLett.102.039902.

[147] J. Sun, A. Ruzsinszky, and J. P. Perdew. "Strongly Constrained and Appropriately Normed Semilocal Density Functional". In: *Physical Review Letters* 115.3 (2015), p. 036402. DOI: 10.1103/PhysRevLett.115.036402.

[148] N. Marom, A. Tkatchenko, M. Rossi, V. V. Gobre, O. Hod, M. Scheffler, and L. Kronik. "Dispersion Interactions with Density-Functional Theory: Benchmarking Semiempirical and Interatomic Pairwise Corrected Density Functionals". In: *Journal of Chemical Theory and Computation* 7.12 (2011), pp. 3944–3951. DOI: 10.1021/ct2005616.

[149] H. Peng, Z.-H. Yang, J. P. Perdew, and J. Sun. "Versatile van der Waals Density Functional Based on a Meta-Generalized Gradient Approximation". In: *Physical Review X* 6.4 (2016), p. 041005. DOI: 10.1103/PhysRevX.6.041005.

[150] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". In: *The Journal of Chemical Physics* 132.15 (2010), p. 154104. DOI: 10.1063/1.3382344.

[151] 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". In: *Physical Review B* 94.11 (2016), p. 115144. DOI: 10.1103/PhysRevB.94.115144.

[152] E. Flahaut, J. Sloan, S. Friedrichs, A. I. Kirkland, K. S. Coleman, V. C. Williams, N. Hanson, J. L. Hutchison, and M. L. H. Green. "Crystallization of 2H and 4H PbI_2 in Carbon Nanotubes of Varying Diameters and Morphologies". In: *Chemistry of Materials* 18.8 (2006), pp. 2059–2069. DOI: 10.1021/cm0526056.

[153] T. B. Rymer and P. G. Hambling. "The lattice constant of caesium iodide". In: *Acta Crystallographica* 4.6 (1951), pp. 565–565. DOI: 10.1107/S0365110X51001847.

[154] G. I. Csonka, J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán. "Assessing the performance of recent density functionals for bulk solids". In: *Physical Review B* 79.15 (2009), p. 155107. DOI: 10.1103/PhysRevB.79.155107.

[155] G.-X. Zhang, A. M. Reilly, A. Tkatchenko, and M. Scheffler. "Performance of various density-functional approximations for cohesive properties of 64 bulk solids". In: *New Journal of Physics* 20.6 (2018), p. 063020. DOI: 10.1088/1367-2630/aac7f0.

[156] R. J. Sutton, M. R. Filip, A. A. Haghighirad, N. Sakai, B. Wenger, F. Giustino, and H. J. Snaith. "Cubic or Orthorhombic? Revealing the Crystal Structure of Metastable Black-Phase $CsPbI_3$ by Theory and Experiment". In: *ACS Energy Letters* 3.8 (2018), pp. 1787–1794. DOI: 10.1021/acsenergylett.8b00672.

[157] R. Jinnouchi, J. Lahnsteiner, F. Karsai, G. Kresse, and M. Bokdam. "Phase Transitions of Hybrid Perovskites Simulated by Machine-Learning Force Fields Trained on the Fly with Bayesian Inference". In: *Physical Review Letters* 122.22 (2019), p. 225701. DOI: 10.1103/PhysRevLett.122.225701.

[158] N. Michaud-Agrawal, E. J. Denning, T. B. Woolf, and O. Beckstein. "MDAnalysis: A toolkit for the analysis of molecular dynamics simulations". In: *Journal of Computational Chemistry* 32.10 (2011), pp. 2319–2327. DOI: 10.1002/jcc.21787.

[159] R. J. Gowers et al. "MDAnalysis: A Python Package for the Rapid Analysis of Molecular Dynamics Simulations". In: *Proceedings of the 15th Python in Science Conference* (2016), pp. 98–105. DOI: 10.25080/Majora-629e541a-00e.

[160] V. Calandrini, E. Pellegrini, P. Calligari, K. Hinsen, and G. R. Kneller. "nMoldyn - Interfacing spectroscopic experiments, molecular dynamics simulations and models for time correlation functions". In: *École thématique de la Société Française de la Neutronique* 12 (2011), pp. 201–232. DOI: 10.1051/sfn/201112010.

[161] P. d. Buyl. "tidynamics: A tiny package to compute the dynamics of stochastic and molecular simulations". In: *Journal of Open Source Software* 3.28 (2018), p. 877. DOI: 10.21105/joss.00877.

Appendices

A. The phases of $CsPbI_3$

In this appendix, we will explain the rich phase behaviour of metal halide perovskites. In particular, we zoom into $CsPbI_3$, the metal halide perovskite that is central to the work in this thesis. An overview of the phase evolution observed in experiments [99, 100] is shown in Figure A.1, what follows is a detailed explanation of the different polymorphs that $CsPbI_3$ can adopt.

