UNIVERSITY OF CALIFORNIA
RIVERSIDE

Atomistic Modeling of Amorphous Materials

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Materials Science & Engineering

by

Sarah I. Allec

September 2020

Dissertation Committee:

    Dr. Peter Alexander Greaney, Chairperson
    Dr. Sinisa Coh
    Dr. Mahesh Neupane

The Dissertation of Sarah I. Allec is approved:

_____

_____

_____

Committee Chairperson

University of California, Riverside

Acknowledgments

I am grateful to my advisor, without whose help, I would not have been here.

To my parents for all the support.

ABSTRACT OF THE DISSERTATION

Atomistic Modeling of Amorphous Materials

by

Sarah I. Allec

Doctor of Philosophy, Graduate Program in Materials Science & Engineering
University of California, Riverside, September 2020
Dr. Peter Alexander Greaney, Chairperson

Amorphous (non-crystalline) materials offer new capabilities and applications relative to their corresponding crystalline phases due to their novel and tunable material properties, including but not limited to larger surface area, wider substrate compatibility, stronger corrosion resistance, and lower temperature processing than their crystal counterparts. However, even the most advanced experimental techniques (e.g., *in situ* NMR [1, 2] and "inverse" approaches such as Reverse Monte Carlo (RMC) modeling of diffraction data [3, 4, 5]) cannot directly measure the bulk atomic structure of such complex, non-crystalline materials. This, coupled with the lack of a structural theory of amorphous materials, has limited our understanding about the structure-property relationships required for the optimization of realistic devices. In this thesis, we compare and combine several atomistic modeling techniques for generating accurate and general structures of amorphous systems, both solid and liquid. We develop several metrics for measuring efficiency of configuration space sampling to guide and standardize the computational study of amorphous materials, and use these to differentiate between different samples of amorphous structures. Lastly, we use these structures to help elucidate structure-property relationships in these systems. In amor-

phous metal oxides, we characterize the types of coordination defects in amorphous metal oxide systems and subsequently identify potential electron and hole traps in their electronic structures. In a liquid system, water-in-salt electrolytes, we evaluate how the hydrogen bonding network and ion coordination shells change with composition, and how these are directly correlated to battery performance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Amorphous (non-crystalline) materials offer new capabilities and applications relative to their corresponding crystalline phases due to their novel and tunable material properties, including but not limited to larger surface area, wider substrate compatibility, stronger corrosion resistance, and lower temperature processing than their crystal counterparts. However, even the most advanced experimental techniques (e.g., *in situ* NMR [1, 2] and "inverse" approaches such as Reverse Monte Carlo (RMC) modeling of diffraction data [3, 4, 5]) cannot directly measure the bulk atomic structure of such complex, non-crystalline materials. This, coupled with the lack of a structural theory of amorphous materials, has limited our understanding about the structure-property relationships required for the optimization of realistic devices.

In this thesis, we compare and combine several atomistic modeling techniques for generating accurate and general structures of amorphous systems, both solid and liquid. We develop several metrics for measuring efficiency of configuration space sampling to guide and standardize the computational study of amorphous materials, and use these to determine

what constitutes a "good" amorphous structure. Having identified the most representative structures for each system, we compute various material properties, including glass transition temperatures and density of states, to elucidate the structure-property relationships in these systems and guide experimental synthesis and design. We have found that the level of theory (classical or *ab initio*) as well as the particular details of the simulation (e.g. quench rate) severely bias the final amorphous structure, necessitating the generation of a diverse set of amorphous structures in order to properly describe structure-property relationships of these complex systems. This proves to be particularly important for the identification of electron and hole traps in amorphous metal oxide systems, as the density of traps in realistic systems is low relative to the system sizes accessible in simulations. In particular, we *i)* discover a possible *deep* trap in amorphous indium oxide which has not been previously identified in the literature and *ii)* possible traps in amorphous vanadium oxide which cannot be categorized as deep, but are electronic structure outliers, which we coin *topographic* traps.

In the remainder of this introduction, we give a definition of amorphous materials (i.e. what it means for a system to be amorphous), describe some ways that order is quantified, both experimentally and theoretically, and provide some background information on the systems and methods covered in this thesis. In chapter 2, we give a detailed description of the various simulation methods employed, followed by a presentation of the metrics we have developed to measure efficiency of configuration space sampling in chapter 3. In chapter 4, we present our results on two amorphous metal oxide systems, amorphous indium (III) oxide (a-In_2O_3) and amorphous vanadium (V) oxide a-V_2O_5, and in chapter 5, we present our results on a liquid system, zinc chloride water-in-salt electrolytes. Finally,

in chapter 6, we conclude with a summary of our findings, including a comparison of the different modeling techniques evaluated in this thesis and a guide to generating accurate amorphous structures entirely *in silico*.

1.1 Properties of Amorphous Materials

In order to properly model noncrystalline materials, we need to be able to distinguish between crystalline and amorphous phases of a given material. The fundamental difference between periodic materials, which encompasses single crystal, polycrystals, and quasicrystals, and amorphous materials is that amorphous materials *lack both rotational symmetry and translational symmetry (periodicity)*, examples of which include liquids and glasses. Single crystals are both ordered (rotationally symmetric) and periodic over all distances, while polycrystals are ordered and periodic over smaller distances (i.e. within a grain) and quasicrystals are ordered but not periodic. Amorphous materials display short-range order and medium-range order to an extent, but no long-range order or periodicity. A comparison of these types of materials is given in Table 1.1.

Whether a material is periodic, ordered, both, or neither dictates how easily we can determine its atomic structure. For a perfect crystal, which is both periodic and symmetric,

	Ordered	Periodic
Single crystals	Yes	Yes
Polycrystals	Yes	Locally
Quasicrystals	Yes	No
Amorphous materials	No	No

Table 1.1: Order and periodicity in crystalline and amorphous materials

we can identify a pattern (a symmetry) M that is repeated in three dimensions:

$$\vec{r}_{atom} = M(\vec{r}_0, atom) + u \cdot \vec{a} + v \cdot \vec{b} + w \cdot \vec{c}, \tag{1.1}$$

for some starting motif M and lattice vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$. For polycrystals, equation 1.1 holds true within each grain of the material. On the other hand, for quasicrystals, we can identify patterns M, but such patterns are not repeated periodically. Lastly for amorphous materials, we can neither identify such patterns M nor predict the atomic positions by any relationship like equation 1.1 (c.f. Figure 1.1). Such a lack of periodicity means that we cannot reduce the atomic structure to a finite unit cell, i.e. materials that lack periodicity (quasicrystals and amorphous materials) have *infinite* unit cells. Because atomic simulations require that the atomic coordinates fit in a finite simulation box, it is impossible to simulate a realistic non-periodic material. However, in addition to the challenges of modeling an extremely large unit cell, amorphous materials also do not display rotational symmetry, as the coordination of a given atom varies from site to site due to an excessive amount of coordination defects.

However, this lack of long-range order does not mean that amorphous materials are completely disordered. In fact, all amorphous materials exhibit *short-range order*, a regular, predictable atomic arrangement over distances of one or two atoms. We can think of short-range order as the first coordination shell of an atom, defined by the number N_j of nearest-neighbor atoms of type j (each with connectivity α_j) within some bond distance r_{ij} of an atom i. For example, in the ideal crystal structure of silicon, every atom is tetrahedrally coordinated, while in the amorphous phase, there is a range of predominantly 3-, 4-, and 5-coordinated atoms. Thus, the atomic structure of amorphous materials is *not*
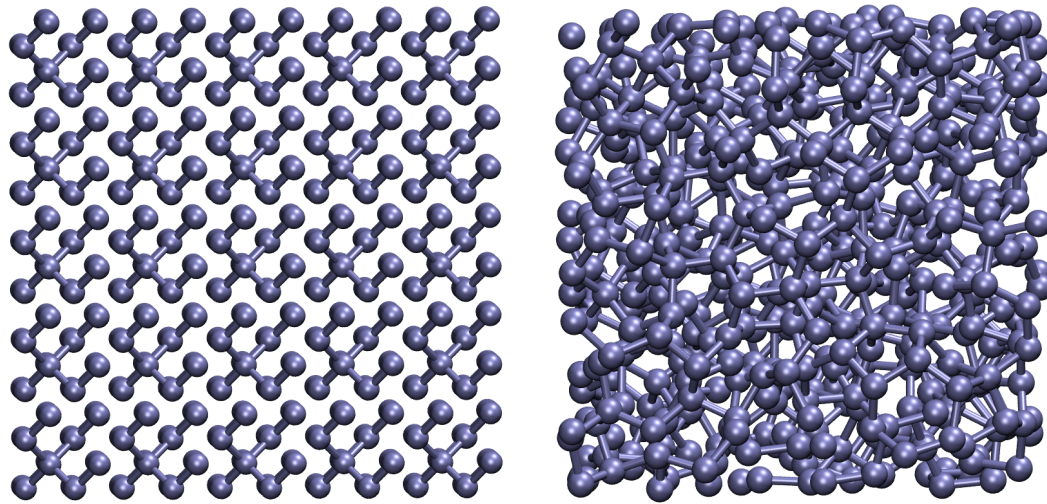
Figure 1.1: Supercells of crystalline silicon *(left)* and amorphous silicon *(right)*.

random but rather excessively disordered or defective. These parameters of short-range order also define a coordination polyhedron, which links together with other polyhedra to build the entire structure. As we will see in the next sections, while amorphous materials retain individual polyhedra, the network of polyhedra is distorted such that there is no longer any long-range order.

These two types of order can be probed (to an extent) with experimental techniques based on how atoms scatter radiation in the material: Long-range order manifests itself in sharp diffraction peaks and short-range order in radial distribution functions. In terms of diffraction, amorphous solids and liquids exhibit *smooth* structure factors, as opposed to the sharper, more well-defined peaks of crystalline structure factors. As the radial distribution function is related to the structure factor by a Fourier transform, we can examine the radial distribution functions to see these differences. Looking at Figure 1.2, which displays the radial distibution functions of crystalline and amorphous indium oxide, we see the retention

of short-range order at lower r values and the loss of long-range order as r increases. Because of this, the traditional wave theory used for determining the atomic structure by inverting the diffraction peaks of crystalline materials is insufficient for such structure factors. Thus, computational modelling plays a huge role in the determination of the atomic structure of amorphous materials and the discovery of resulting structure-property relationships.
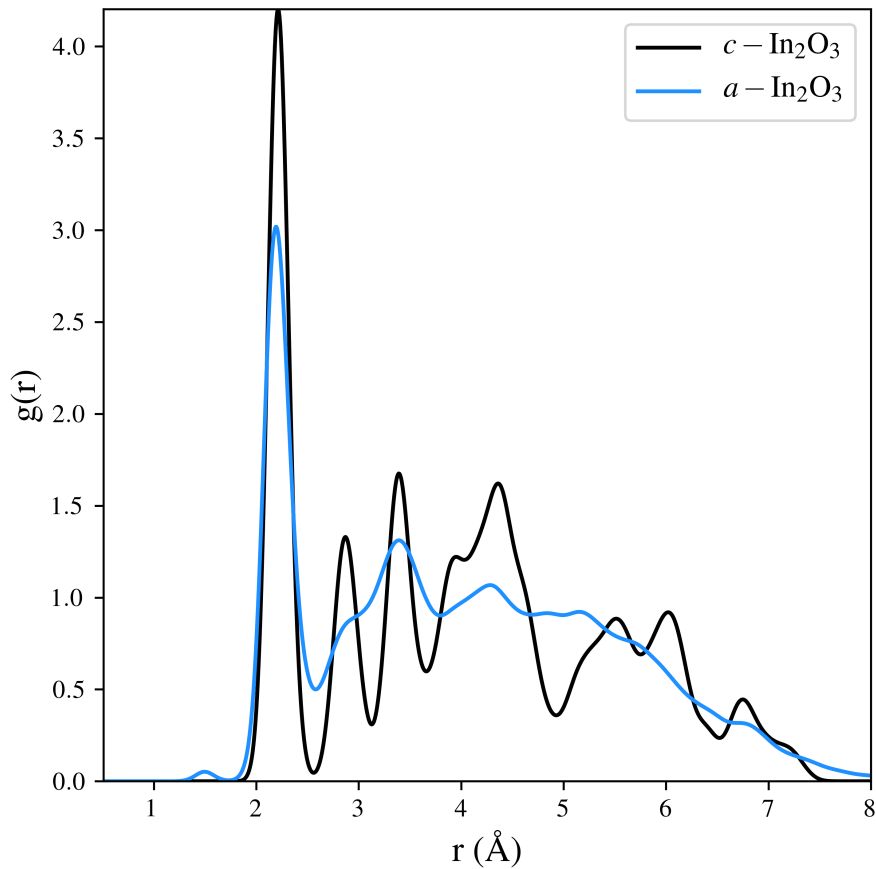
Figure 1.2: The radial distribution functions $g(r)$ of crystalline c-In$_2$O$_3$ (black) and amorphous a-In$_2$O$_3$ (blue).

In general, there are two paradigms for determining the atomic structure of amorphous materials computationally: the "simulation paradigm" and the "information paradigm".

As the name implies, the simulation paradigm employs some type of computer simulation, usually molecular dynamics (MD) or Monte Carlo (MC), with suitable interatomic potentials. On the other hand, the "information paradigm" seeks to invert experimental diffraction data via Reverse Monte Carlo (RMC) without any interatomic potential, i.e. purely with "information" from the diffraction data. Historically, these two paradigms operated independently of each other, each with its own drawbacks: In general, RMC models tend to be maximally disordered and chemically unrealistic, while MD and MC simulations are limited by size scale and time-scale limitations, imperfect interatomic potentials, and a physically unrealistic amorphization process. More recently, approaches bridging the two have been developed [6, 7, 8, 9, 10, 11, 12, 13, 14], with *ab initio* force-enhanced atomic refinement (AIFEAR) [15] leading the way to a more general framework for the inversion of diffraction data.

Nonetheless, an accurate and efficient purely computational approach still offers several advantages, particularly the speed and low cost of *in silico* experiments relative to physical experiments. In this thesis, we address the challenge of atomic structure prediction faced in the modelling of two subcategories of amorphous materials: i) amorphous solids and ii) liquids.

1.1.1 Amorphous Metal Oxides

Traditionally, the study of amorphous solids has primarily been restricted to the study of glasses - materials that exhibit a glass transition temperature. The most common glasses are either metallic (e.g. bulk metallic glasses) or covalently-bonded structures (e.g. silicate), which exhibit different types of packing and consequently different properties

relative to each other. In general, metallic glasses are well-described by the dense random packing of hard spheres, while covalent glasses have sparsely packed, strongly bound network structures.

These two categories of glasses represent two extremes of amorphous solids. In this thesis, we will concern ourselves with an intermediate of these two extremes: amorphous metal oxides (a-MOs). This family of amorphous solids has widespread applications in batteries, supercapacitors, electronics, conducting films, multilayered transistors, electrochromic displays, and nonvolatile memories [16, 17, 18]. Such a wide variety of applications of (a-MOs) is due to the diversity of this family of materials, whose properties can be tuned by changing the metal cation, as well as the level of disorder.

The basic building blocks of (a-TMOs) are M-O$_x$ polyhedra, where M is a transition metal, O is oxygen, and x denotes the number of oxygen atoms coordinated to a single M atom. The polyhedra are interconnected through different types of sharing configurations known as edge-sharing, corner-sharing, and face-sharing configurations, based on the number of shared oxygen atoms. In general, corner-sharing polyhedra (one shared oxygen) are the most stable because the distance between the metal cations is largest. The next stable are edge-sharing polyhedra (two shared oxygens), and the least stable are face-sharing polyhedra (three shared oxygens), which are virtually nonexistent in ionic structures. Examples of corner-sharing and edge-sharing polyhedra are shown in Figure 1.3 for In$_2$O$_3$.
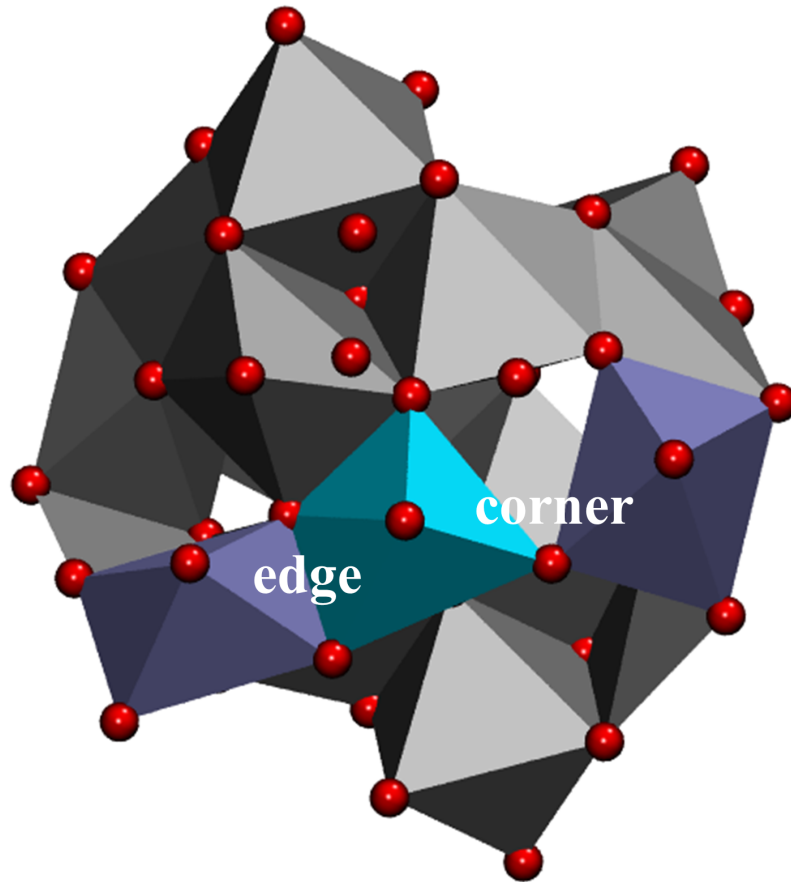
Figure 1.3: Polyhedra network of c-In$_2$O$_3$.