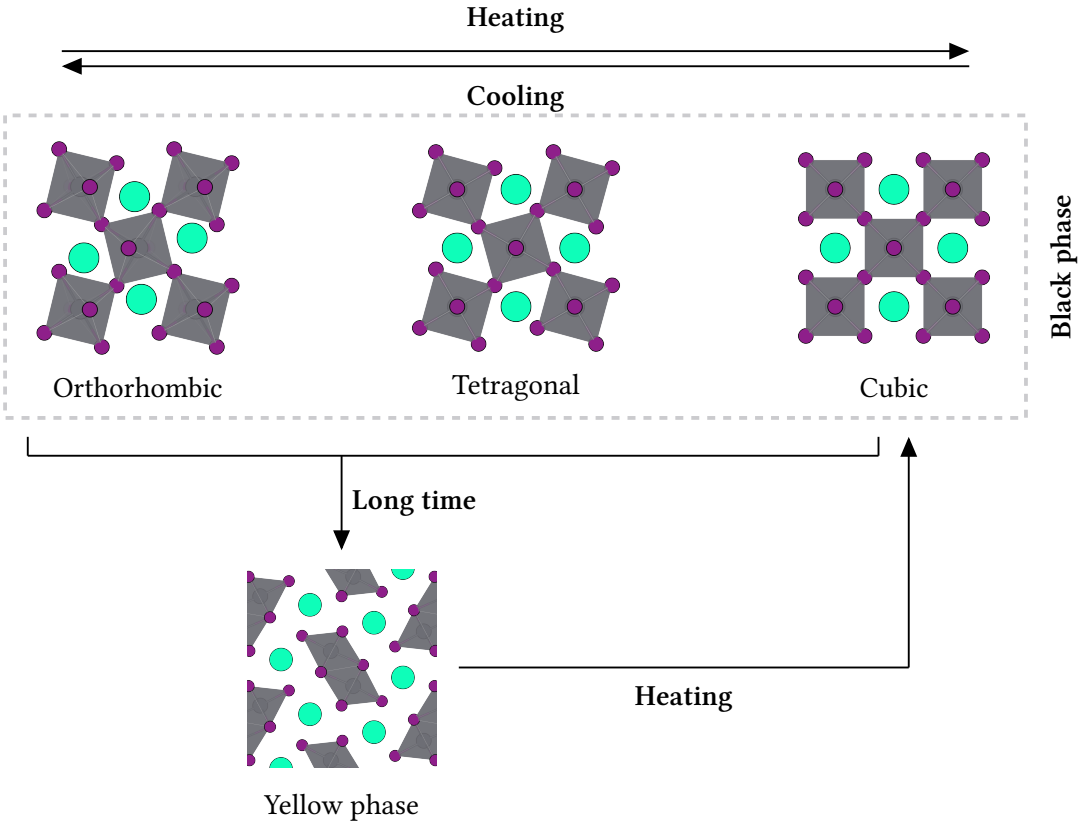

Figure A.1. The phase evolution of $CsPbI_3$ as observed from experiments [99, 100]. Two types of structures are shown: the metastable black phase perovskite structures (cubic, tetragonal and orthorhombic) and the yellow phase non-perovskite structure. The c-axis is taken as the viewing direction of the structures.

$CsPbI_3$ is known to adopt two different types of structures: non-perovskite and perovskite. The non-perovskite phase of $CsPbI_3$ is the thermodynamically stable phase at room temperature and is often named the yellow phase as a result of its color [101, 109]. This phase consists of one-dimensional chains of face-sharing metal halide octahedra ($Pnma$) [101, 102]. Because of its

poor optoelectronic properties, which include a large indirect bandgap, this polymorph of $CsPbI_3$ has no practical use in solar cells [109].

Studies have shown that this non-perovskite phase can readily be transformed into a metastable perovskite phase at elevated temperatures [99–102]. The perovskite phases have the right optoelectronic properties to be used in solar cells [109], and are typically named black phase $CsPbI_3$ as a result of their colour. The phase consists of a network of interconnecting metal halide octahedra (PbI_6 octahedra) of which the cuboids formed by these octahedra are occupied by a monovalent cation (Cs cation). Since the cation of $CsPbI_3$ is undersized in comparison to the space it occupies in the lattice, the perovskites are predicted to show a tilting of the metal halide octahedra [106].

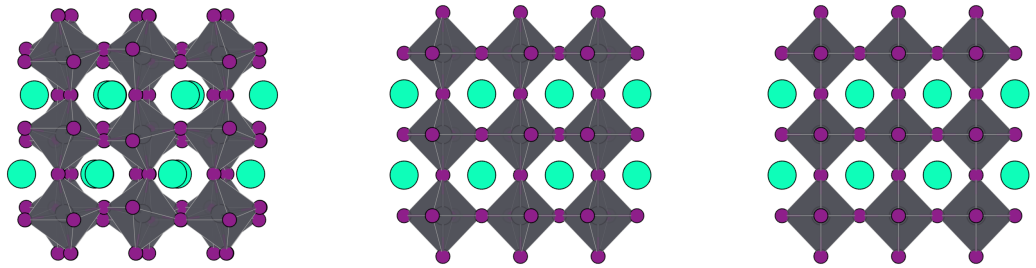

Figure A.2. A side view of the different perovskite phases of $CsPbI_3$. The a-axis is taken as the viewing direction of the structures.

At high temperatures, the perovskite has an ideal cubic perovskite ($Pm\bar{3}m$) phase, where the metal halide octahedra do not show any tilting [99, 100]. Upon cooling this high-temperature perovskite phase results in two subsequent octahedral tilting transitions. First, the octahedra obtain an in-phase tilting around one common axis of the unit cell, the c-axis, resulting in the tetragonal phase ($P4/mbm$) of $CsPbI_3$ [99, 100]. Further cooling of the tetragonal phase causes the octahedra to show an additional out-of-phase tilting around the a- and b-axis of the unit cell, resulting in the orthorhombic phase ($Pnma$) of $CsPbI_3$ [99, 100]. An overview of the tilting patterns can be obtained from the top view in Figure A.1 and the side view in Figure A.2.

Table A.1. Summary of the different perovskite phases of $CsPbI_3$. For each of the phases we note the space group, lattice vectors and pseudo-cubic lattice vectors of the unit cells.

Phase	Space group	Lattice vectors	Pseudo-cubic lattice vectors
Cubic	$Pm\bar{3}m$	(a, a, a)	(a, a, a)
Tetragonal	$P4/mbm$	(a, a, c)	$\left(\frac{a}{\sqrt{2}}, \frac{a}{\sqrt{2}}, c\right)$
Orthorhombic	$Pnma$	(a, b, c)	$\left(\frac{a}{\sqrt{2}}, \frac{b}{\sqrt{2}}, \frac{c}{2}\right)$

As a result of these octahedral tilting patterns, the lattice vector lengths evolve as a function of the temperature. In the cubic perovskite phase, the lattice vectors all have the same length

($a = b = c$). The tilting of the octahedra causes the lattice vector around which the octahedra tilt to be of a different length than the others ($a = b \neq c$ and c). Finally, with the additional octahedral tilting in the orthorhombic phase, all lattice vectors have a different length ($a \neq b \neq c$). These lattice vectors can be converted to pseudo-cubic lattice vectors to allow for a comparison between the different phases. An overview of the space group, lattice vectors and pseudo-cubic lattice vectors of the unit cells is given in Table A.1.

B. Settings in the training set

As ReaxFF force fields are typically trained against density functional theory (DFT) data, the quality of the DFT data in the training set is critical for the accuracy of the obtained force field. To assess the quality of different levels of theory within the framework of DFT, we tested a variety of methods as implemented in the Vienna Ab-Initio Simulation Package (VASP) [72–75] in this appendix. Noting that we aim to obtain a ReaxFF force field that can accurately describe the phase and defect behaviour of the $CsPbI_3$ metal halide perovskite, we focused in particular on the accuracy of the computational methods for a variety of experimental phases of $CsPbI_3$ and its precursors.

Since training sets for force fields require an extensive amount of data, we limited our tests to the computationally efficient local and semi-local functionals that correspond to the bottom three rungs of the Jacob's ladder of DFT: local density approximation (LDA), generalized gradient approximation (GGA) and meta-GGA [42, 43]. Each successive rung incorporates additional exact constraints in the exchange-correlation (XC) interaction of the electrons, which results in a more accurate approximation at the expense of some computational cost. The XC-functionals that we tested include LDA (LDA) [144], PBE (GGA) [76, 77], PBEsol (GGA) [145, 146] and SCAN (meta-GGA) [147]. Although the tested functionals perform well for a large number of compounds, it is nevertheless found that they perform sub-optimally in systems where long-range dispersive interactions play an important role [148, 149]. Therefore, to potentially overcome this deficiency, our tests also covered the explicit inclusion of these dispersive interactions through the DFT-D3 [150], DFT-D3(BJ) [78] and rVV10 [149] methods[19].