Identification of Electron & Hole traps

The motivation of this research is the use of vanadium pentoxide (V$_2$O$_5$) as a surface electron acceptor in diamond field-effect transistors (FETs) to effectively dope diamond via *surface transfer doping* [19]. Optimization of these devices depends on being able to tune the electron traps in the V$_2$O$_5$ layer. A trap is any location in a solid that can restrict (immobilize) an electron or hole carrier, requiring a significant amount of energy to remove the carrier from the trap. A trap occurs either at a chemical impurity or at an imperfection

in the crystal lattice. Because amorphous materials contain an excessive number of defects, they can contain a large number of possible traps.

The contributions of this thesis toward this goal are *i)* the generation of a variety of amorphous V_2O_5 (a-V_2O_5) structures and *ii)* the identification of possible traps in these samples. Because the atomic structure of (a-V_2O_5) is not well-studied, we first validate our computational procedure by generating structures of a well-studied amorphous oxide, indium oxide (a-In_2O_3). Thus, in this thesis, we exclusively focus on two amorphous oxides, a-V_2O_5 and a-In_2O_3.

1.1.2 Water-in-Salt Electrolytes

Another broad family of amorphous materials are liquids, which only display short-range order and exhibit lower connectivity between atoms/molecules than solids. They are also characterized by higher molecular motion than solids. In this thesis, we focus on one liquid system, water-in-salt electrolytes (WiSE). Electrolyte composition has direct impact on the safety, sustainability, and performance of batteries. Organic electrolytes exhibit excellent performance but are plagued by safety concerns due to their flammability and toxicity. On the other hand, aqueous electrolytes are a safe alternative but have low energy densities due to the narrow electrochemical stability window of water. The recent discovery of highly concentrated aqueous electrolytes, WiSE, extended the voltage window of aqueous electrolytes and enabled the development of the first reversible Zn-metal batteries. These batteries are safer, more sustainable alternatives to lithium-ion batteries, particularly for energy storage on the grid and electric vehicles.

The mechanism underlying the enhanced performance of WiSE arises from the

short-range order around the cations of the salt: In such highly concentrated solutions, all water molecules are a part of ion solvation shells, leaving no "free" water molecules to participate in parasitic reactions that degrade the electrodes. However, the chemistry of such highly concentrated solutions are outside of the realm of basic solubility rules, hindering the optimization of the battery electrochemistry such that the cell voltage, cyclability, and reversibility remain below industry standards. Interestingly, Wang et al. recently discovered that use of a bi-salt WiSE, consisting of a lithium salt and a zinc salt, drastically improves the performance of Zn-metal batteries. However, the interplay between the concentrations of the two salts is not well understood. Here, we use *ab initio* molecular dynamics to correlate the local water environments in these systems directly to measures of battery performance.

1.1.3 Overview of Methods

We generated the various atomic structures of each amorphous metal oxide non-empirically using *ab initio* and classical liquid-quench molecular dynamics (LQMD) simulations. In particular, we developed neural network interatomic potentials (NNP) fit to the *ab initio* LQMD data, which were subsequently used in classical LQMD in order to access larger system sizes and longer simulation times. The use of NNPs allowed us to generate more general models while retaining (to a degree) the accuracy of DFT. The flow of the computational framework is shown in Figure 1.4.

As WiSE are liquids, LQMD is not an applicable method for structure generation. In this case, we utilized a packing optimization strategy of the molecular components that
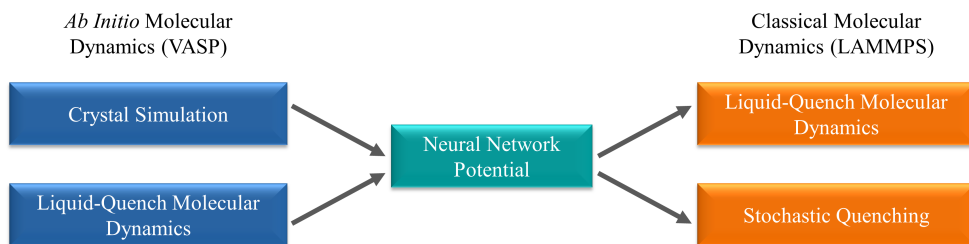
Figure 1.4: Computational framework utilized in this thesis.

balances the short-range interactions between the molecules in the system. Each system was

packed into a box consistent with the experimental densities of these systems. An example

of such a system is shown in Figure 1.5.
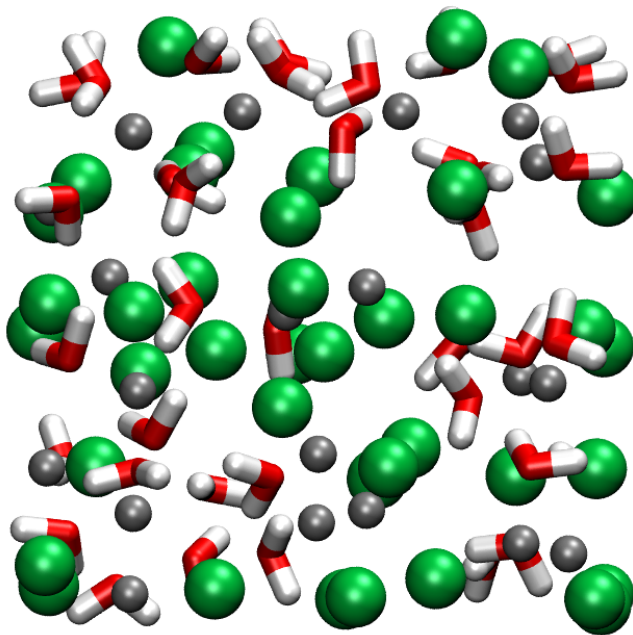
Figure 1.5: ZnCl$_2$ water-in-salt electrolytes (WiSE) studied in this thesis. Chlorine atoms
are shown in green, zinc atoms in gray, oxygen atoms in read, and hydrogen atoms in white.

Chapter 2

Methods

The basis of all the simulations in this thesis is *ab initio* molecular dynamics (AIMD). One of the biggest challenges in simulating amorphous systems is the generation of the atomic structure (i.e. the *xyz*-coordinates for every atom in the system). As discussed in the previous chapter, there is no overarching model of the atomic structure for amorphous materials like there is for crystalline materials. Because of this, we have to perform some modeling "tricks" with AIMD in order to effectively search the configuration space of these systems: liquid-quench MD for the amorphous TMOs and metadynamics for the water-in-salt electrolytes. We start with an overview of AIMD.

2.1   Ab Initio Molecular Dynamics

Molecular dynamics (MD) is a general term that describes any simulation method that evolves a system of particles in time according to Newton's second law, $\vec{F} = m\vec{a}$. From MD simulations, we can compute equilibrium thermodynamic and transport properties.

The "particles" in an MD simulation can be proteins, molecules, or atoms, depending on the desired resolution and computational limitations of the problem at hand. For the entirety of this thesis, we will only concern ourselves with atomistic MD, i.e. the particles are individual atoms. In this case, we can think of MD as a computational microscope of atomic resolution.

One of the caveats of molecular dynamics is the choice of interatomic potential, which provides the forces on all of the atoms (i.e. the $\vec{F}$ in $\vec{F} = m\vec{a}$). In traditional classical MD, the interatomic potential is parametrized either from experimental data or from more accurate simulations; either way, the potential is restricted to the domain of systems on which it is parametrized. To put this more succinctly, parametrized potentials are not transferable or universal. Nonetheless, for well-parametrized potentials, classical MD can be transformative and illuminating for a variety of problems, particularly in biology, and its computational efficiency is unprecedented in atomistic modeling.

Nonetheless, traditional interatomic potentials cannot treat problems that are very sensitive to changes in the electronic charge distribution over the system. This is particularly challenging for the generation of amorphous structures, which involves a significant degree of bond breaking/forming, and for systems that involve hydrogen bonding, which require an accurate description of polar molecules. In these cases, ab initio MD (AIMD) is ideal, as the forces are computed on-the-fly from the electronic structure and the electronic charge distribution is re-computed at each timestep. Of course, this increase in accuracy comes with severe limitations in system size and timescale; we will discuss how we compensate or overcome these challenges later in the text.

While it may seem like a stretch to describe material properties simply from the

movement of electrons and nuclei, at a fundamental level, that is all materials are: collections of electrons and nuclei. It is often helpful to keep the following conceptual equation in mind:

$$\text{materials} = \text{electrons} + \text{nuclei}. \tag{2.1}$$

Within the framework of AIMD, there are two levels of theory for the two types of particles in equation 2.1: classical mechanics for the nuclei and quantum mechanics for the electrons. Because nuclei and electrons typically move on different timescales, the motion of the nuclei and the motion of the electrons are decoupled, according to the Born-Oppenheimer approximation [20]. The equation of motion for the nuclei is Newton's second law and the equation of motion for the electrons is the time-independent Schrödinger equation. We proceed first with the treatment of the nuclei (section 2.1.1) and subsequently with that of the electrons (section 2.1.2). Treating the atoms classically is justified for most a wide range of systems, as quantum effects for nuclei only become significant for light atoms or for molecules with vibrational frequencies greater than $k_B T$.

2.1.1 Molecular Dynamics

The general idea behind molecular dynamics (MD) simulations shares many similarities with real experiments, where we prepare a sample of a material, take measurements of the properties of this sample, and then average these measurements over some period of time. In an MD simulation, our sample is a model system consisting of M atoms, for which we numerically solve the classical equation of motion, Newton's second law, until the property of interest has equilibrated. For such a system consisting of M atoms, this is given

by

$$M_I \mathbf{R}_I = \mathbf{F}_I \tag{2.2}$$

for each atom, with $\mathbf{R}_I$ and M_I the position and mass of the Ith atom, respectively. The force on each atom is derived from some potential energy function U:

$$\mathbf{F}_I = -\frac{\partial U}{\partial \mathbf{R}_I}. \tag{2.3}$$

In the case of AIMD, U is derived from the electron density obtained via density functional theory, as described at the end of the next section. Ideally, we would like to consider the contributions to $\mathbf{F}_I$ due to all neighbors of atom I; however, practically, we only consider interactions within some cutoff distance r_c of each atom.

From the derived atomic forces, we evolve the system in time according to the Verlet algorithm: At every timestep Δt, we move the atoms according to

$$\mathbf{R}_I^{j+1} = 2\mathbf{R}_I^j - \mathbf{R}_I^{j-1} + \frac{\mathbf{F}_I^j}{M_I}\Delta t^2 + \mathcal{O}(\Delta t^4) \tag{2.4}$$

where j indexes time and $\mathbf{R}_I^j$ denotes the position of atom I at timestep j. We can similarly compute the trajectories of the velocities $\mathbf{v}$ with the same form as equation 2.4. Initial velocities for each atom are randomly assigned according to the desired initial temperature of the system at the beginning of the simulation, utilizing the following definition of instantaneous temperature:

$$k_B T(t) = \sum_{I=1}^{M} \frac{M_I v_{\alpha,I}^2(t)}{N_f}, \tag{2.5}$$

where $v_{\alpha,I}(t)$ is the α component of the velocity of atom I at time t and $N_f = 3M - 3$ is the number of degrees of freedom in the system (here, we use t to denote time instead of j for indexing to avoid clutter of superscripts).

The main task of MD is to compute the expectation value of an arbitrary operator

$\langle A \rangle$:

$$\langle A \rangle = \frac{\int d\mathbf{R} d\mathbf{P} A(\mathbf{R}, \mathbf{P}) e^{-\beta E(\mathbf{R}, \mathbf{P})}}{\int d\mathbf{R} d\mathbf{P} e^{-\beta E(\mathbf{R}, \mathbf{P})}}, \tag{2.6}$$

with Boltzmann distribution $e^{-\beta E(\mathbf{R}, \mathbf{P})}$. The energy of the system is given by

$$E(\mathbf{R}, \mathbf{P}) = \sum_{I=1}^{M} \frac{\mathbf{P}_I^2}{2M_I} + U(\mathbf{R}), \tag{2.7}$$

where the first term denotes the nuclear kinetic energy and the second term is the potential

energy.

Solving the high-dimensional integrals of equation 2.6 is often impractical and

inefficient for large, complex systems. Thankfully, under the ergodic hypothesis, we can

equivalently compute the temporal average. Armed with the trajectories $\mathbf{R}(t)$ and momenta

$\mathbf{P}(t)$, we can compute the thermal average of A as

$$\langle A \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int dt A(\mathbf{R}(t), \mathbf{P}(t)) \tag{2.8}$$

From equation 2.8, it is obvious that we must be able to express the property of interest

as a function of the positions and momenta of the atoms in the system. The connection

between MD and thermodynamics is given by the equipartition theorem:

$$\left\langle \frac{1}{2} \sum_{I}^{M} M_I^2 \dot{\mathbf{R}}_I^2 \right\rangle = \frac{3}{2} N k_B T \tag{2.9}$$

In practice, we must ensure that the length scale and time scale of our systems are appro-

priate for the property to be computed. For example, only processes with relaxation times

much less than the total simulation time and with characteristic correlation lengths much

less than the simulation box can be reliably computed with MD. This is a significant chal-

lenge for AIMD, thus requiring some creativity in computing many properties, leading to

17

the aforementioned "tricks" that can help accelerate the search of (or enhance the sampling of) configuration space.

The Potential Energy Surface

The potential energy surface (PES) is a convenient conceptual tool used in the analysis of MD trajectories, particularly for understanding structural stability and reaction dynamics. In general, the PES is a multi-dimensional real-valued function of some set of internal coordinates that yields the potential energy of the system. For our purposes, the internal coordinates are the positions of all the atoms in the system, i.e., the PES is given by $U(\mathbf{R})$ in equation 2.3, where $\mathbf{R}$ denotes the coordinates of all the nuclei, for all relevant values of $\mathbf{R}$.

Of primary importance are the stationary points of the PES, where the gradients of U with respect to all atomic coordinates are zero:

$$\nabla U(\mathbf{R}_I) = 0, \qquad (2.10)$$

for all I nuclei. In particular, minima correspond to stable structures (e.g. reactants and products) and first-order saddle points (where U is a maximum in one direction and a minimum in all other directions) correspond to transition states.

As one could easily imagine, once a system has arrived in a local minimum of the PES, it tends to stay there, as a typically large amount of energy must be supplied for the system to escape a valley. For systems where the atomic structures or reaction pathways of interest are approximately known *a priori*, this generally does not present an issue, as the starting structure(s) for AIMD will be close to the stationary state/pathway of interest or can be directed to it via some sort of biasing technique. However, for amorphous systems, we

18

generally do not have an accurate picture of the atomic structure *a priori*. Simple geometry relaxation or MD equilibration will only lead us to the nearest local minimum, which may not be representative of the real structure. Thus, efficient and accurate sampling of the PES is of critical importance for atomic structure determination of amorphous structures.

2.1.2 Density Functional Theory

We now transition to our calculation of the electron density of our system, from which we obtain the potential energy function U utilized to evolve the nuclei as in equation 2.3. Density functional theory (DFT), or the "*ab initio*" in AIMD, is the electronic structure method of choice for AIMD. In essence, within the framework of AIMD, we perform a DFT calculation at each timestep and move the atoms according to the forces generated by the converged electronic density. As previously mentioned, the equation of motion of concern here is the time-independent Schrödinger equation for the electrons:

$$\hat{H}_{el}\psi_{el} = E_{el}\psi_{el}. \tag{2.11}$$

$\hat{H}_{el}$ is the quantum mechanical Hamiltonian operator and ψ_{el} is the wave function for all of the electrons in the system. What we need to "find" in this equation is ψ_{el}, as the behavior of quantum particles is given by their wave function. As equation 2.11 is an eigenvalue problem, the set of all E_{el} is the spectrum of $\hat{H}_{el}$ with corresponding eigenvectors ψ_{el}. The lowest-energy eigenstate is the ground state of the system, ψ_{ground}.

For many-electron systems in the absence of any external potentials, we know the kinetic and potential energy terms that make up $\hat{H}$:

$$\left(-\sum_i \frac{\nabla_i^2}{2} + \sum_i V_n(\mathbf{r}_i) + \frac{1}{2}\sum_{i \neq j} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \right)\psi = E\psi, \tag{2.12}$$

19

where i runs over all electrons in the system and we have dropped the subscript for E and ψ. The first term on the left-hand side of the equation is the kinetic energy of the electrons, the second term is the electron-nuclear Coulomb attraction, and the third term is the electron-electron Coulomb repulsion.

If we knew the exact mathematical forms of each term in $\hat{H}$ and of ψ, a single diagonalization would provide us with the exact ground state wave function. Unfortunately, we do not know the analytical forms of either for many-electron systems, and solving for the exact forms numerically is infeasible, even with modern supercomputers. One restriction on ψ is the Pauli exclusion principle: two fermions cannot occupy the same quantum state simultaneously, thus ψ must be antisymmetric: if we exchange the positions and spin of any two electrons, ψ must change signs. This is taken care of by using a Slater determinant for ψ built from single-particle electronic wavefunctions. For example, for a simple two-electron system, where the wavefunction for the first electron is ϕ_1 and the wavefunction for the second electron is ϕ_2, the total electronic wavefunction ψ in the form of a Slater determinant is

$$\psi(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(\mathbf{r}_1) & \phi_2(\mathbf{r}_2) \\ \phi_2(\mathbf{r}_1) & \phi_1(\mathbf{r}_2) \end{vmatrix} \tag{2.13}$$

It is straightforward to show that upon swapping $\mathbf{r}_1$ and $\mathbf{r}_2$, the sign of ψ changes.

Much of the difficulty in the Hamiltonian comes from the electron-electron repulsion, which couples all of the electrons to each other - i.e. the behavior of each electron depends on the behavior of all the other electrons *simultaneously*. To alleviate this difficulty, we proceed with two approximations: i) First, we neglect entirely the Coulomb repulsion, which will give a set of decoupled Schrödinger equations for each electron and ii) then we re-

incorporate the presence of other electrons through a classical, mean-field approach. These approximations produce a set of N decoupled single-electron wave functions, a great simplification from the original $3N$-dimensional many-body Schrödinger equation. The resulting single-electron wave functions are then used to build up the Slater determinant for the total wave function ψ.

Independent Electrons Approximation

The first approximation is to remove the Coulomb repulsion from equation 2.12:

$$\left(-\sum_i \frac{\nabla_i^2}{2} + \sum_i V_n(\mathbf{r}_i) \right) = \mathrm{E}\psi. \tag{2.14}$$

In essence, the electrons have become blind to each other and no longer interact. We now introduce the *single-particle Hamiltonian* $\hat{\mathrm{H}}_0$:

$$\hat{\mathrm{H}}_0(\mathbf{r}) = -\frac{1}{2}\nabla^2 + V_n(\mathbf{r}). \tag{2.15}$$

With this notation, equation 2.14 becomes

$$\sum_i \hat{\mathrm{H}}_0(\mathbf{r}_i)\psi = \mathrm{E}\psi. \tag{2.16}$$

Recall that ψ is given by the Slater determinant of the single-electron wave functions ϕ_i. We can easily determine the set of ϕ_i as the solutions of the single-electron Schrödinger equations:

$$\hat{\mathrm{H}}_0\phi_i = \varepsilon_i\phi_i \tag{2.17}$$

The set of N decoupled equations given by 2.17 is much easier to solve than the coupled system of equation 2.12. Furthermore, within this approximation, we can easily write the

21

electron density as

$$n(\mathbf{r}) = \sum_i^{N_{occ}} |\phi_i(\mathbf{r})|^2, \tag{2.18}$$

where the sum is over all occupied states. However, the approximation of entirely non-interacting electrons is abhorrently physically unrealistic. In the next section, we show how we can keep this single-particle description, but take into account the Coulomb repulsion in a mean-field way.

Mean-Field Approximation

Classical electrostatics tells us that a distribution of electronic charge $n(\mathbf{r})$ will generate an electrostatic potential $\Phi(\mathbf{r})$ according to Poisson's equation:

$$\nabla^2 \Phi(\mathbf{r}) = 4\pi n(\mathbf{r}) \tag{2.19}$$

Electrons in the presence of this potential have a potential energy $V_{\mathrm{H}}(\mathbf{r}) = -\Phi(\mathbf{r})$, called the Hartree potential. The formal solution for V_{H} is

$$V_{\mathrm{H}}(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.20}$$

Modeling the presence of other electrons as an effective charge distribution, every electron in our system experiences the Hartree potential. We then improve our single-particle description given by equation 2.17 by adding V_{H}:

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{n}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \tag{2.21}$$

The Self-Consistent Kohn-Sham Equations

Equation 2.21 only takes into account a classical electron-electron interaction, the Hartree potential, which as we should expect, is not sufficiently accurate for describing

quantum particles such as electrons. Here, we briefly show how writing the energy as a functional of the electron density produces another potential term, the exchange-correlation (XC) potential. This is the last piece we need to complete the DFT-description of the electrons.

We can obtain the total electronic energy E according to

$$E = \langle\psi|\hat{H}|\psi\rangle = \int d\mathbf{r}_1...d\mathbf{r}_N\psi^*(\mathbf{r}_1...\mathbf{r}_N)\hat{H}\psi(\mathbf{r}_1...\mathbf{r}_N), \tag{2.22}$$

with $\hat{H}$ the same as in equation 2.12. Because $\hat{H}$ is independent of the choice of material, any change in E must be associated with a change in ψ, or in other words, E is a functional of ψ ($E = \mathcal{F}[\psi]$). This result is generally true for for any quantum state. However, for the ground state in particular, we also have that the energy E_{ground} is a functional of the electron density $n(\mathbf{r})$:

$$E_{\text{ground}} = \mathcal{F}[n(\mathbf{r})] \tag{2.23}$$

This was discovered by Hohenberg and Kohn in 1964 and is known as the first Hohenberg-Kohn theorem [21]. This observation is particularly significant because all we need to determine the total energy of the ground state is the electron density n, which depends only on three variables, as opposed to the $3N$-dimensional many-body wavefunction ψ.

We can write the ground state energy functional as

$$E_{\text{ground}}[n] = \langle\psi|\sum_i V_{\text{n}}(\mathbf{r}_i)|\psi\rangle + \langle\psi|\hat{T} + \hat{W}|\psi\rangle$$
$$= \int d\mathbf{r}n(\mathbf{r})V_{\text{n}}(\mathbf{r}) + \langle\psi|\hat{T}[n] + \hat{W}[n]|\psi\rangle, \tag{2.24}$$

where we have used the symbolic notation

$$\hat{T} = -\sum_i \frac{1}{2}\nabla_i^2, \; \hat{W} = \frac{1}{2}\sum_{i\neq j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.25}$$

for the kinetic energy and electron-electron repulsion, respectively. While we know the explicit dependence of the first term on the density n (see the second line of equation 2.24), we do not know the dependence of $\hat{T}$ and $\hat{W}$ on n. In order to keep the simplicity of the independent electrons approximation while accounting for quantum electron-electron interactions, Kohn and Sham [22] proposed to split the implicit terms into the kinetic energy and Hartree energy of independent electrons, and an extra term that accounts for the difference between the independent electron approximation and equation 2.24:

$$E_{\mathrm{KS}}[n] = \int d\mathbf{r} n(\mathbf{r}) V_{\mathrm{n}}(\mathbf{r}) - \sum_i \int d\mathbf{r} \phi_i^*(\mathbf{r}) \frac{\nabla^2}{2} \phi_i(\mathbf{r}) + \frac{1}{2} \int \int d\mathbf{r} d\mathbf{r}' \frac{n(\mathbf{r}) n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + E_{xc}[n]. \quad (2.26)$$

The first three terms of equation 2.26 correspond to the terms in the mean-field approximation with independent electrons (equation 2.21), and the last term, called the *exchange-correlation* energy, contains everything left out of the classical description of the electron-electron interaction.

Lastly, to obtain the total energy of the ground state E_{ground}, we use the result that the ground state density n_{ground} minimizes the energy functional of equation 2.26 [21]. Such a minimization leads to a set of equations for the single-electron wavefunctions ϕ_i, from which the density is constructed according to equation 2.18:

$$\left[-\frac{1}{2} \nabla_i^2 + V_{\mathrm{n}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \varepsilon_i \phi_i(\mathbf{r}) \quad (2.27)$$

The set of equations given by equation 2.27 are known as the *Kohn-Sham equations* and form the basis of DFT calculations. If we knew the exact form of the exchange-correlation energy $E_{xc}[n]$, the solutions of 2.27 would give the *exact* ground-state energy and electron density. Unfortunately, the exact mathematical form of this functional is unknown and must be approximated. For the purposes of this thesis, we use the general gradient approximation

(GGA) to $E_{xc}[n]$ of Perdew, Burke, and Ernzerhof, the details of which can be found in reference [23].

In practice, the equations given by 2.27 must be solved self-consistently in an iterative fashion, as the total potential $V_{tot} = V_{\mathrm{n}}(\mathbf{r}) + V_{\mathrm{H}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r})$ depends on the electron density n, which is a function of the unknown ϕ_i. This is achieved by initializing n to some reasonable guess, such as the superposition of the densities of the individual atoms, and then iteratively solving for the ϕ_i until the density given by these solutions and the density used to solve for these solutions are within some error tolerance. Thus, within the framework of DFT at a given level of approximation (by choice of exchange-correlation functional and basis set size), one can systematically improve n to whatever level of accuracy is desirable or feasible with respect to computational cost.

Atomic forces

We now bridge the gap between the electron density and the motion of the atoms by defining the atomic forces in terms of the electron density $n(\mathbf{r})$. Under the Born-Oppenheimer approximation, we separated the motion of the nuclei and electrons, which amounts to separating the total wavefunction $\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, \mathbf{R}_1, ..., \mathbf{R}_M)$ into a wavefunction for the electrons $\psi_{\mathrm{el}}(\mathbf{r}_1, ..., \mathbf{r}_N)$ and a wavefunction for the nuclei $\chi_{\mathrm{nuc}}(\mathbf{R}_1, ..., \mathbf{R}_M)$:

$$\Psi(\mathbf{r}_1, ..., \mathbf{r}_N, \mathbf{R}_1, ..., \mathbf{R}_M) = \psi_{\mathrm{el}}(\mathbf{r}_1, ..., \mathbf{r}_N)\chi_{\mathrm{nuc}}(\mathbf{R}_1, ..., \mathbf{R}_M) \tag{2.28}$$

Without going into details, this separation enables us to write the complete time-independent Schrödinger equation as

$$E_{\mathrm{el}}\psi_{\mathrm{el}}\chi_{\mathrm{nuc}} + \left[-\sum_I \frac{\nabla_I^2}{2M_I} + \frac{1}{2}\sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \right] \psi_{\mathrm{el}}\chi_{\mathrm{nuc}} = E_{\mathrm{tot}}\psi_{\mathrm{el}}\chi_{\mathrm{nuc}}, \tag{2.29}$$

Multiplying both sides of equation 2.29, we obtain a many-body Schrödinger equation for the nuclei in the effective potential of the electrons:

$$\left[-\sum_I \frac{\nabla_I^2}{2M_I} + \frac{1}{2}\sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + E_{\mathrm{el}} \right] \chi_{\mathrm{nuc}} = E_{\mathrm{tot}} \chi_{\mathrm{nuc}}. \tag{2.30}$$

Upon inspection of equation 2.30, we see that the total potential energy function for the nuclei is given by:

$$U(\mathbf{R}_1, ..., \mathbf{R}_M) = \frac{1}{2}\sum_{I \neq J} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} + E_{\mathrm{el}}(\mathbf{R}_1, ..., \mathbf{R}_M), \tag{2.31}$$

where we explicitly write E_{el} as a function of the nuclear coordinates as a reminder that $\mathbf{R}_1, ..., \mathbf{R}_M$ enter into the electronic Schrödinger equation as parameters.

To summarize, in AIMD, we start with an initial set of nuclear coordinates $\mathbf{R}_I$, for which we compute E_{el} from the solution of the Kohn-Sham equations of 2.27. We can then construct the potential energy function U for the nuclei by adding to E_{el} the nuclear-nuclear repulsion, as in equation . Finally, we update the positions of the atoms using the forces from equation 2.3 and solve for E_{el} for these new nuclear coordinates. This process is repeated for some number of timesteps, over which thermal averages are computed according to equation 2.8.

2.1.3 Liquid-Quench Molecular Dynamics

Liquid-quench molecular dynamics (LQMD) is a computational method for generating the atomic structure of amorphous solids by performing subsequent MD (classical or *ab initio*) simulations that mimic experiments. LQMD has been successfully applied to amorphous silicon, amorphous carbon, and various amorphous oxides, to give some examples. Meant to model the general deposition process of amorphous solids, the system is first

melted at an extremely high temperature (well-above the melting temperature) in order to erase the crystal structure and then subsequently quenched (rapidly cooled) to a low temperature. A schematic for LQMD is shown in Figure 2.1. The validity of LQMD comes

Figure 2.1: Workflow of liquid-quench molecular dynamics (LQMD)

from the notion of a *thermal spike*: localized melting produced by the bombardment of ions during deposition, as depicted in Figure 2.2. A single impacting ion temporarily melts tens to hundreds of atoms near the impact site, forming a quasiliquid that rapidly cools to a solid structure. This is analagous to the *quench* step of LQMD. Furthermore, it has been demonstrated that the cooling time for thermal spikes is on the order of subpicoseconds (e.g. 10^{-1} picoseconds) for 100 eV ions, which is a reasonable timescale for classical MD. However, for AIMD, we have to settle for unrealistically fast quench rates. This, coupled with the severe limitations on system size for AIMD, greatly restricts the accuracy of amorphous structures generated by *ab initio* LQMD. In an effort to overcome this limitation while retaining the intricacy of *ab initio* interatomic potentials, we developed a neural network interatomic potential (NNP) using *ab initio* LQMD data to be used in classical LQMD and other types of simulations. This procedure is described in the following section.
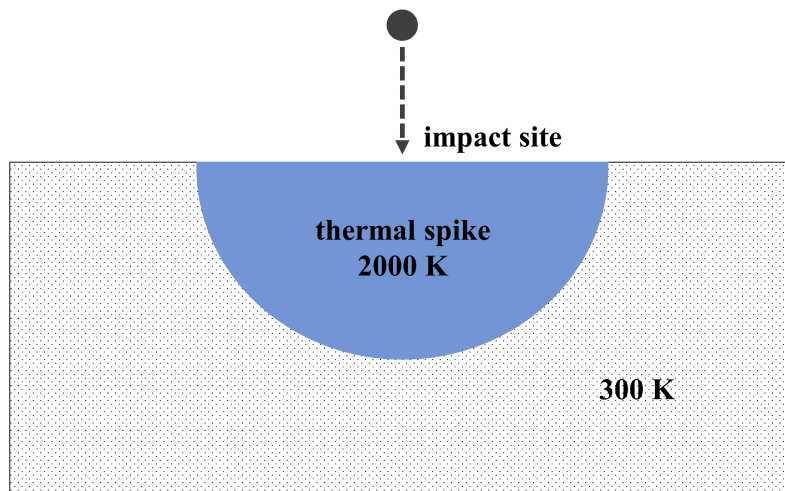
Figure 2.2: Depiction of a thermal spike (blue) due to ion bombardment. The temperature of the melted region is approximately the melting temperature of the material and is much higher than the temperature of the surrounding substrate.

2.2 Neural Network Interatomic Potential

The $\mathcal{O}(N^3)$ scaling with respect to system size N of DFT limits both the quench rate and the simulation cell size of *ab initio* LQMD to 10-100 K/ps and a few hundred atoms, respectively - both of which are unrealistic for accurately modelling amorphous systems. On the other hand, the construction of classical interatomic potentials requires extensive experience in potential development and a deep understanding of the system at hand. Furthermore, these potentials often do no capture the effects of intricate quantum interactions, which are particularly important for modelling amorphous materials and liquids.