To benchmark the accuracy of the different density functional approximations, we looked at the material geometries and relative stability of the different perovskite phases. Note that all structures were fully relaxed according to the calculation settings[20] that we describe for the equations of state in our training set in subsection 3.3.1. In the geometry comparison we referenced a variety of material phases, including PbI_2 [152], CsI [153], cubic, tetragonal and orthorhombic $CsPbI_3$ (black phase) [99] and yellow phase $CsPbI_3$ [102], to their experimentally reported geometries. Here we used the relative change in unit cell volume with respect to the experimental unit cell to measure the accuracy of the calculations. An overview of this comparison is shown in figure

[19] Since VASP by default does not contain the correct dispersive parameters to properly account for dispersive interactions with the SCAN functional, we used parameters from literature for the DFT-D3/DFT-D3(BJ) [151] and rVV10 [149] dispersive interactions.

[20] The k-space grid of the yellow phase of $CsPbI_3$ was not mentioned in this section, we employed a $13 \times 6 \times 4$ grid for the yellow phase of $CsPbI_3$ which resulted in a convergence of the energy to within 1 meV/atom.

Figure B.1. The phase stability was determined from the formation energies of the perovskites. We define the formation energy per formal unit of $CsPbI_3$ as

$$E_f = E_{CsPbI_3} - E_{CsI} - E_{PbI_2}, \tag{B.1}$$

where E_{CsPbI_3}, E_{CsI} and E_{PbI_2} are, respectively, the energies per formal unit of $CsPbI_3$, CsI and PbI_2. Not only did this formation energy give us a measure of the stability of the different phases of $CsPbI_3$ with respect to its precursors, it also allowed for a comparison of the relative energies between the different $CsPbI_3$ material phases. Without any experimental data on the formation energies of the compounds, the different approximations were compared amongst themselves. The resulting formation energies are shown in figure Figure B.2.

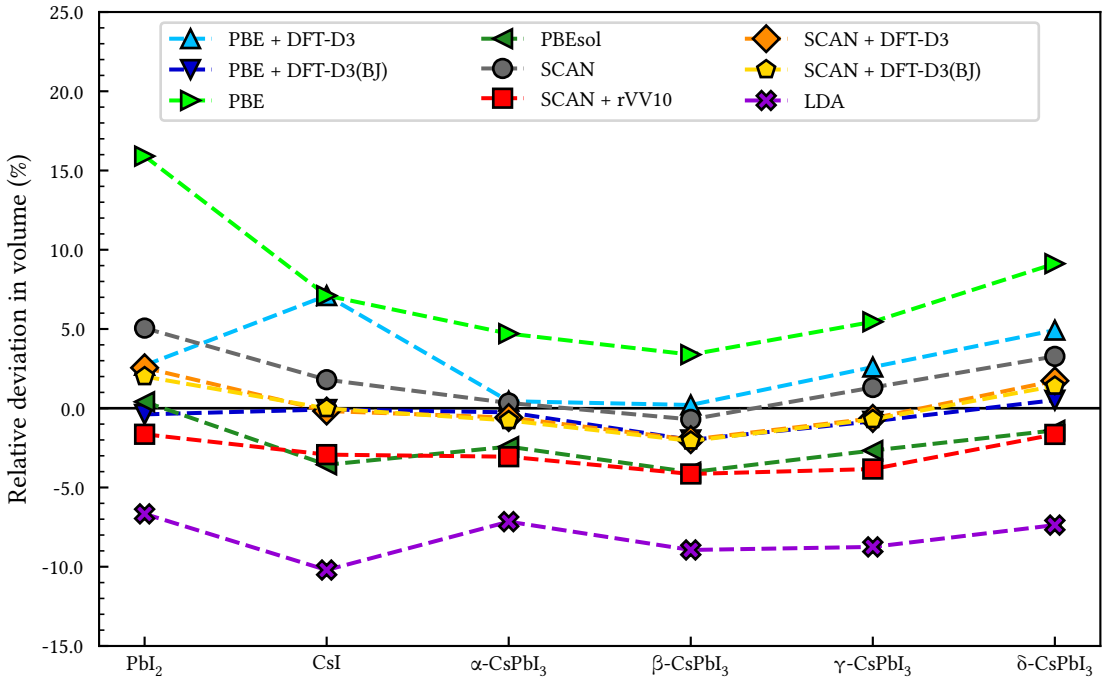

Figure B.1. Unit cell volumes for a variety of density functional approximations compared to experimentally determined unit cell volumes. With α-, β-, γ-, δ-$CsPbI_3$ denoting the cubic, tetragonal, orthorhombic and yellow phase of $CsPbI_3$, respectively.

From the comparison of the geometries, we learn that the majority of density functional approximations accurately predicts the experimental unit cell geometries (to within 5% volume), with PBE + DFT-D3(BJ) performing the best (to within 2% volume). Two notable outliers in this comparison are LDA and PBE. However, the mismatch of LDA and PBE with the experimental geometries is of no surprise, since they are known for under- and overpredicting geometries due to a respective over- and underbinding [154, 155]. Our geometry comparison demonstrates that the underbinding of PBE can be overcome with some refinements to the density functional approximation. The reparametrization of PBE for solids (PBEsol) and the explicit inclusion of dispersion interactions in a post-correction (DFT-D), result in geometries that match well with experiments.

The improvement is most apparent for the layered PbI_2 compound. Dispersive interactions are key to the material geometry of PbI_2. The refinements either implicitly or explicitly account for these long-range interactions, resulting in more realistic interlayer distances and thus unit cell geometries. Nevertheless, the refinement of PBE with DFT-D3 does not improve the result for all material geometries: the underbinding of PBE seems to persist for a calculation of CsI with PBE + DFT-D3[21]. In contrast to LDA and PBE, the SCAN meta-GGA functional predicts geometries that are rather close to experiments, which is supported by the fact that approximations higher up the Jacob's ladder of DFT are generally deemed more accurate, due to the additional exact constraints [42, 43]. Although SCAN itself already accurately predicts geometries, the explicit inclusion of dispersive interactions, either through DFT-D methods or rVV10, results in predictions that agree slightly better with experiments, but the improvement is not as distinct as for PBE.