A recent development for bridging the gap between the more accurate but less efficient AIMD and the less accurate but more efficient classical MD involves the application of machine learning (ML) to potential development. When trained with AIMD data, an ML potential can reproduce the *ab initio* potential energy surface (PES) at a much lower

computational cost. Another advantage of ML potentials is they do not have any prefixed functional form, giving them a flexibility to fit a variety of PES shapes. The two most common ML potentials are the Gaussian approximation potential (GAP) [24, 25, 26] and the neural network potential (NNP) [27, 28, 29]. While both exhibit the aforementioned benefits of ML potentials, GAP is more computationally demanding and is less widely applied than NNP. NNPs have been applied to a wide variety of phenomena, including surface reactions [30, 31], diffusion in the amorphous phase [32], phase changes [33], amorphous alloys [34], and various multi-component systems [35, 36]. In this thesis, we developed NNPs using a high-dimensional neural network [29] with atom-centered symmetry functions [37] as descriptors, as applied in the SIMPLE-NN Python module [38].

2.2.1   Theory

The fundamental idea behind NNPs is that the total energy can be written as the sum of the individual atomic energies. Early NNPs were applied to molecular adsorption on surfaces and thus used input features such as distance from the surface and angle with respect to the surface normal and the lateral coordinate [39]. However, using these features more generally is infeasible, as the number of degrees of freedom increases rapidly with increasing system size. Hobday et al. extended this idea but used the parameters related to the local bonding environment as the input to the network and fit to a many-body potential [40].

The general form of a neural network is shown in Figure 2.4. The network consists of an input layer, multiple hidden layers, and an output layer. The input layer corresponds
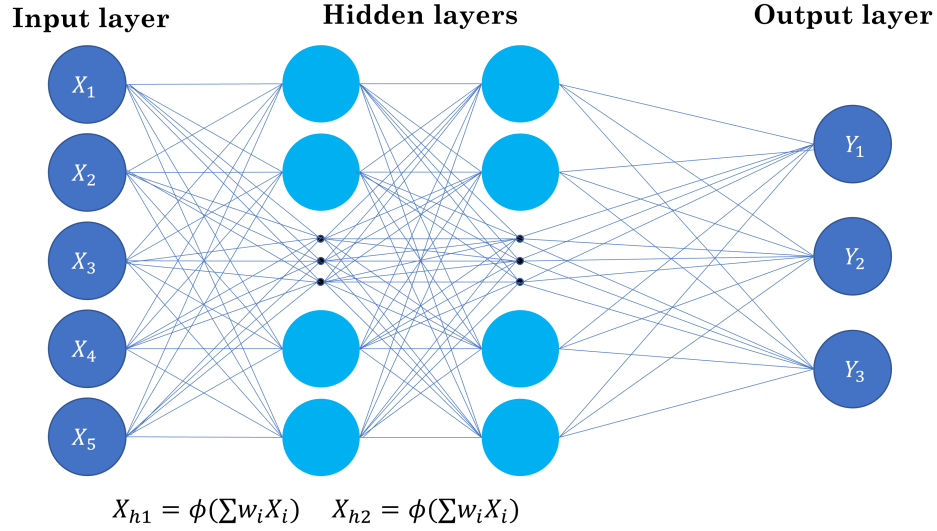
Figure 2.3: Artificial neural network.

to the training data set from AIMD and the output layer corresponds to the potential energy (one output node) in our case. Each layer is composed of nodes (circles), which are connected by synapses (lines). A synapse carries the output value from one node to the input of a node in the next layer. The input X_j^{L+1} to a given node is a weighted sum of all the nodes in the previous layer:

$$X_j^{L+1} = \sum_i w_i X_i^L + b^L. \tag{2.32}$$

Here, L denotes the layer number, thus X_j^{L+1} is the jth node in the $L+1$th layer, w_i is the weight for the ith node, b_L is the bias for the Lth layer, and the sum is over all the nodes in the previous layer L. At each node, X_j^{L+1} is transformed by some differentiable non-linear activation function ϕ. In this work, we use the sigmoid function, given by

$$\phi(X_j^{L+1}) = \frac{1}{1 + \exp(X_j^{L+1})}. \tag{2.33}$$

Thus the output of the jth node in the $L + 1$th layer is

$$X_j^{L+1,\text{out}} = \phi(\sum_i w_i X_i^L + b^L). \tag{2.34}$$

$X_j^{L+1,\text{out}}$ is then input to the next layer as one of the X_i in equation 2.32. The training of a neural network refers to the optimization of the weights w_i in this equation.

Atom-Centered Symmetry Functions

An NNP is essentially a regression of the relationship between atomic configurations (nuclear coordinates) and the total energy. Because the total energy is invariant under rotation and translation of the system, as well as permutations of atoms of the same element, we first must transform the atomic coordinates before inputting into the neural network. In the formulation used here, we use the symmetry function descriptors developed by Behler and Parrinello with one radial and two angular components [37]:

$$G_i^{\text{radial}} = \sum_j e^{\eta(R_s - R_{ij})^2} \cdot f_c(R_{ij}), \tag{2.35}$$

$$G_i^{\text{angular,1}} = 2^{1-\xi} \sum_{j,k \neq j} (1 + \lambda \cos\theta_{ijk})^\xi \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}) \cdot f_c(R_{jk}), \tag{2.36}$$

$$G_i^{\text{angular,2}} = 2^{1-\xi} \sum_{j,k \neq j} (1 + \lambda \cos\theta_{ijk})^\xi \cdot e^{-\eta(R_{ij}^2 + R_{ik}^2)} \cdot f_c(R_{ij}) \cdot f_c(R_{ik}), \tag{2.37}$$

where i is the index of the central atom, j and k the indices of neighboring atoms, R_{ij} the distance between atom i and atom j, and θ_{ijk} the angle formed by atoms i, j, and k. The parameters η and R_s specify the width and center of the Gaussian functions, respectively, and ξ and λ change the shape of the angular symmetry functions. The function f_c is a

cutoff function with radius R_c:

$$f_c(R_{ij}) = \begin{cases} \frac{1}{2}\cos(\pi\frac{R_{ij}}{R_c}) + \frac{1}{2} & \text{if } R_{ij} \leq R_c \\ \\ 0 & \text{if } R_{ij} > R_c \end{cases}. \tag{2.38}$$

This function restricts the symmetry functions 2.35, 2.36, and 2.37 to a radius R_c around

each atom. A set of symmetry functions with different values of $(\eta, R_s, \xi, \lambda)$ constitutes a

vector $\mathbf{G}_i(\mathbf{R}_i)$ that defines the local environment around the ith atom. The set $\{\mathbf{G}_i\}$ of

such vectors for all atoms in the system constitutes the input of the neural network.

The quality and generalizability of the network are governed by its *complexity*

(number of hidden nodes and synapses between layers) and by the quality and size of the

training data set. Practically, both of these are limited by computational resources. We

discuss both of these issues in the following section.

Atomic Neural Network

The structure of the neural network consists of an atomic neural network (ANN)

for each chemical species in the system, where the same ANN is used for every atom of the

same chemical species. As shown in Figure, each ANN takes in the descriptor vector $\mathbf{G}_i$ of

its corresponding atom i and outputs the atomic energy.

The target loss function Γ to be minimized is given by:

$$\Gamma = \frac{1}{K}\sum_{i=1}^{K}\left(\frac{E_i^{\text{DFT}} - E_i^{\text{NNP}}}{M_i}\right)^2 + \frac{\mu}{3\sum_i^K N_i}\sum_{i=1}^{K}\sum_{j=1}^{N_i}|\mathbf{F}_{ij}^{\text{DFT}} - \mathbf{F}_{ij}^{\text{NNP}}|^2$$

$$= \text{RMSE(energy)}^2 + \frac{\mu}{3}\text{RMSE(force)}^2 \tag{2.39}$$

Here, K is the total number of structures in the training set, N_i is the number of atoms

in the ith structure, and $E_i^{\text{DFT(NNP)}}$, $F_{ij}^{\text{DFT(NNP)}}$ is the total energy, atomic force of the
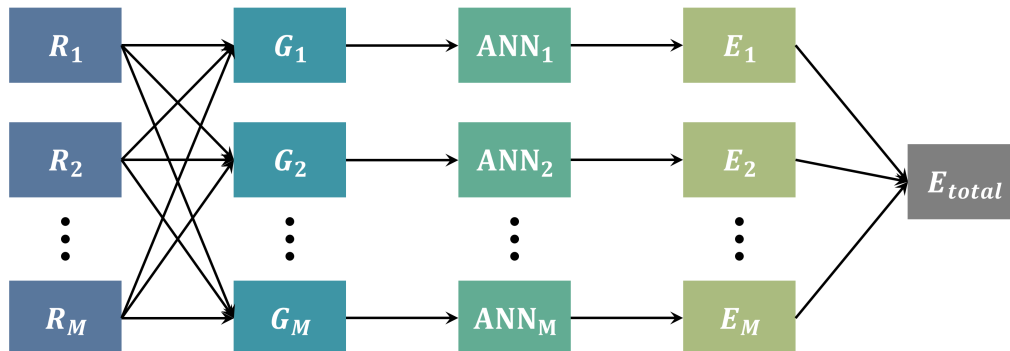
Figure 2.4: Atomic neural network

jth atom from DFT (NNP), respectively. The scaling parameter μ controls the relative importance between the energy error and the force error.

2.2.2 NNP Development

There are three main stages of NNP development: i) database construction, ii) NNP training, and iii) NNP validation.

Database construction

In order for the NNP to properly fit the *ab initio* PES, we must include a sufficient amount of relevant and distinct atomic configurations. Depending on the problem at hand, this can mean including configurations of different phases (e.g. liquid, crystalline solid, amorphous solid, etc.), of different geometries (e.g. bulk, slab, cluster, structures with defects), and/or of different chemistries (e.g. varying concentrations). To ensure proper sampling of the thermodynamics of each configuration type, an AIMD simulation is performed for each configuration. For example, a reasonable database for carbon could

consist of AIMD simulations for each of the following: bulk diamond, various diamond slabs, graphite, graphene, and amorphous carbon.

The inputs (or features) for the NNP are the *xyz*-coordinates of configurations sampled from the various AIMD runs, and the output of the NNP is the AIMD potential energy for each sample. Therefore, it is important that the AIMD simulations used as the training set are sufficiently accurate. In this work, all of the AIMD training data was produced using the projector-augmented wave (PAW) pseudopotentials as implemented in the Vienna Ab initio Simulation Package (VASP) [41, 42]. Specific details about each simulation will be given in their corresponding chapters.

NNP validation

There are two ways by which one should validate a developed NNP. First, we need to confirm that our NNP accurately describes the training data, i.e. it reproduces the *ab initio* PES for configurations in the training data. This is achieved by running the same simulations used in the training data with the NNP, and then comparing the potential energies given by the NNP to DFT. Secondly, and more importantly, we need to check that the NNP has not *overfit* the training set and can reasonably describe the PES of new configurations not in the training set. This is achieved by running longer simulations or simulations with larger simulation boxes than included in the training data.

For comparing the NNP PES to the *ab initio* PES, we choose the root mean square error (RMSE) as our measure of accuracy:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(E_i^{NNP} - E_i^{DFT})^2}{n}}, \tag{2.40}$$

where i denotes a sample configuration and E_i^{NNP}, E_i^{DFT} the NNP, DFT potential energy

for sample i. Additionally, to confirm the stability of the NNP outside of the training data, we monitor how the RMSE changes with simulation time. If necessary, the NNP is re-trained adaptively by including configurations from the NNP simulation for which the RMSE is large.

Chapter 3

Analysis

In this chapter, we give details about the two metrics, the pair distribution function and the polyhedra network, used to measure and describe order in the various phases of the metal oxides and aqueous electrolytes studied in this thesis. Each of these analyses was coded in Python, the codes of which are presented in the Appendix.

3.1 Pair Distribution Function

The first tool we will use to characterize our amorphous structures is the pair distribution function $g(r)$. The pair distribution function gives the distribution of distances between pairs of atoms in a given system and provides some insight into the extent of short- and medium-range order of a material. For example, crystalline materials tend to have narrower distributions of distances between bonded atom types, i.e. all bond distances between two atom types are very close to some average bond length. This order arises in the pair distribution function as sharp peaks, with each peak corresponding to distinct

coordination shells around a given atom type. In less ordered materials, such as amorphous materials, the distribution of distances broadens, i.e. there is more variety in the bond distances between two atoms types.

The pair distribution function $g(r)$ is defined as the radial number density of atoms, $\rho(r)$, normalized by the average density ρ_0:

$$g(r) = \frac{\rho(r)}{\rho_0} \tag{3.1}$$

with the radial density given by

$$\rho(r) = \frac{2}{N} \sum_{i=1}^{N} \sum_{j=i+1}^{N} f(r, r_{ij}, \sigma) \tag{3.2}$$

and the average density by

$$\rho_0 = \frac{N}{V}. \tag{3.3}$$

In equations 3.2 and 3.3, N is the number of atoms in the system and V is the volume of the simulation box. The sum for $\rho(r)$ is performed over all atom pairs in the system, and r_{ij} is the shortest distance between the ith and jth atoms, taking into account periodic boundary conditions. The function $f(r, r_{ij}, \sigma)$ in equation 3.2 is a Gaussian of width σ that is used to smear each atom's contribution to $\rho(r)$:

$$f(r, r_{ij}, \sigma) = \frac{1}{4\pi^{3/2} r_{ij}^2 \sigma} e^{-\left(\frac{r - r_{ij}}{\sigma}\right)^2} \tag{3.4}$$

In ordered structures, we observe distinct peaks in $g(r)$, which can be used to determine bond lengths, coordination numbers, and the degree of order. Bond lengths are given by the r-value at which the maximum of the peaks occurs, and coordination numbers can be derived from the area under the peak:

$$N_{\alpha-\beta} = \int_0^{r_{min}} 4\pi r^2 \rho_0 g(r) dr, \tag{3.5}$$

where α and β denote the two atom types for which the coordination number is being computed, r_{min} is the r-value of the first minimum (valley) in g, and ρ_0 is the average density given by equation 3.3. Lastly, the degree of order is given by the sharpness of the peaks, where peaks occurring at distances of one to two bond lengths depict short-range order (typically up to $4-6$ Å) and distances of three to four bond lengths depict medium-range order (typically up to $8-10$ Å).

3.2 Polyhedra network

As previously mentioned, the presence of structural vacancies in the first shell of In leads to two different types of linking in the polyhedra network of c-In$_2$O$_3$, edge-sharing (2 shared O atoms) and corner-sharing (1 shared O atom). There is another possible type of linking, face-sharing (3 shared O atoms), that is not present in c-In$_2$O$_3$, as it has the shortest In-In distance and thus tends to be less stable than the other types of linking. While only edge-sharing and corner-sharing polyhedra have been identified previously in the literature on a-In$_2$O$_3$, we expand our analysis to include face-sharing polyhedra and find the presence of a few such linkings in our a-In$_2$O$_3$ structure. We also investigate the connection between our discovery of face-sharing polyhedra and the bound oxygen pairs discussed in the previous section.

To quantify the polyhedra network, we compute $i)$ the number of each type of polyhedra and $ii)$ the distribution of In-In distances for each type of polyhedra in c-In$_2$O$_3$

and a-In_2O_3. To calculate the number of each type of polyhedra around a given In atom, we use the following algorithm:

1. **Identify all _O_ atoms in the first shell of central In atom.** This amounts to finding all O atoms within some cutoff distance, taken to be the r-value of the first minimum in the In-O $g(r)$ (see Figure).

Algorithm 1 First Shell

---
for $i = 1...N_O$ **do**   ▷ Iterate over all O atoms

    if $dist(In_{central}, O_i) < r^O_{cut}$ **then**   ▷ Check if O atom is within cutoff.

        first_shell.append(O_i)   ▷ Add O atom to first shell of $In_{central}$.

---

2. **Identify all _In_ atoms in the second and third shells of central In atom.** This amounts to finding all In atoms within some cutoff distance, taken to be the r-value of the second minimum in the In-In $g(r)$ (see Figure).

Algorithm 2 Second & Third Shells

---
for $i = 1...N_{In}$ **do**   ▷ Iterate over all In atoms

    if $In_i \neq In_{central}$ **then**   ▷ Do not compute distance between $In_{central}$ and itself.

        if $dist(In_{central}, In_i) < r^{In}_{cut}$ **then**   ▷ Check if In atom is within cutoff.

            outer_shells.append(O_i)   ▷ Add In atom to outer shells of $In_{central}$.

---

3. **Identify how many O atoms are shared between the central In atom and every In atom in its second and third shells.** This is done by iterating through the In atoms identified in 2 and finding the intersection of the first shell O atoms of the central In atom and the first shell O atoms of every other In atom.

4. **If the size of the intersection is 1, meaning there is 1 shared O atom**

between two given In atoms, the count for corner-sharing polyhedra is increased by 1. Likewise, if the size of the intersection is 2 (3), the count for edge-sharing (face-sharing) polyhedra is increased by 1.

Algorithm 3 Polyhedra
---

    for $i = 1...N_{\text{outer}}$ **do**   ▷ Iterate over all In atoms in 2nd and 3rd shells.

        shared_O = first_shell($\text{In}_{\text{central}}$).intersect(first_shell($\text{In}_i$)) ▷ Find shared O atoms

        if shared_O.size == 1 **then**   ▷ Determine polyhedra type and increment.

            corner += 1

        else if shared_O.size == 2 **then**

            edge += 1

        else if shared_O.size > 2 **then**

            face += 1

In crystalline materials, there is typically a narrow range for the number of each type of polyhedra around a given central atom. However, as the order in the material decreases, this range widens, giving a distribution of each polyhedra type around a central atom. Thus, we can measure disorder in terms of the distortion of this network, i.e. the variation of the number of corner-, edge-, and face-sharing polyhedra around the values corresponding to the crystal structure.

Chapter 4

Amorphous Metal Oxides

Metal oxides (MOs) are the most abundant materials in the Earth's crust and exhibit many desirable electronic and mechanical properties, making them important materials for future electronics in terms of sustainability and performance. In comparison to crystalline silicon and other III-V semiconductors, MOs exhibit excellent carrier mobilities (even in the amorphous state), mechanical stress tolerance, compatibility with diverse substrates, room-temperature processing, and high optical transparency. Transition metal oxides (TMOs) are particularly interesting, displaying a broad range of structures and properties owing to the various types of metal-oxygen bonding: Due to the unique nature of d-electrons in the valence shell, bonding in TMOs can range from ionic to highly covalent or metallic. For example, RuO_2 and ReO_3 display metallic properties, while $BaTiO_3$ is highly insulating, and others have properties that are highly tunable via temperature, pressure, and composition. While not discussed in this thesis, the exciting discovery of high-temperature superconductivity in cuprates underscores the significance of TMOs.

This chapter focuses exclusively on two binary MOs: Indium (III) oxide (In_2O_3)

and vanadium (V) oxide (V_2O_5). For both systems, we first present our results on structures generated from *ab initio* liquid-quench molecular dynamics (LQMD). Ab initio simulations are ideal for accurately capturing the structural complexities of amorphous systems, but are limited in system size and timescale because of the computational cost associated with quantum-mechanical simulations. These limitations are particularly restrictive for generating amorphous structures from LQMD because i) theoretical amorphous systems have infinitely large unit cells and the sizes feasible for DFT (up to a few hundred atoms) are not representative of a real amorphous structure, thus severely constraining the system, and ii) the quench times achievable in DFT are unrealistically fast (10^{14} K/s). To help bridge the gap between the quantum-mechanical accuracy of DFT and the efficiency of classical MD, we used our *ab initio* MD (AIMD) to train a neural network interatomic potential (NNP) that was subsequently used in classical LQMD simulations of larger system sizes and timescales. We then utilize the metrics defined in Chapter 2 to compare the structures obtained from *ab initio* LQMD to those obtained from classical MD. We then computed various material properties, including glass transition temperatures and electronic density of states, to elucidate structure-property relationships.

4.1   Computational Details

All *ab initio* molecular dynamics (AIMD) and density functional theory (DFT) calculations were performed using the projector-augmented wave (PAW) pseudopotentials as implemented in the Vienna Ab initio Simulation Package (VASP) [41, 42] with the PBE exchange-correlation functional [23]. For the AIMD calculations, we utilized an energy

cutoff of 400 eV and a timestep of 0.5 fs, and k-point sampling was restricted to the Γ point only. For static DFT calculations, we increased the energy cutoff to 520 eV and used a Monkhorst-Pack grid of 4x4x4 k-points. The conventional 80 atom unit cell for In_2O_3 and a 56 atom supercell of V_2O_5 was used in all AIMD and DFT calculations.

The liquid-quench AIMD simulations were performed as follows. In order to properly melt the structure, we first added 0.5 Å of vacuum spacing in every direction to the simulation box and peformed a 10 ps NVT simulation at 3000 K for In_2O_3 ($T_{melt} = 2183$ K) and at 2000 K for V_2O_5 ($T_{melt} = 963$ K). We subsequently performed a quench to 100 K in the NPT ensemble to allow the system's density to adjust from the constrained density of the melt. We quenched the system at a cooling rate of 200 K/ps for a total simulation time of 14.5 ps. Finally, we equilibrated the system for 6 ps in the NVT ensemble. The final structure of this equilibration run was used in subsequent static DFT calculations of the electronic properties of a-In_2O_3 and a-V_2O_5.

All classical MD simulations were performed using the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [43]. For In_2O_3, a neural network potential trained with the Python module SIMPLE-NN [38] as described in the methods section was utilized, and for V_2O_5, a ReaxFF potential [44] designed to model bond breaking and chemical reactions was used. The liquid-quench MD simulations were performed as follows.

4.2 Indium oxide

Indium (III) oxide (In_2O_3) is a wide-bandgap (3.6-3.75 eV) semiconductor [45, 46] with a single-crystal mobility of 160 $cm^2 \cdot V^{-1} \cdot s^{-1}$ [47] and high (90%) transparency in

the visible region [48]. It forms the basis of most transparent conducting oxides (TCOs) and enabled the first fully transparent thin-film transistor (TFT). Initial applications of TCOs primarily employed the crystalline phase of the semiconductor; however, recently there has been an increasing shift toward use of the amorphous phase. This is primarily due to a simpler deposition process and the ability to scale to larger dimensions than their crystalline counterparts. In particular, amorphous materials can be deposited at lower temperatures, expanding the types of substrates that can be used, and are isotropic and lack grain boundaries, allowing them to be uniformly deposited over large areas. Consequently, 30-40% of flat panel displays now utilize an amorphous TCO. Amorphous materials also often tend to be more pliable, making them desirable for use in flexible electronics.

Astonishingly, these advantages can be realized *without loss to conductivity or transparency* when optimized. This is attributable to the nature of the conduction bands of indium. Due to the large spherically symmetric orbitals characterizing the conduction bands, there is a high degree of wave function overlap, electron delocalization, and electron mobility, regardless of the microstructure. Because the structure and properties of both crystalline indium oxide (c-In$_2$O$_3$) and amorphous indium oxide (a-In$_2$O$_3$) are well-studied in the literature, experimentally and theoretically, we first present In$_2$O$_3$ as a prototype for validating our computational framework for producing representative structures for amorphous MOs.

4.2.1 *Ab initio* LQMD

In$_2$O$_3$ can exist in two phases: a cubic (bixbyite) structure and a rhombohedral (corundum) structure. Most In$_2$O$_3$-based TCOs exhibit the cubic structure, as the rhom-

44

bohedral phase is produced at high temperatures and pressures. Because of this, here we only concern ourselves with the bixbyite phase (space group Ia3, number 206). The bixbyite structure (shown in Figure 4.1) is a fluorite-type structure with one-quarter of the anions missing, where the oxygen atoms are octahedrally coordinated around indium with indium cations surrounded by six oxygen atoms (c-$N_{\text{In-O}} = 6$) and two structural vacancies. As shown in Figure 4.2, these six oxygen atoms form an octahedron around each indium atom, which forms the building block of the c-In_2O_3 structure. This is the first coordination shell around indium.
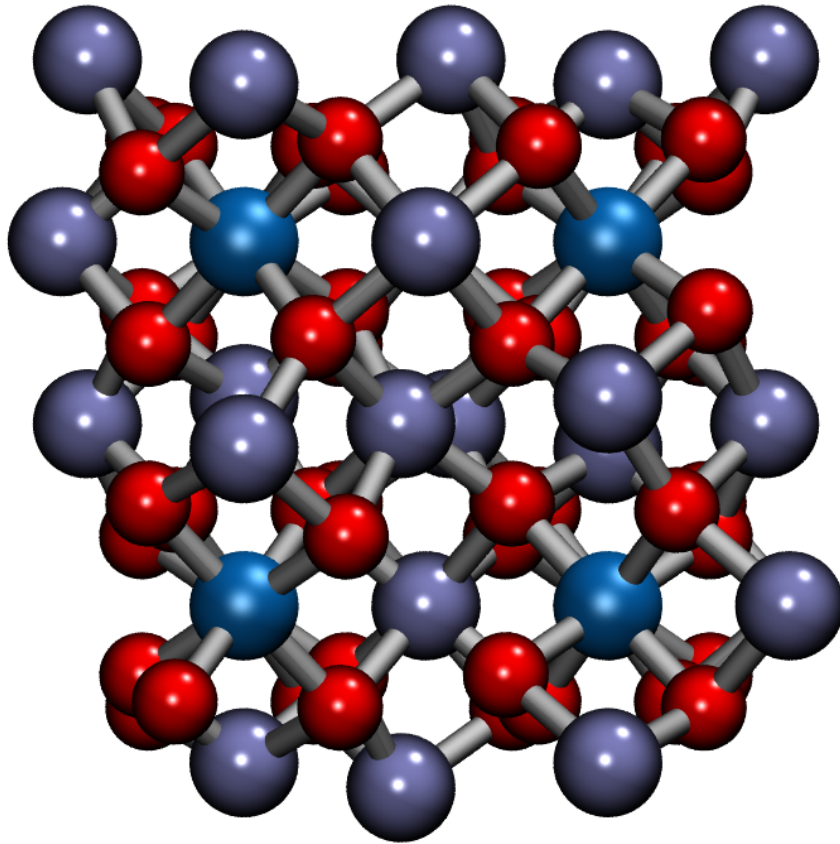

Figure 4.1: Bixbyite structure of In2O3. d-site In atoms are shown in lavender, b-site In atoms in blue, and O atoms in red.

There are two different configurations for the two structural vacancies around indium: b-sites, where the structural vacancies are along the body diagonal, and d-sites, where the structural vacancies are along a face diagonal. In the ideal crystal structure, 25% of indium atoms are in b-sites and 75% are in d-sites. Because of the presence of structural vacancies, the InO_6 octahedra can link together in two ways: either by sharing one oxygen atom and one structural vacancy, thus being joined by corners of the two octahedra, or by sharing two oxygen atoms, thus being joined by edges of the two octahedra. Each indium atom participates in 6 edge-sharing polyhedra, with an average In-In distance of 3.34 Å, and in 6 corner-sharing polyhedra, with an average In-In distance of 3.83 Å(see Figure 4.3). These form the second and third shells of indium, respectively.
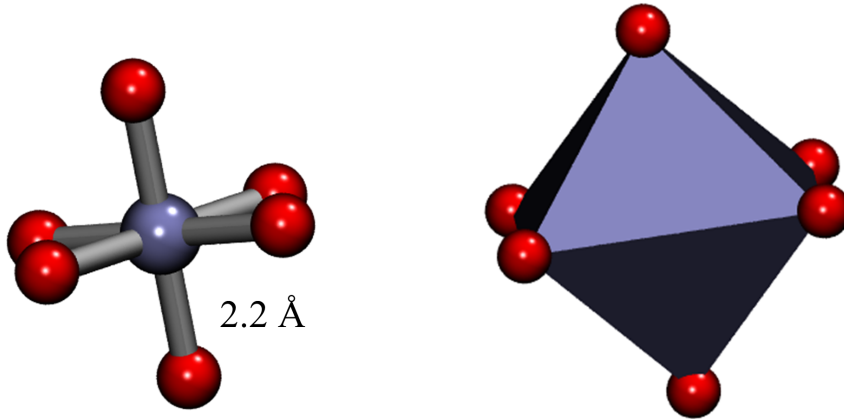

2.2 Å

Figure 4.2: Building block of c-In_2O_3.

In Figure 4.4, we plot total energy versus time for each step of the *ab initio* LQMD process, showing the energy for an MD simulation of c-In_2O_3 for reference at the top left. We see that there is a sharp increase in energy during the melting process of c-In_2O_3, which
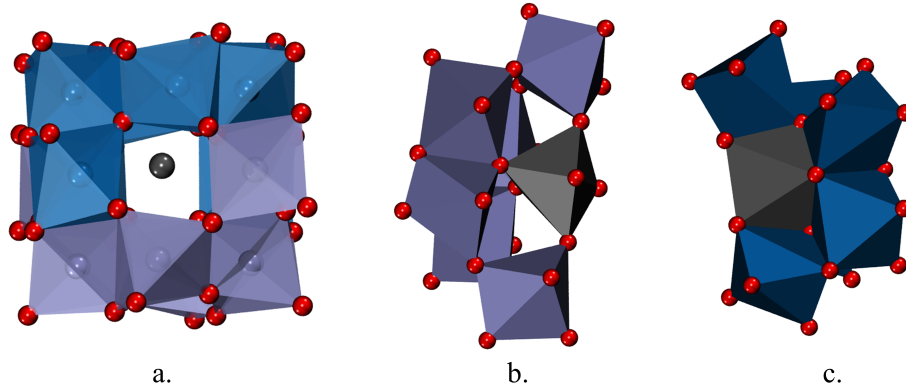
Figure 4.3: Polyhedra network of c-In$_2$O$_3$

is very likely due to the instantaneous increase in temperature and initial expansion, as we added 0.5 Åof spacing in every direction of the simulation box to allow for proper melting. During the quenching simulation, we observe a steady decrease in energy with time, as expected for a stepwise decrease in temperature over time. Lastly, we see that a-In$_2$O$_3$ equilibrates quickly after the quenching.

We next examine the extent of short- and medium-range order in each phase of In$_2$O$_3$ via the pair distribution function. In Figure 4.6, we show the total $g(r)$ for c-In$_2$O$_3$, $melt$-In$_2$O$_3$, and a-In$_2$O$_3$. In c-In$_2$O$_3$, we observe distinct, sharp peaks up to 7 Å, indicative of the short- and medium-range order of a crystal. Upon melting, we observe a complete loss of medium-range order and some retention of short-range order, demonstrated by a smoothing of all peaks above 3 Åand by a broadening of the first peak. Lastly, upon quenching and equilibration, $g(r)$ for a-In$_2$O$_3$ re-gains short-range order and medium-range order to a lesser extent, as exhibited by an increase in sharpness of the first peak relative to the melt and by a noticeable presence of the second and third peaks, as well as a small
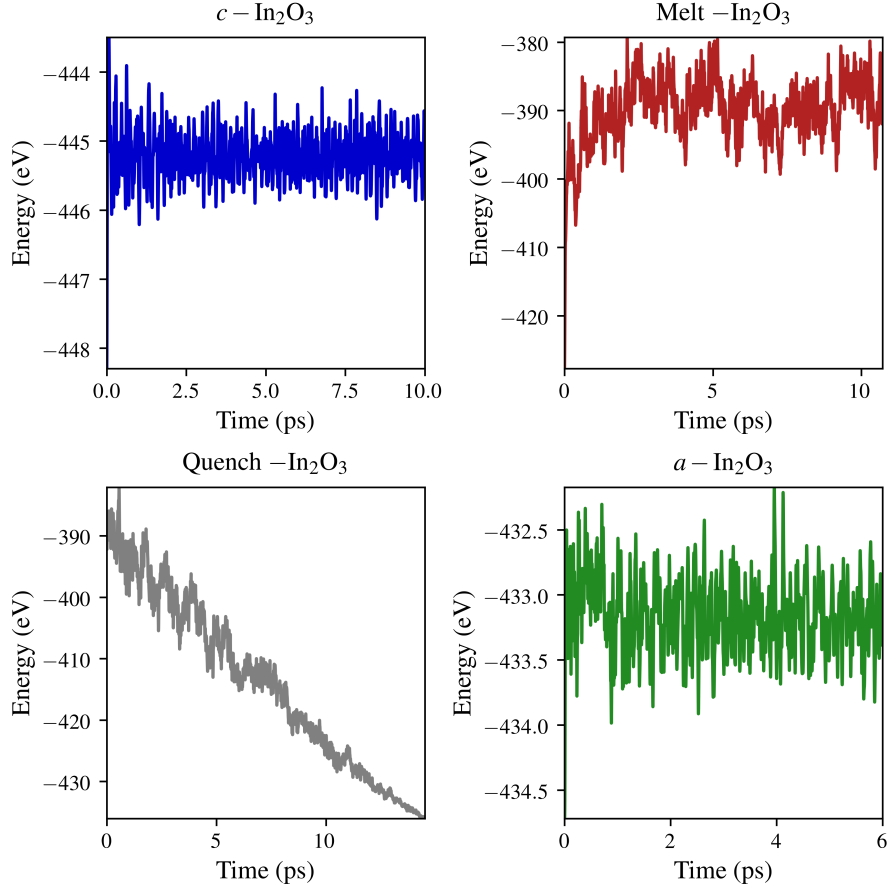
Figure 4.4: Total energy for each simulation of LQMD.

distinguishable peak above 4 Å.