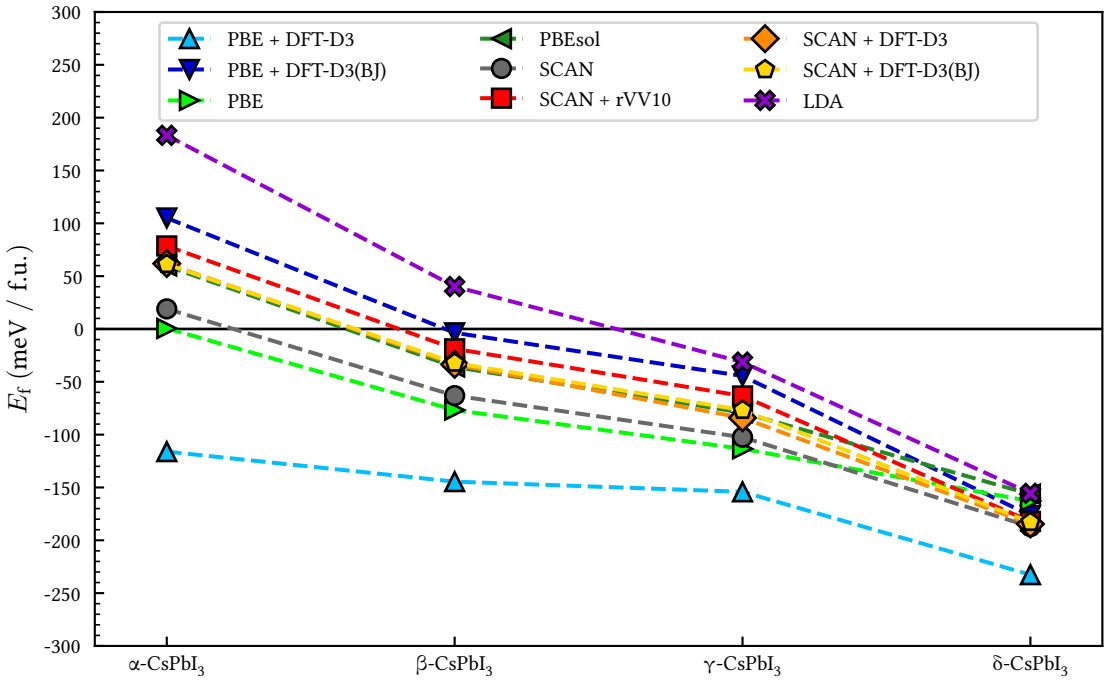

Figure B.2. The formation energies of the different material phases of $CsPbI_3$ per formal unit as calculated with different density functional approximations. With α-, β-, γ-, δ-$CsPbI_3$ denoting the cubic, tetragonal, orthorhombic and yellow phase of $CsPbI_3$, respectively.

A first glance at the comparison of the formation energies shows that they are sensitive to XC-functional choice. A similar sensitivity was reported on for the $CsPbI_3$ metal halide perovskite in literature [156]. Despite this sensitivity to the choice of XC-functional, all density functional approximations predict a similar stability trend for the different $CsPbI_3$ phases. When ordered from least to most stable, the phases are cubic, tetragonal, orthorhombic and yellow phase $CsPbI_3$,

[21] A detail that is likely caused by a bug in VASP when using PBE + DFT-D3 together with CsI, where the dispersive interactions do not seem to affect the forces and stress tensor during the geometry optimization.

a trend that matches well with experiments [99, 100]. Nonetheless, when the stability of $CsPbI_3$ with respect to its precursors is concerned, PBE + DFT-D3 appears to be an outlier. It predicts all black phases of $CsPbI_3$ to be stable with respect to its precursors, whereas most approximations predict at least a part to be unstable to its precursors. This apparent overstabilization is of geometric origin. In the benchmark of the geometries, we found that PBE + DFT-D3 underbinds CsI. Naturally, this results in a destabilization (higher energy) of this precursor. Thus, all the formation energies for this density functional approximation are overstabilized due to this erroneous reference structure.

When we consider the results of both benchmarks, we judge the PBE + DFT-D3(BJ) approximation to be the most suitable for the generation of the reference data in the training set. Not only does it result in geometries that best resemble experimental geometries, but it also predicts a phase stability for the phases of $CsPbI_3$ that is in line with experimental observations with a reasonable computational cost. Therefore, we generate all reference data in our training set with the PBE + DFT-D3(BJ) density functional approximation.

C. I/Pb/Cs ReaxFF parameters

```
Reactive MD-force field: I/Pb/Cs, M. Pols et al. - MSc Thesis
 39        ! Number of general parameters
   50.0000 !p_boc1 Eq(4c): Overcoordination parameter
    9.5469 !p_boc2 Eq(4d): Overcoordination parameter
   26.5405 !p_coa2 Eq(15): Valency angle conjugation
    1.7224 !p_trip4 Eq(20): Triple bond stabilisation
    6.8702 !p_trip3 Eq(20): Triple bond stabilisation
   60.4850 !k_c2 Eq(19): C2-correction
    1.0588 !p_ovun6 Eq(12): Undercoordination
    4.6000 !p_trip2 Eq(20): Triple bond stabilisation
   12.1176 !p_ovun7 Eq(12): Undercoordination
   13.3056 !p_ovun8 Eq(12): Undercoordination
  -70.5044 !p_trip1 Eq(20): Triple bond stabilization
    0.0000 !Lower Taper-radius (must be 0)
   10.0000 !R_cut Eq(21): Upper Taper-radius
    2.8793 !p_fe1 Eq(6a): Fe dimer correction
   33.8667 !p_val6 Eq(13c): Valency undercoordination
    6.0891 !p_lp1 Eq(8): Lone pair param
    1.0563 !p_val9 Eq(13f): Valency angle exponent
    2.0384 !p_val10 Eq(13g): Valency angle parameter
    6.1431 !p_fe2 Eq(6a): Fe dimer correction
    6.9290 !p_pen2 Eq(14a): Double bond/angle param
    0.3989 !p_pen3 Eq(14a): Double bond/angle param
    3.9954 !p_pen4 Eq(14a): Double bond/angle param
   -2.4837 !p_fe3 Eq(6a): Fe dimer correction
    5.7796 !p_tor2 Eq(16b): Torsion/BO parameter
   10.0000 !p_tor3 Eq(16c): Torsion overcoordination
    1.9487 !p_tor4 Eq(16c): Torsion overcoordination
   -1.2327 !p_elho Eq(26a): electron-hole interaction
    2.1645 !p_cot2 Eq(17b): Conjugation if tors13=0
    1.5591 !p_vdW1 Eq(23b): vdWaals shielding
    0.1000 !Cutoff for bond order (*100)
    2.1365 !p_coa4 Eq(15): Valency angle conjugation
    0.6991 !p_ovun4 Eq(11b): Over/Undercoordination
   50.0000 !p_ovun3 Eq(11b): Over/Undercoordination
    1.8512 !p_val8 Eq(13d): Valency/lone pair param
    0.5000 !X_soft Eq(25): ACKS2 softness for X_ij
   20.0000 !d Eq(23d): Scale factor in lg-dispersion
    5.0000 !p_val Eq(27): Gauss exponent for electrons
    0.0000 !1 Eq(13e): disable undecoord in val angle
    2.6962 !p_coa3 Eq(15): Valency angle conjugation
```

```
 4    ! Nr of atoms; cov.r; valency;a.m;Rvdw;Evdw;gammaEEM;cov.r2;#
           alfa;gammavdW;valency;Eunder;Eover;chiEEM;etaEEM;n.u.
           cov r3;Elp;Heat inc.;bo131;bo132;bo133;softcut;n.u.
           ov/un;val1;n.u.;val3,vval4
I     1.9000    1.0000 126.9000    3.1000    0.0861    0.8009   -1.0000    7.0000
     10.3555    5.4304    1.0000    0.0000    0.0000    9.8607    6.2778    2.0000
     -1.0000    2.3407   35.1770    6.2293    5.2294    0.1542    0.8563    0.0000
    -10.2080    2.9867    1.0338    6.2998    2.5791    0.0000    0.0000    0.0000
Pb    1.9590    2.0000 207.2000    2.1357    0.3301    0.3236   -1.0000    2.0000
     12.1777    4.2077    2.0000    0.0000    0.0000    4.7057    5.0078    0.0000
     -1.0000    0.0000 133.1770   27.2704    1.8727    0.1586    0.8563    0.0000
    -13.3352    2.3781    1.0338    5.0000    2.5791    0.0000    0.0000    0.0000
Cs    0.0100    1.0000 132.9054    2.6059    0.4360    0.6004   -1.0000    1.0000
      8.9210    3.8710    1.0000    0.0000    0.0000   -2.7745    8.8075    0.0000
     -1.0000    0.0000   23.0445  100.0000    1.0000    0.0000    0.8563    0.0000
     -2.5000    3.9900    1.0338    8.0000    2.5791    1.0000    0.0100   13.0000
X    -0.1000    2.0000    1.0080    2.0000    0.0000    0.0100   -0.1000    6.0000
     10.0000    2.5000    4.0000    0.0000    0.0000    8.5000  999.0000    0.0000
     -0.1000    0.0000   -2.3700    8.7410   13.3640    0.6690    0.9745    0.0000
    -11.0000    2.7466    1.0338    4.0000    2.8793    0.0000    0.0000    0.0000
 6    ! Nr of bonds; Edis1;LPpen;n.u.;pbe1;pbo5;13corr;pbo6
                        pbe2;pbo3;pbo4;n.u.;pbo1;pbo2;ovcorr
 1  1 116.1043    0.0000    0.0000    0.9312   -0.3500    0.0000   25.0000    0.9492
       -0.1922   -0.2500   15.0000    1.0000   -0.1772    6.8516    0.0000    0.0000
 1  2  96.2233    0.0000    0.0000   -0.2358   -0.2000    0.0000   16.0000    0.2032
        3.0008   -0.2000   15.0000    1.0000   -0.0358    4.8860    0.0000    0.0000
 1  3  26.8805    0.0000    0.0000    0.4653   -0.2000    0.0000   16.0000    0.2920
        0.1064   -0.2000   15.0000    1.0000   -0.0488    4.5473    0.0000    0.0000
 2  2  65.8420    0.0000    0.0000   -0.3550   -0.2000    0.0000   16.0000    0.2382
        1.7003   -0.2000   15.0000    1.0000   -0.0640    5.7778    0.0000    0.0000
 2  3   0.0000    0.0000    0.0000    0.5000   -0.3000    1.0000   16.0000    0.5000
        0.5000   -0.2500   15.0000    1.0000   -0.1000    9.0000    0.0000    0.0000
 3  3  17.6609    0.0000    0.0000    0.5700    0.3000    0.0000   25.0000    0.6157
        0.4946   -0.4000   12.0000    1.0000   -0.0557    4.5153    0.0000    0.0000
 3    ! Nr of off-diagonal terms; Ediss;Ro;gamma;rsigma;rpi;rpi2
 1  2   0.1958    2.2289   13.6252    2.1484   -1.0000   -1.0000
 1  3   0.4324    2.3251   12.1698    1.9784   -1.0000   -1.0000
 2  3   0.2337    2.4896   12.1378    0.0100   -1.0000   -1.0000
 7    ! Nr of angles;at1;at2;at3;Thetao,o;ka;kb;pv1;pv2
 1  1  2  97.2120    2.5724    1.6934    0.0000    0.6628    0.0000    1.5679
 1  2  2  99.3893    3.8960    3.0085    0.0000    1.3692    0.0000    1.0780
 1  2  1 113.4265   15.3364    0.8860    0.0000    0.0513    0.0000    1.2651
 2  1  2 102.7582    9.1798    1.8220    0.0000    0.6699    0.0000    1.9751
 1  3  1  70.0899    0.6959    1.0539    0.0000    0.5191    0.0000    1.6486
 3  1  3  97.5407    1.8600    3.3689    0.0000    0.5922    0.0000    1.0959
 3  1  2  81.6045    6.7382    1.7103    0.0000    0.6253    0.0000    1.3576
 0    ! Nr of torsions;at1;at2;at3;at4;;V1;V2;V3;V2(BO);vconj;n.u;n
 0    ! Nr of hydrogen bonds;at1;at2;at3;Rhb;Dehb;vhb1
```

D. Lattice vector analysis

In this appendix, we cover the method with which the lattice vectors of $CsPbI_3$ were determined during the *discrete*-approach to the phase diagram. The analysis method closely follows the method outlined by Jinnouchi *et al.* [157].

During the constant temperature simulations of the perovskite, the lattice vectors fluctuated considerably[22], making it difficult to determine the lattice vectors at these temperatures uniquely. Therefore, we binned the time evolution of the pseudo-cubic lattice vectors. The binned lattice vector distributions were consequently fit with Gaussian distributions to determine the lattice vectors. The Gaussian distributions ρ have the following functional form

$$\rho(x) = A \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right], \tag{D.1}$$

with x the binned latticed vector distribution, A a scaling factor for the distribution, μ the equilibrium lattice vector and σ the standard deviation of the lattice vector.

First, all pseudo-cubic lattice vectors were binned individually, where each binned distributions was fit with their own Gaussian distribution[23]. If the difference between the equilibrium lattice vectors was larger than $\frac{1}{3}\sqrt{\sigma_1^2+\sigma_2^2+\sigma_3^2}$, unit cell was judged orthorhombic and consequently a, b and c were all determined to be different. If not, the lattice vectors with a smaller difference were grouped, and two Gaussian distributions were fit. In case the difference between the average values of these two distributions was larger than $\frac{1}{2}\sqrt{\sigma_1^2+\sigma_2^2}$, the unit cell was judged tetragonal, resulting in a value for a and c. Else, the unit cell is judged as cubic. All distributions were grouped, and a single Gaussian function was fit, resulting in the same lattice vector values for a, b and c, which we call a.

[22] Fluctuations of the lattice vector are possible due to the use of an *NpT*-ensemble in which the system volume is allowed to change.

[23] Each of the distributions has its own average value μ_i and standard distribution σ_i with $i = 1, 2$ or 3 indexing the different distributions.