Next, we examine more closely the $g(r)$ for each interaction in each phase of In_2O_3. In Figure 4.6, we show $g(r)$ for each pair of elements (In-O, O-O, In-In) in c-In_2O_3. The first peak corresponds to the first shell of O around In, with the peak occurring at an r-value of 2.21 Å, the average In-O bond length in c-In_2O_3. We also point out the first two peaks in the In-In $g(r)$, which correspond to the second and third shells of In, at r-values of 3.41 Åand 3.91 Å, respectively. Next, we plot $g(r)$ for a-In_2O_3 in Figure 4.7. We first notice the
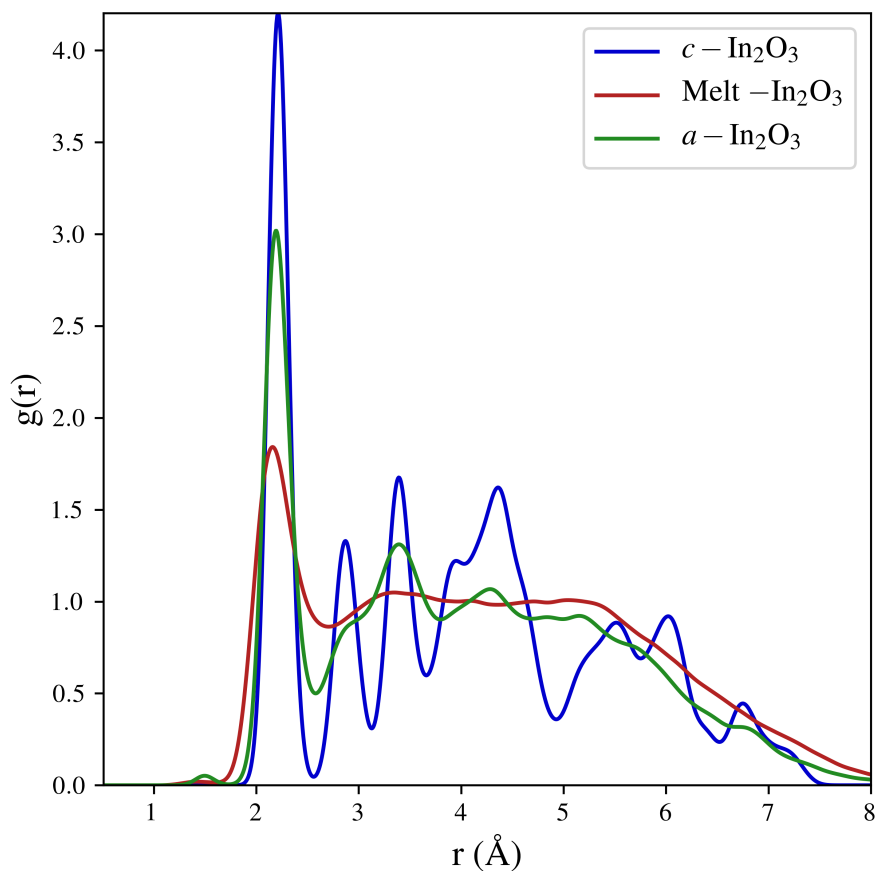
Figure 4.5: Pair distribution functions $g(r)$ for c-In$_2$O$_3$, $melt$-In$_2$O$_3$, and a-In$_2$O$_3$

retention of the first In-O peak and a broadening of all peaks above 3 Å, which is indicative of a short-range order (up to the first shell of In) and of a loss of medium-range order. We also notice that we no longer observe two In-In peaks in the 3-4 Årange, making it difficult to distinguish between a second and third shell of indium. Interestingly, we also observe a new, small peak in the O-O $g(r)$ at 1.5 Å. To investigate the origin of this peak, we next examine the $g(r)$ for $melt$-In$_2$O$_3$ in Figure 4.8. As expected for a liquid, we observe a larger broadening of all peaks, as well as a reduction in peak height, even for the first In-O peak.
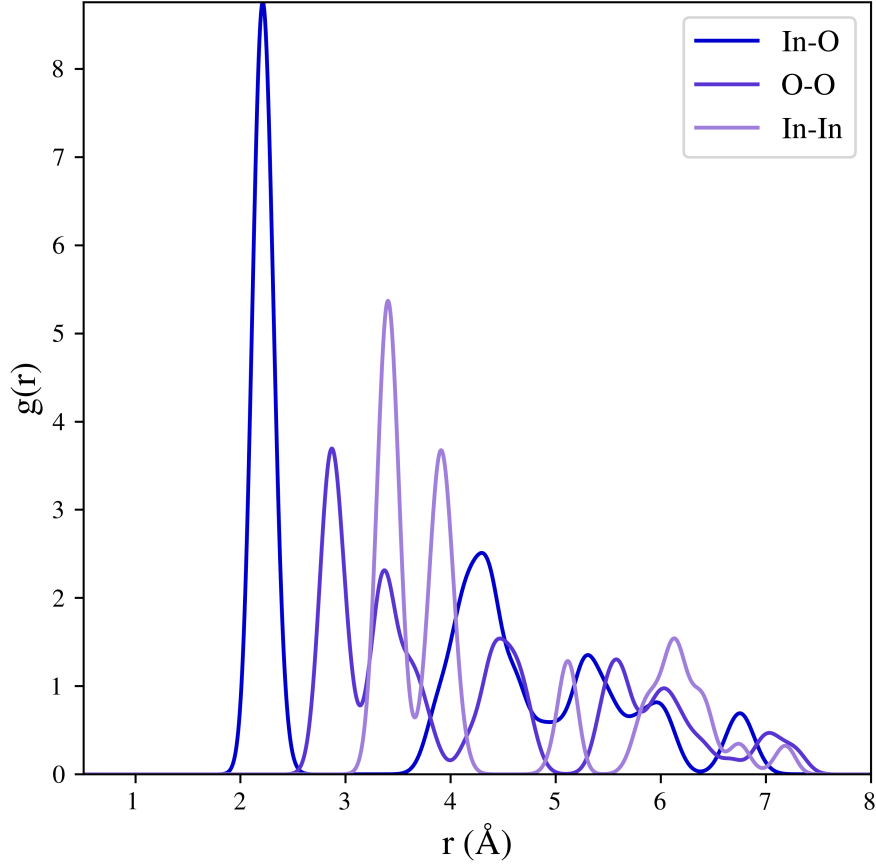
Figure 4.6: Pair distribution functions $g(r)$ for c-In$_2$O$_3$

We also again observe a small peak in the O-O $g(r)$ that is not present in c-In$_2$O$_3$ and, to the best of our knowledge, has not been observed in the literature. However, recently, such a peak was found in simulations of amorphous alumina [49] and was attributed to bound O-O pairs, which were found to serve as hole traps. As such bound O-O pairs have also been observed in amorphous silica, and In, Al, and Si are close to each other in the periodic table, it is very likely that such pairs exist in many amorphous metal oxides that contain group 5-6 metals. Nonetheless, in the case of amorphous alumina, *ab initio* LQMD did not produce such an artifact in the O-O $g(r)$, while another simulation method, stochastic quenching,
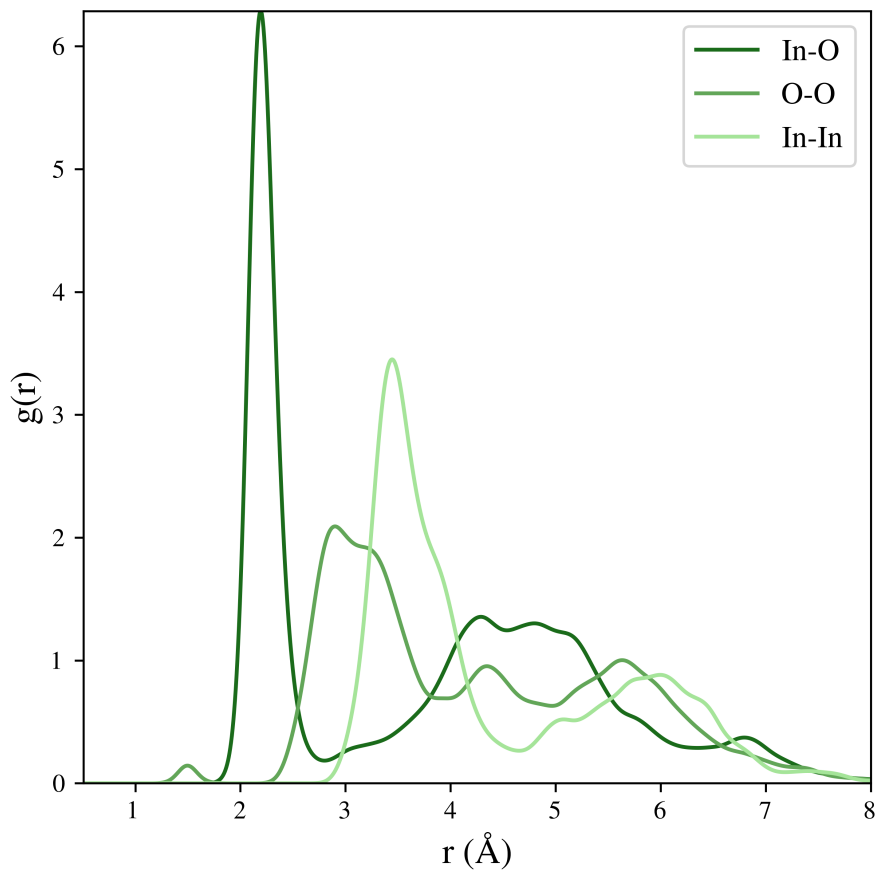
Figure 4.7: Pair distribution functions $g(r)$ for a-In$_2$O$_3$

was able to. The authors claim that LQMD is incapable of capturing such subtleties, while stochastic quenching is, as it should produce maximally amorphous structures, which is not guaranteed in LQMD.

To uncover why our *ab initio* LQMD simulations were able to produce such bound oxygens in a-In$_2$O$_3$ while other such simulation studies have not, we plot the O-O $g(r)$ for *melt*-In$_2$O$_3$ in Figure 4.9 for three simulation conditions - NVT with vacuum spacing, NVT without vacuum spacing, and NPT - as the bound O-O peak first occurs in the melting
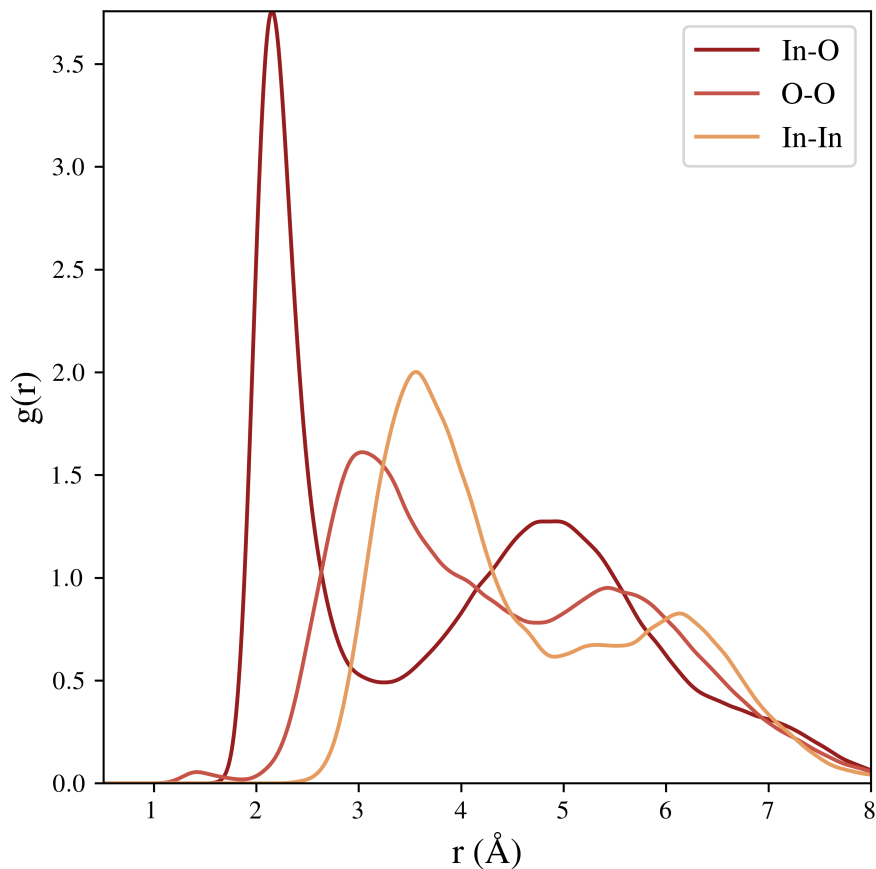
Figure 4.8: Pair distribution functions $g(r)$ for $melt$-In$_2$O$_3$

simulation. From the figure, we see that the bound O-O peak is sharpest in the NVT with vacuum spacing simulation, which is the simulation we used for subsequent quenching and equilibration to produce a-In$_2$O$_3$. We note that in the literature, most ab $initio$ LQMD simulations are done in the NVT ensemble with the density of the crystal. Not allowing for thermal expansion during the melting process seems to prohibit diffusion so that such bound O-O pairs cannot form.

In addition to general trends, we can also obtain coordination numbers and bond
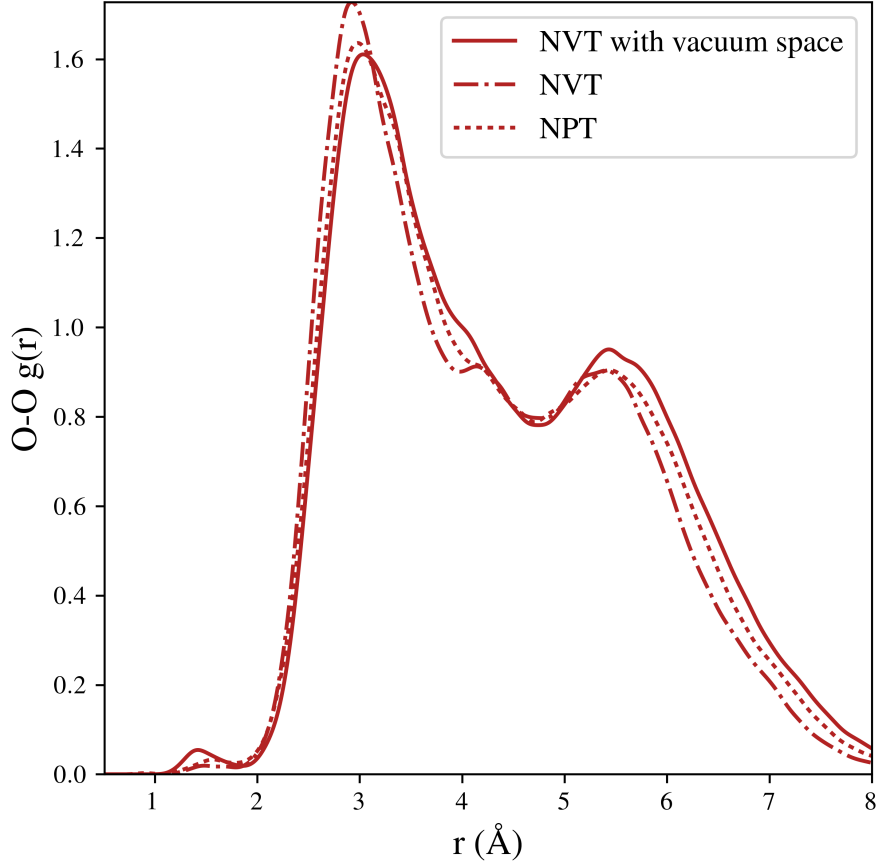
Figure 4.9: O-O $g(r)$ for *melt*-In$_2$O$_3$

lengths from each of the pair distribution functions discussed. In Table 4.1, we provide the coordination numbers and bond lengths of c-In$_2$O$_3$, *melt*-In$_2$O$_3$, and a-In$_2$O$_3$. This will be investigated in future research. As expected, upon melting, the In-O coordination number $N_{\text{In-O}}$ and bond length $R_{\text{In-O}}$ decreases from 6 to 5 and from 2.2 to 2.15, respectively. However, in a-In$_2$O$_3$, both of these values increase to values close to those of c-In$_2$O$_3$, while we expected these values to be lower than the crystalline values, which has been observed in the literature. We suspect that the presence of the bound O$_2$ in our simulations, which

	$N_{\text{In-O}}$	$R_{\text{In-O}}$ (Å)	$R_{\text{In-In}}$ (Å)	$R_{\text{In-In}^*}$ (Å)
c-In$_2$O$_3$	6.0	2.21	3.41	3.91
$melt$-In$_2$O$_3$	5.0	2.15	3.56	NA
a-In$_2$O$_3$	5.8	2.2	3.44	NA

Table 4.1: Coordination number and bond lengths in c-In$_2$O$_3$, $melt$-In$_2$O$_3$, and a-In$_2$O$_3$.