E. Sensitivity of the phase diagrams

In subsection 4.1.1 we investigate the phase diagrams that we obtained from ReaxFF simulations with our force field for $CsPbI_3$. This appendix elaborates on some details that are not covered in the main text. First we dive into the sensitivity of the phase diagram as determined with the *temperature ramp*-approach to a range of molecular dynamics parameters. Then, we investigate the non-reversibility of some of the cooling runs with the *temperature ramp*-approach. Note that all of the continuous heating and cooling simulations covered in this appendix were done with a $6 \times 6 \times 6$-supercell of orthorhombic $CsPbI_3$.

Molecular dynamics parameters

In this section, we examine the influence of a variety of molecular dynamics parameters on the continuous phase diagram of $CsPbI_3$.

First of all, one might argue that since we smoothed the data of the continuous phase diagrams by taking a running average over 10 ps, we impacted the phase diagram. In Figure E.1, we show a comparison between the averaged and raw data for the continuous heating phase diagram from Figure 4.1. The effect of this data smoothing is that the phase diagram becomes more interpretable, without changing the actual phase behaviour from the ReaxFF simulations of $CsPbI_3$. The lattice vector fluctuations that result from the molecular dynamics simulations are damped, but the phase transition temperatures remain the same.

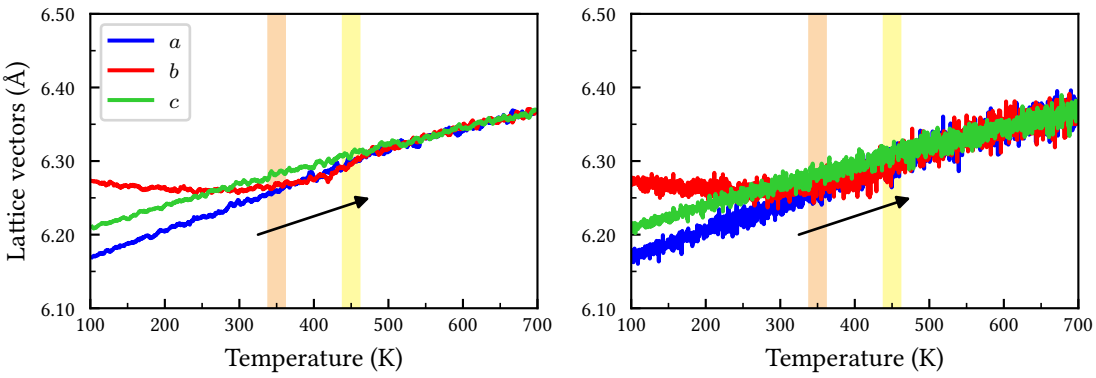

Figure E.1. A comparison of continuous phase diagrams, with (left) smoothed data from a 10 ps running average and (right) the raw data.

Up to now, we have only covered continuous phase diagrams with a heating or cooling rate of $5.0 \cdot 10^{-4}\,\mathrm{K\,fs^{-1}}$, which might incidentally result in a phase behaviour that matches well with experimental behaviour. Thus, to confirm we are looking at the robust and intrinsic phase behaviour of our $CsPbI_3$ ReaxFF parameter set, we generated phase diagrams with different heating rates. A comparison of the different rates is found in Figure E.2. From the comparison, we learn that all continuous phase diagrams result in identical phase behaviour with similar phase transition temperatures irrespective of the heating rate. So, we have confirmed that the continuous phase diagrams in subsection 4.1.1 indeed show the intrinsic phase behaviour of our $CsPbI_3$, which is robust to the used temperature ramp rates.

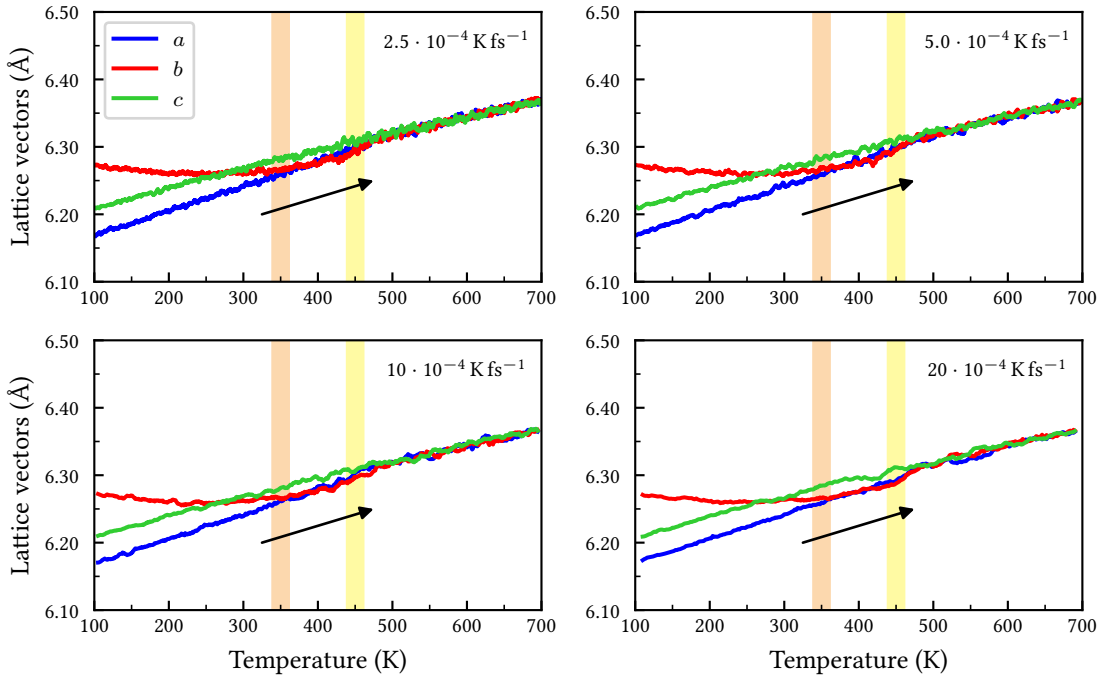

Figure E.2. An overview of continuous phase diagrams from heating runs, each with a different heating rate.

Finally, we check the sensitivity of the continuous phase diagrams to the types of temperature and pressure control in the *NpT*-ensemble. All our previous phase diagrams with the *temperature ramp*-approach, had the temperature and pressure controlled with an NHC-thermostat and MTK-barostat. To determine whether the phase behaviour is sensitive to the type of thermo- and barostat, we carried out some reactive molecular dynamics simulations in which we used a Berendsen thermostat and Berendsen barostat. The damping constants were identical to the ones used in our other simulations: $\tau_\mathrm{T} = 100\,\mathrm{fs}$ and $\tau_\mathrm{p} = 2500\,\mathrm{fs}$. Figure E.3 shows the result of the comparison of the phase diagrams from the different temperature and pressure control mechanisms. The Berendsen-Berendsen controlled phase diagram exhibits a noticeably lower amount of lattice vector fluctuations than the NHC-MTK controlled phase diagram. This results from the suppression of fluctuations in the temperature and pressure in a Berendsen-Berendsen controlled phase diagram [50]. Nevertheless, the phase diagram appears to be the same for both

types of phase diagrams, demonstrating that the phase behaviour appears to be insensitive to the used temperature and pressure control mechanisms.

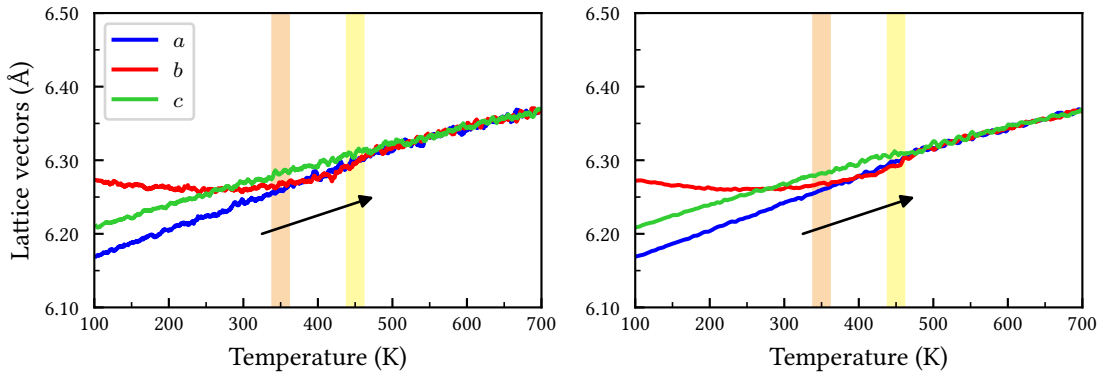

Figure E.3. A comparison of a different thermo- and barostat for the phase behaviour from a continuous phase diagram. The simulations used (left) an NHC-thermostat and MTK-barostat and (right) a Berendsen thermostat and Berendsen barostat.

Cooling runs

Having noticed that the lattice vectors of our $CsPbI_3$ structure are not fully reversible in the simulations with our ReaxFF parameter set, we aim to obtain some atomistic insights into this incomplete reversibility in this section.

To this end, we compare the phase diagrams and structures of a cooling run that exhibits complete reversibility to those from one that only shows partial reversibility. The investigated structures were obtained from the atomic positions of the final simulation frame, representing the structures into which the $CsPbI_3$ system was cooled. Both, the phase diagrams and a snapshot of the top and side view of a part of the structures are shown in Figure E.4. The phase two types of phase diagrams exhibit a distinct difference. In the partially reversible phase diagram, the lattice vectors do not show an as distinct splitting of the lattice vectors at low temperatures as seen in a completely reversible phase diagram. From this observation, we hypothesized that in the partially reversible cooling run, the $CsPbI_3$ structure does not adopt a structure with a single orthorhombic domain at low temperatures.

To understand this behaviour at the atomic scale, we shift our attention to the atomic structures. A closer inspection of the top view of $CsPbI_3$ for the completely reversible cooling run shows that the octahedra in each consecutive layer of the perovskite are rotated in phase. The perovskite has thus adopted an orthorhombic phase with a single orientation (denoted with: □), and is thus a single orthorhombic domain. When we take a closer look at the partially reversible cooling run's top view, we observe a staggered arrangement of the PbI_6 octahedra. This arrangement results from an out of phase rotation of the octahedra in the consecutive perovskite layers.

Thus, the perovskite adopts a structure composed of differently oriented orthorhombic domains, something that is also found in metal halide perovskite experimentally, where it is referred to as *'twinning'* [102, 103]. In the snapshot, two different domains are depicted (denoted with: □ and △).

All in all, this comparison has made it clear that whenever a high-temperature perovskite structure is cooled down to a lower temperature, the entire perovskite adopts the structure of the orthorhombic phase. In a completely reversible cooling run, the perovskite obtains an orthorhombic structure that consists of a single orthorhombic domain. Whenever we named the cooling run reversible, the $CsPbI_3$ crystal also fully adopts an orthorhombic structure. However, the overall structure then consists of multiple orthorhombic $CsPbI_3$ domains with a different orientation. The fact that we observe these two types of cooling runs, a completely and partially reversible run, is presumably caused by the structural fluctuations that occur within the perovskite during the cooling process. Whenever the perovskite is cooled down, the material is frozen into the less dynamic structures at a lower temperature. These low-temperature structures can be a single domain with a single orientation or one structure with multiple differently oriented domains, depending on the local fluctuations occur within the simulation whenever the structure freezing (read: phase transition) occurs.

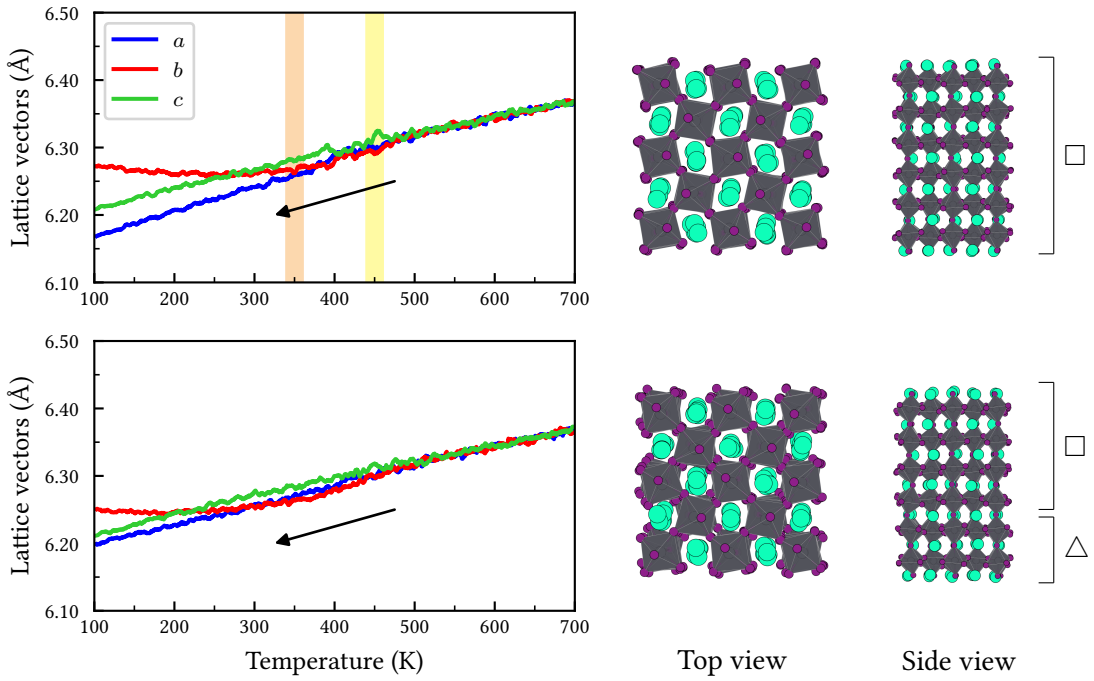

Figure E.4. A comparison of two cooling runs of $CsPbI_3$, with (top) a run that exhibits complete reversibility of the phase transitions and (bottom) a simulation that only shows partial reversibility. Snapshots of parts of the low-temperature structures of these cooling runs are shown next to the phase diagrams.