has not been found in the literature for a-In$_2$O$_3$, could play a role in our observation of higher coordination numbers and bond lengths in a-In$_2$O$_3$.
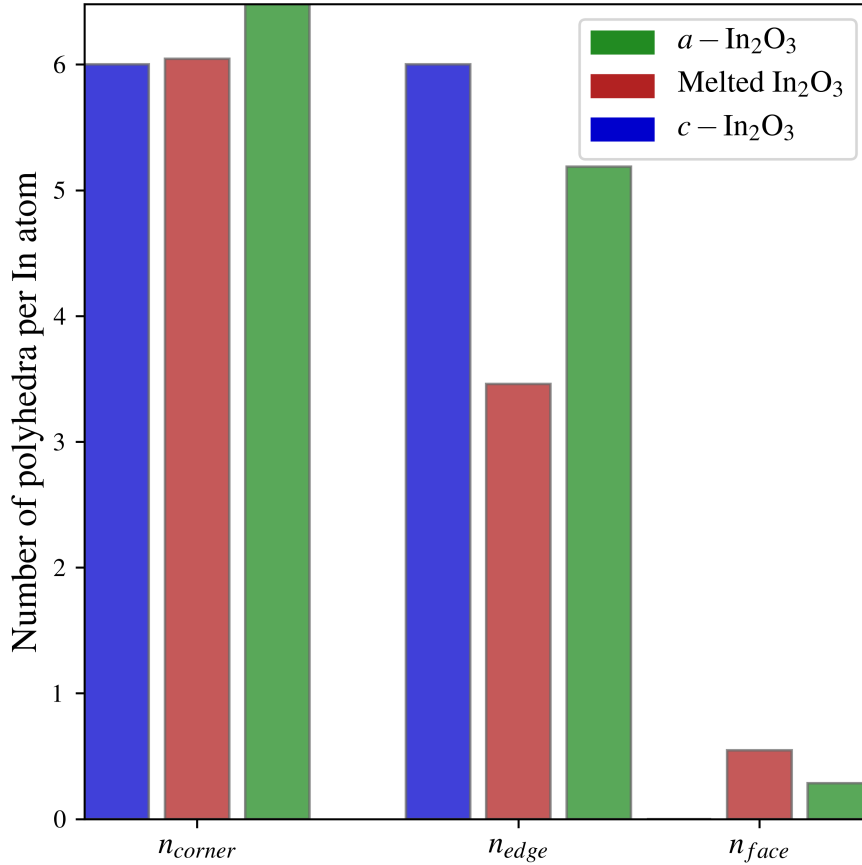

Figure 4.10: Average number of each type of polyhedra (n_{corner}, n_{edge}, and n_{face}) in c-In$_2$O$_3$, $melt$-In$_2$O$_3$, and a-In$_2$O$_3$.

We next examine the distortion of the polyhedra network in $melt$-In_2O_3 and a-In_2O_3 for further insight into our novel a-In_2O_3 structure. We recall that in c-In_2O_3, the polyhedra network consists of exactly 6 corner-sharing and 6 edge-sharing polyhedra around each In atom. We observe in Figure 4.10 that on average $melt$-In_2O_3 exhibits less edge-sharing polyhedra than the crystal, and a-In_2O_3 both more corner-sharing polyhedra and less edge-sharing polyhedra than the crystal. However, as shown in Figure 4.11, $melt$-In_2O_3 and a-In_2O_3 exhibit a distribution of corner-sharing and edge-sharing polyhedra around these averages, unlike the crystal. On the x-axis of this figure, we display the possible numbers of polyhedra around a single In atom, and on the y-axis we show the frequency, with the black dotted line showing the value of 6 corresponding to the crystal. In both systems, we observe a shift toward higher values of corner-sharing polyhedra and toward lower values of edge-sharing polyhedra. The wide gap between the two in $melt$-In_2O_3 explains the decrease in coordination number around each In atom, as corner-sharing polyhedra share only one O atom with each other and edge-sharing polyhedra share two O atoms. This decrease in coordination number further explains the decrease in In-O bond length observed in $melt$-In_2O_3. In addition, the increased overlap between the corner-sharing and edge-sharing polyhedra upon amorphization shown on the right of Figure 4.11 explains the return to coordination numbers and bond lengths close to those of c-In_2O_3.

We next examine the presence of under- and over-coordinated defects in a-In_2O_3 in Figure 4.12 and explicitly show the relationship between coordination and polyhedra type (edge- or corner-sharing). With undercoordinated ($N_{In-O} < 6$) sites shown in yellow, overcoordinated ($N_{In-O} > 6$) shown in blue, and octahedrally coordinated sites ($N_{In-O} = 6$)
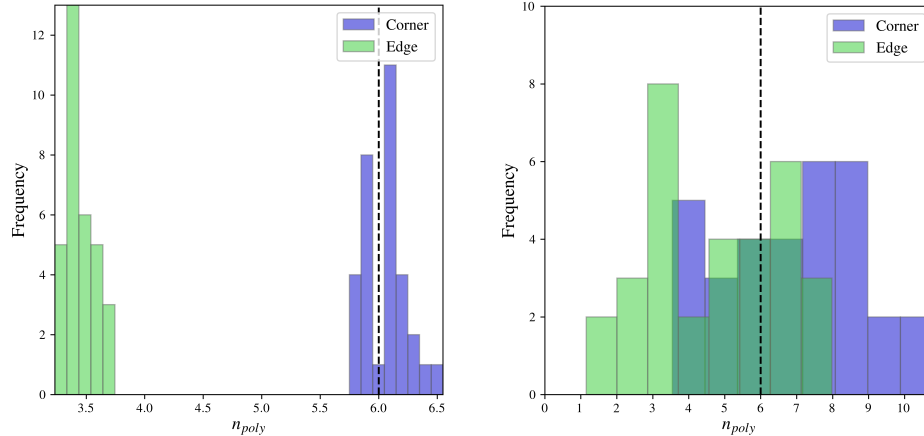
Figure 4.11: Distribution of corner-sharing and edge-sharing polyhedra in $melt$-In$_2$O$_3$ $left$ and a-In$_2$O$_3$ $right$.

shown in green. We clearly see that increasing coordination (going from yellow to green to blue) is negatively correlated with number of corner-sharing polyhedra, as shown in the left plot, and is positively correlated with number of edge-sharing polyhedra, as shown in the right plot. We also observe that in our ab $initio$ LQMD-produced a-In$_2$O$_3$, the majority of In atoms are octahedrally coordinated, with many more undercoordinated sites than overcoordinated sites. We will investigate the relationship between the coordination defects in a-In$_2$O$_3$ and the existence of bound O_2 pairs in future research.

Lastly, because the bound O_2 pair in amorphous alumina was identified as a hole trap, we investigate the possibility of such a pair serving as an electron or hole trap in a-In$_2$O$_3$. In Figure 4.13, we show the electronic density of states (DOS) for c-In$_2$O$_3$ (top) and a-In$_2$O$_3$ (bottom) projected onto In and O atoms, as well as the bound O_2 pair found in a-In$_2$O$_3$. It is clear that the O_2 pair is the main contributor to the state that arises in the gap near the valence band of a-In$_2$O$_3$. However, due to the size and timescale limitations of our ab $initio$ LQMD simulations, we must examine the DOS for other samples of a-In$_2$O$_3$
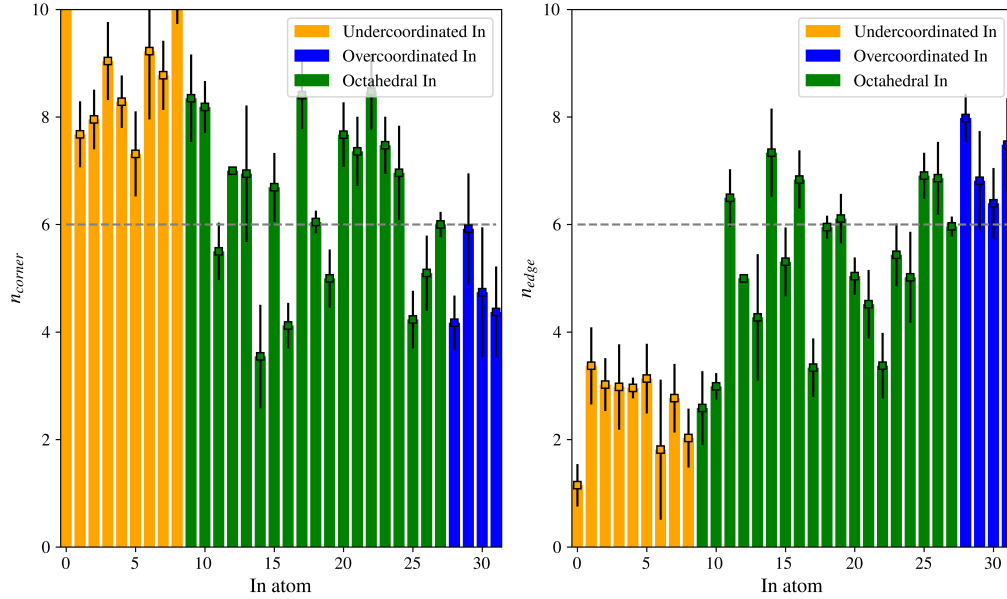
Figure 4.12: Number of corner-sharing polyhedra (*left*) and edge-sharing polyhedra (*right*) around each In atom in a-In$_2$O$_3$. We see that undercoordinated In atoms tend to have more corner-sharing polyhedra and less edge-sharing polyhedra than overcoordinated In atoms.

in order to get a general, representative idea of the DOS for a real a-In$_2$O$_3$ system before we can identify the bound O_2 as a trap.

4.2.2   NNP LQMD

Next, we evaluate our classical LQMD results using a neural network potential (NNP) trained with the *ab initio* LQMD results described in the previous section. In Figures 4.14 - 4.16, we plot the total $g(r)$ for c-In$_2$O$_3$, *melt*-In$_2$O$_3$, and a-In$_2$O$_3$ from LAMMPS using the NNP, with the DFT $g(r)$ shown for comparison and observe exceptional matching between the classical LQMD and the *ab initio* LQMD. These simulations were ran with the same starting configurations as the DFT simulations, and the systems evolved according to
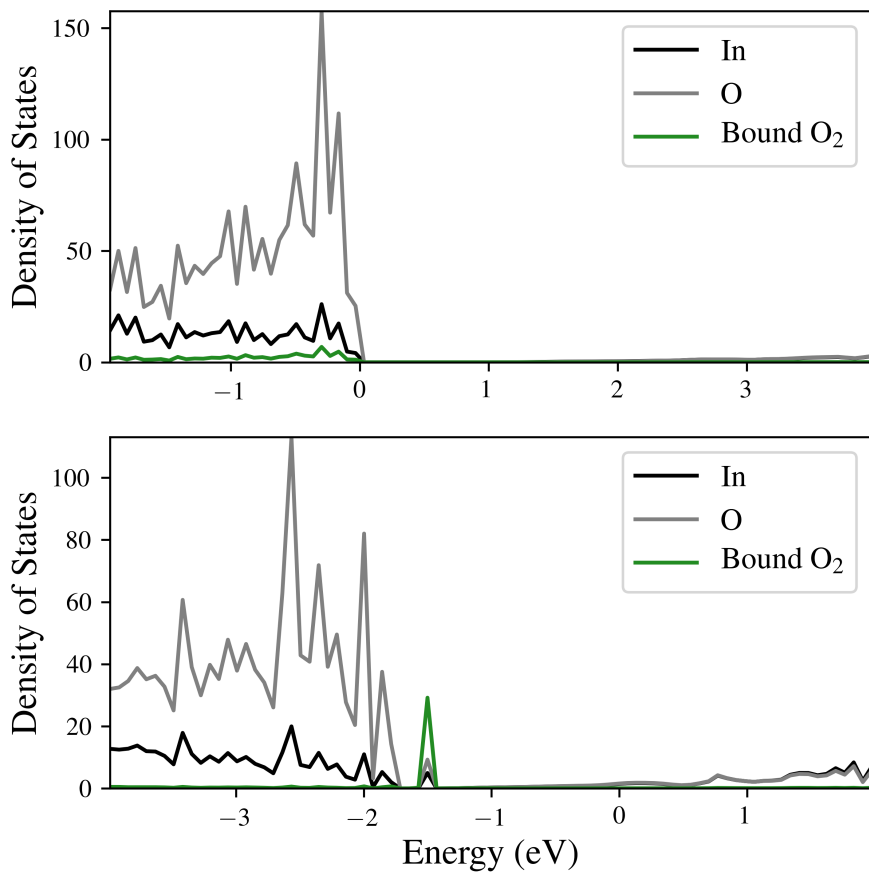
Figure 4.13: Projected density of states onto the bound O_2 pair, all other O atoms, and all In atoms for c-In_2O_3 *(top)* and a-In_2O_3 *(bottom)*

the same ensemble conditions.

Next, we use the NNP to generate new amorphous structures and compare their electronic structures to the VASP a-In_2O_3 structure discussed in the previous section. To generate these structures, we used the melted structure from VASP and quench the system at different quench rates. In Figure 4.17, we show the $g(r)$ for each of these a-In_2O_3 systems, verifying that we have obtained different structures.

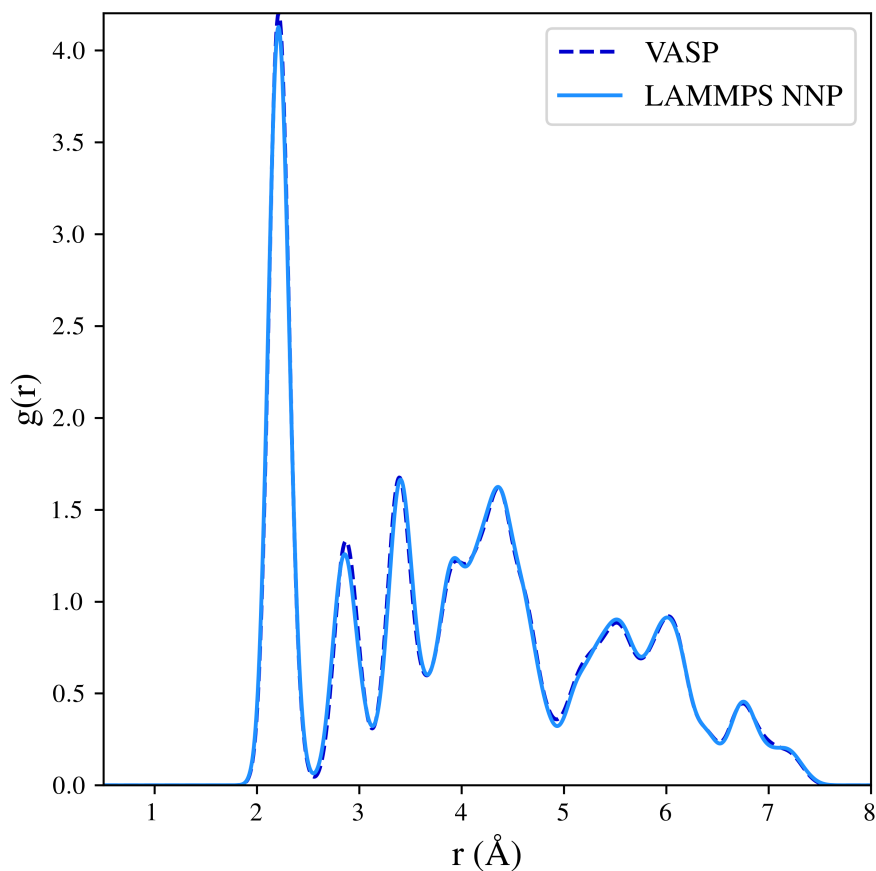Lastly, we examine the HOMO and LUMO energies of samples from each a-In_2O_3

Figure 4.14: $g(r)$ for c-In_2O_3 from LAMMPS NNP simulations, with VASP $g(r)$ shown for comparison.

simulation to identify potential traps in these different samples. In Figure 4.18, we plot the distance of the HOMO and LUMO energies from the Fermi energy and observe variance from the average in a few cases. The VASP-generated sample corresponds to structure 11 in this plot, for which we observe an outlier in the LUMO distance. This seems to confirm our suspicion that this structure contains a trap, namely the bound O_2 pair discussed in the previous section. However, further analysis is necessary to validate this trap, as well as the other outliers in this plot as traps.