F. Framework harmonicity

This appendix shows the derivation of the probability distribution for the displacement parameter δ that results from the harmonic approximation. The final analytical form of the harmonic probability distribution was not explicitly shown in the work of Carignano *et al.* [105], the content of this appendix should thus provide some clarity on the exact equation that was used to fit the data from Figure 4.2.

It is important to note that this derivation is done in a polar coordinate system. The line that connects the neighbouring Pb species is taken as the origin and the displacement of the I species away from this line, which we referred to as δ, can then be described in the polar coordinates: r and ϕ. In the harmonic approximation, we express the potential U that describes the displacement of the iodine atom as

$$U(r, \phi) = \frac{1}{2} k r^2, \tag{F.1}$$

where k is an arbitrary force constant. The lack of any angular dependence on ϕ is a result of the isotropy of a harmonic oscillator. The combination of this potential with the Boltzmann distribution allows us to construct a polar probability density p as

$$p(r, \phi) \propto \exp\left(-\frac{U}{k_\mathrm{B} T}\right) = \exp\left(-\frac{r^2}{2\sigma^2}\right) \tag{F.2}$$

where in the final step we filled out the harmonic potential and rewrote the expression so that it mimics the form of a normal distribution with $\sigma = \sqrt{\frac{k_\mathrm{B} T}{k}}$. The probability density can subsequently be used to obtain $P(r)\, dr$, the probability of finding the iodine atom in the interval from r to $r + dr$, as

$$P(r)\, dr = \left[\int_0^{2\pi} p(r, \phi)\, r\, d\phi\right] dr \propto r \exp\left(-\frac{r^2}{2\sigma^2}\right) dr\,, \tag{F.3}$$

where we integrated over the polar angle ϕ to eliminate it from the expression. The probability distribution for r is then

$$P(r) \propto r \exp\left(-\frac{r^2}{2\sigma^2}\right), \tag{F.4}$$

which we use to probe the anharmonicity in the model systems in subsection 4.1.2.

G. Determining diffusion constants

This appendix serves as a detailed overview of the application of Einstein's method of diffusion on the defect simulations in $CsPbI_3$. Here, we cover the computation of the mean square displacement (MSD) and the subsequent calculation of the diffusion constants in some more detail.

The MSD of the different atom species was obtained using a Fast Fourier Transform-based algorithm as implemented in the `MDAnalysis` library library [158–161]. The implementation takes into account both the time- and ensemble-average seen in Equation 4.3. Unwrapped trajectories were used in the computation of the MSD to prevent the back mapping of atoms from causing any artefacts in the MSD of the atoms. A comparison of the MSDs obtained for different types of iodine point defects at 500 K is shown in Figure G.1. In both simulations, the MSD of the iodine species at intermediate timescales is found to scale linearly with time, indicating diffusion of iodide ions through the perovskite. The absence of any slope for Pb and Cs means these species exhibited no net migration through the bulk. During simulations of a perfect crystalline $CsPbI_3$ lattice, the MSD of iodine did not scale linearly with time, which emphasizes the importance of defects in metal halide perovskites for ion migration to occur. Finally, the MSD was used to identify perovskite decomposition. In cases of the decomposition of the perovskite structure, the MSD of Pb and Cs species also exhibited a time dependence, allowing for the identification of perovskite decomposition based on the MSDs.

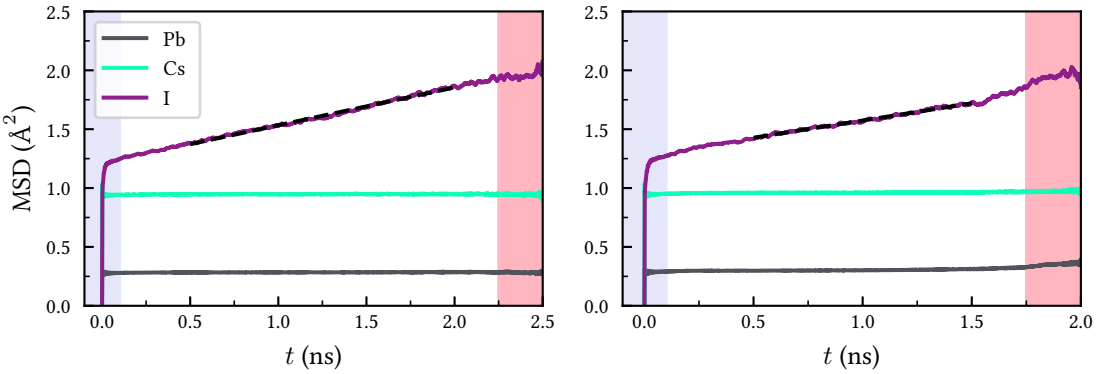

Figure G.1. MSDs obtained for $CsPbI_3$ at 500 K containing (left) two iodine interstitials and (right) two iodine vacancies, with the time section divided into three distinct regions: ballistic (blue), diffusive (white), noisy (red). A linear fit (dashed line) is made to the diffusive regime of iodine to determine the diffusion constant.

Having recognized a '*diffusive*' region at intermediate timescales, we now take a step back to investigate the regions that exhibit some anomalous behaviour. The diffusive region is preceded by a range of short timescales during which we observe a rapid increase of the displacement of the particles. This displacement is not the result of the diffusion of particles through the material, but is connected to the vibrations of the atoms around their equilibrium position in the lattice. Here the magnitude of the MSD is an indication of the spatial extent with which the atoms vibrate[24]. Since the motion of the atoms in such vibrations typically is not impeded by its surroundings, we name this region of the MSD '*ballistic*'. At long timescales, we notice that the MSDs become progressively noisier. This is the result of a lack of data points for these long timescales that can dampen out the fluctuations with a time-average and therefore this region is named '*noisy*'. The inclusion of either of these anomalous regions in determining the diffusion coefficients negatively impacts the values found, thus leading to inaccuracies in the analysis.

To determine the diffusion coefficients, the slope of each MSD was determined with a linear fit. Considering the above discussion of the different regions, we only applied the Einstein analysis to the diffusive region. A linear fit was therefore made on the time interval from 0.5 ns to 2.0 ns and from 0.5 ns to 1.5 ns for the $CsPbI_3$ systems with interstitials and vacancies, respectively. A complete overview of the values determined from this analysis is shown in Table 4.1.

[24]Notice that in $CsPbI_3$ the atoms, when ordered from largest to smallest vibrations around their equilibrium positions in the lattice, are I, Cs and Pb. An observation that is supported by the rapid successive tilting seen at high temperatures for $CsPbI_3$ which results in considerable displacements for the iodine ions in particular.