Figure 4.15: $g(r)$ for *melt*-In$_2$O$_3$ from LAMMPS NNP simulations, with VASP $g(r)$ shown for comparison.
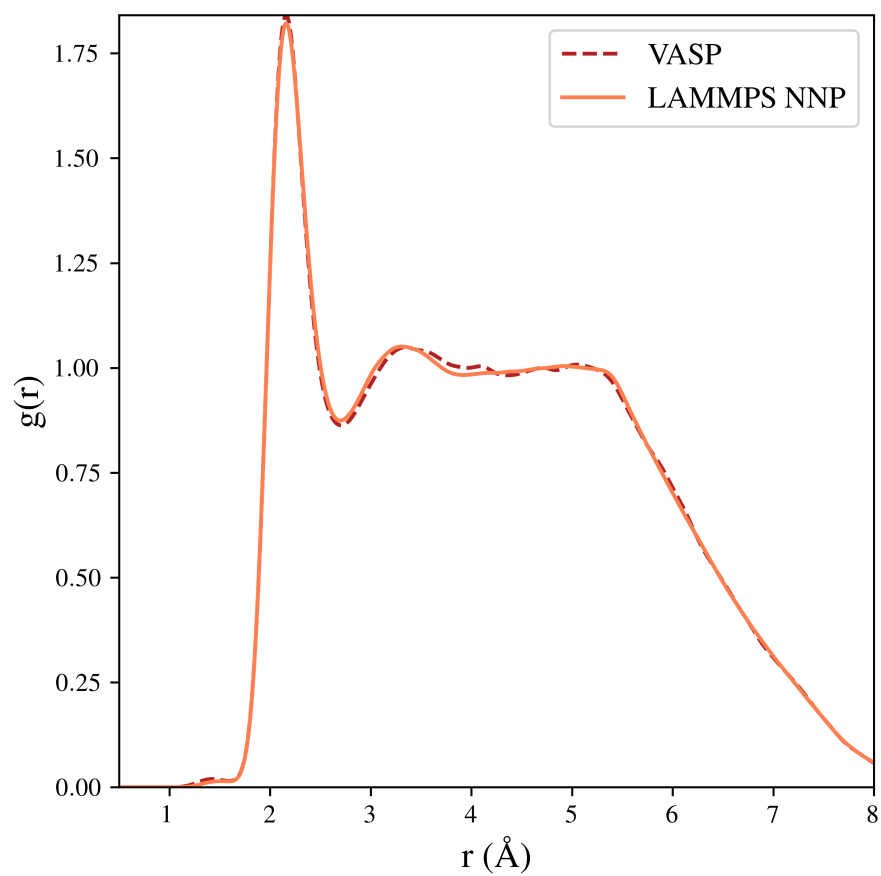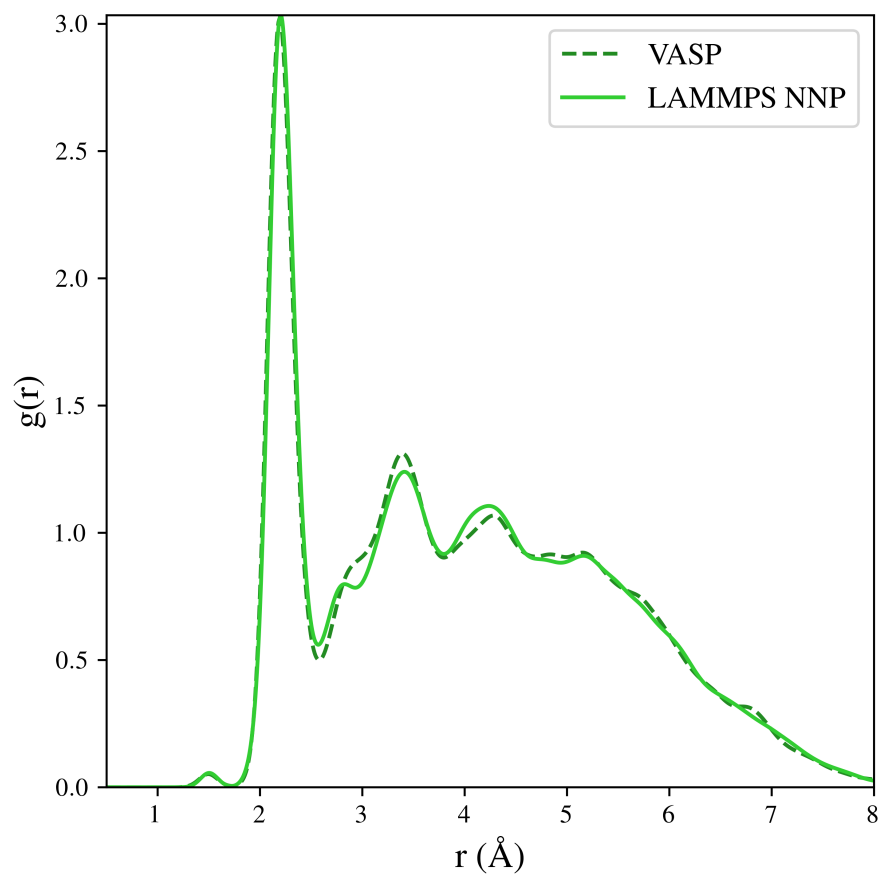
Figure 4.16: $g(r)$ for a-In$_2$O$_3$ from LAMMPS NNP simulations, with VASP $g(r)$ shown for comparison.
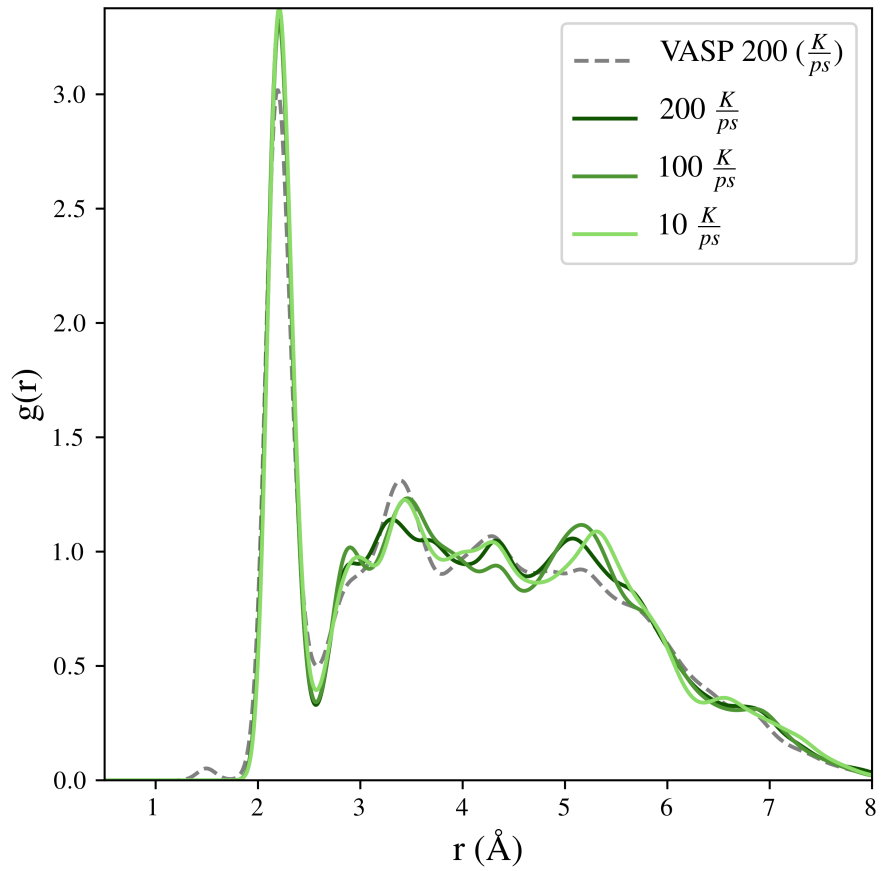
Figure 4.17: $g(r)$ for a-In$_2$O$_3$ from LAMMPS NNP simulations of different quench rates, with VASP $g(r)$ shown for comparison.
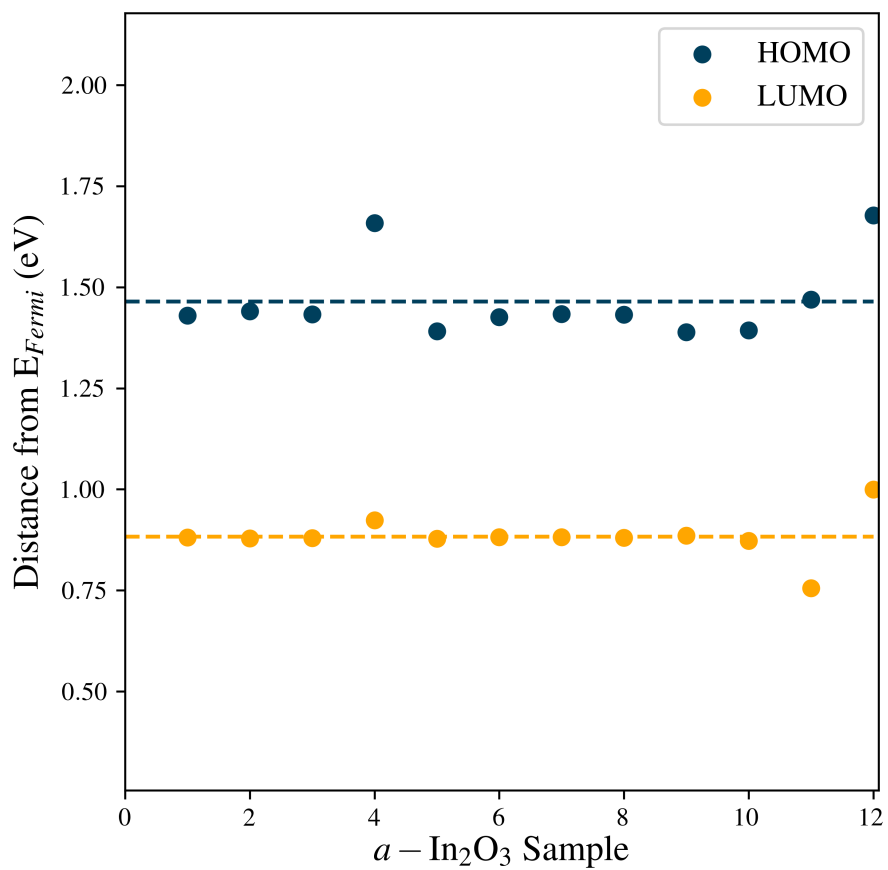
Figure 4.18: The distance of the HOMO and LUMO energies from the Fermi energy as an indicator of electron/hole traps.

4.3 Vanadium oxide

Vanadium (V) oxide (or vanadium pentoxide), V_2O_5 is a multi-functional material that has found widespread applications in fields such as catalysis, energy storage, and biomedical devices. Electronically, V_2O_5 undergoes a fast and reversible transition from a semiconductor to a metal phase near 530 K, drawing appeal for its use in thermally activated electrical and optical switching devices [50]. With regard to energy storage, its layered structure (see Figure 4.19) makes it especially desirable for devices requiring ion intercalation. Amorphous V_2O_5 in particular has a shorter diffusion path and higher capacity than crystalline V_2O_5, drawing attention for use in lithium-ion batteries and in electrochemical energy storage (EES) devices.

In this work, we focus on the use of amorphous V_2O_5 as a surface electron acceptor for the surface transfer doping of diamond in field-effect transistors (FETs). Diamond is of interest for use in high-power electronic devices, but is limited by various fabriaction challenges, particularly mature doping techniques. Due to the high electron affinity of V_2O_5, it has been shown to be effect in doping diamond through the trapping of electrons. In an effort to accelerate the optimization of these devices, here we follow similar steps as in the previous section to identify potential electron traps.

4.3.1 *Ab initio* LQMD

Unlike In_2O_3, the crystal structure of V_2O_5 is layered, as shown in Figure 4.19. The building block of c-V_2O_5 is a VO_5 square pyramid, displayed in Figure 4.20, corresponding to the first shell of V. Thus, as we see in Figure 4.21, the c-V_2O_5 structure consists of layers
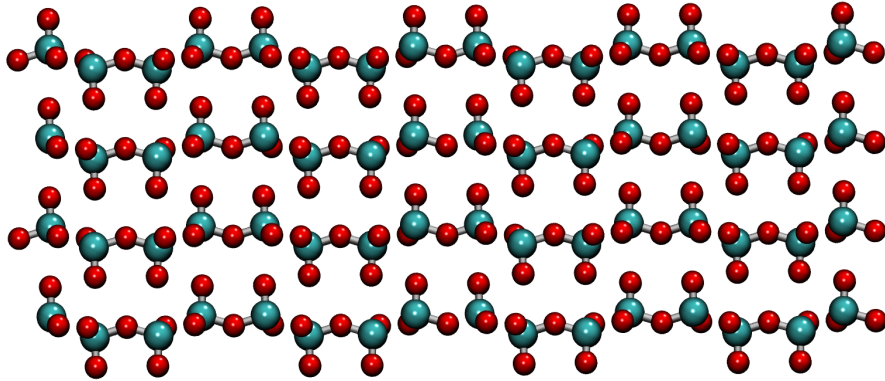
Figure 4.19: Crystal structure of crystalline V_2O_5.

of pairs of VO_5 pyramids.
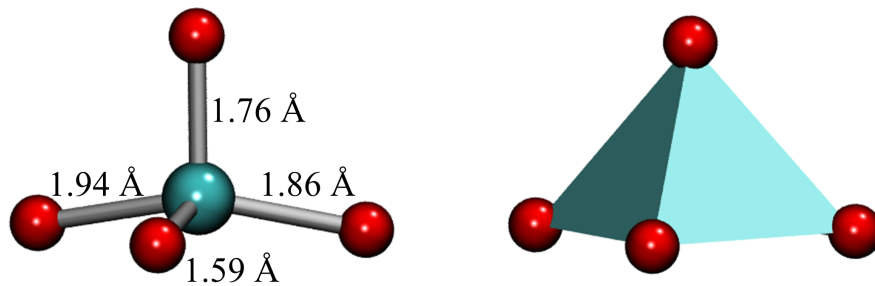
Figure 4.20: Building block of c-V_2O_5, shown as a ball-stick structure *left* and as a polyhedron (square pyramid) *right*.

Similarly to c-In_2O_3, the second and third shells correspond to edge- and corner-sharing polyhedra, respectively, with 2 V atoms in the second shell and 3 V atoms in the third shell.

In Figure 4.23, we plot total energy versus time for each step of the *ab initio* LQMD process, showing the energy for an MD simulation of c-V_2O_5 for reference at the top left. We see that there is a sharp increase in energy during the melting process of
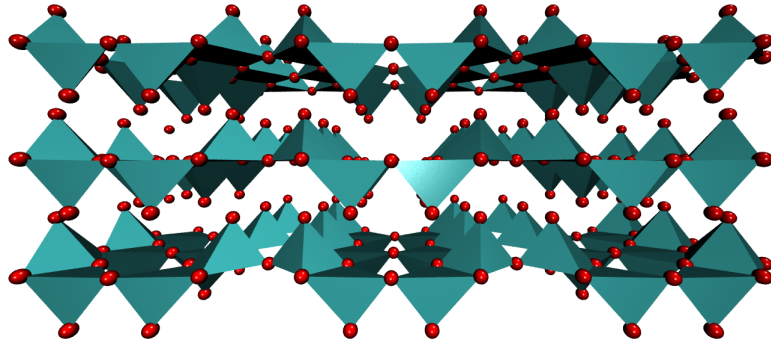
Figure 4.21: Polyhedra network of c-V_2O_5

a.   b.   c.
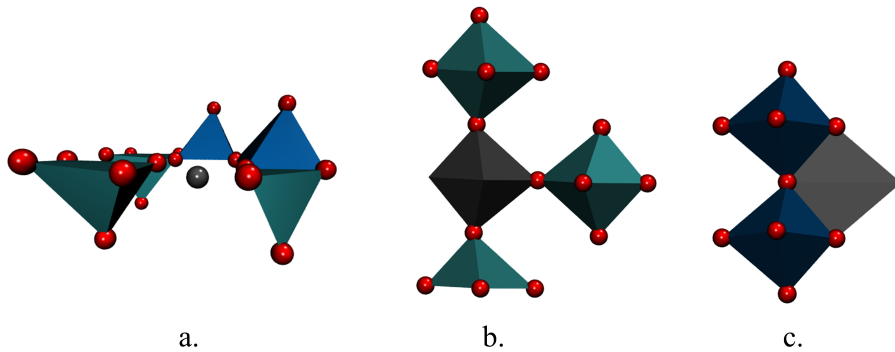
Figure 4.22: Polyhedra network of c-V_2O_5.

c-V_2O_5, which is very likely due to the instantaneous increase in temperature and initial expansion, as we added 0.5 Åof spacing in every direction of the simulation box to allow for proper melting. During the quenching simulation, we observe a steady decrease in energy with time, as expected for a stepwise decrease in temperature over time. Lastly, we see that a-V_2O_5 equilibrates quickly after the quenching.

Next, we plot the pair distribution functions $g(r)$ for c-V_2O_5, $melt$-V_2O_5, and
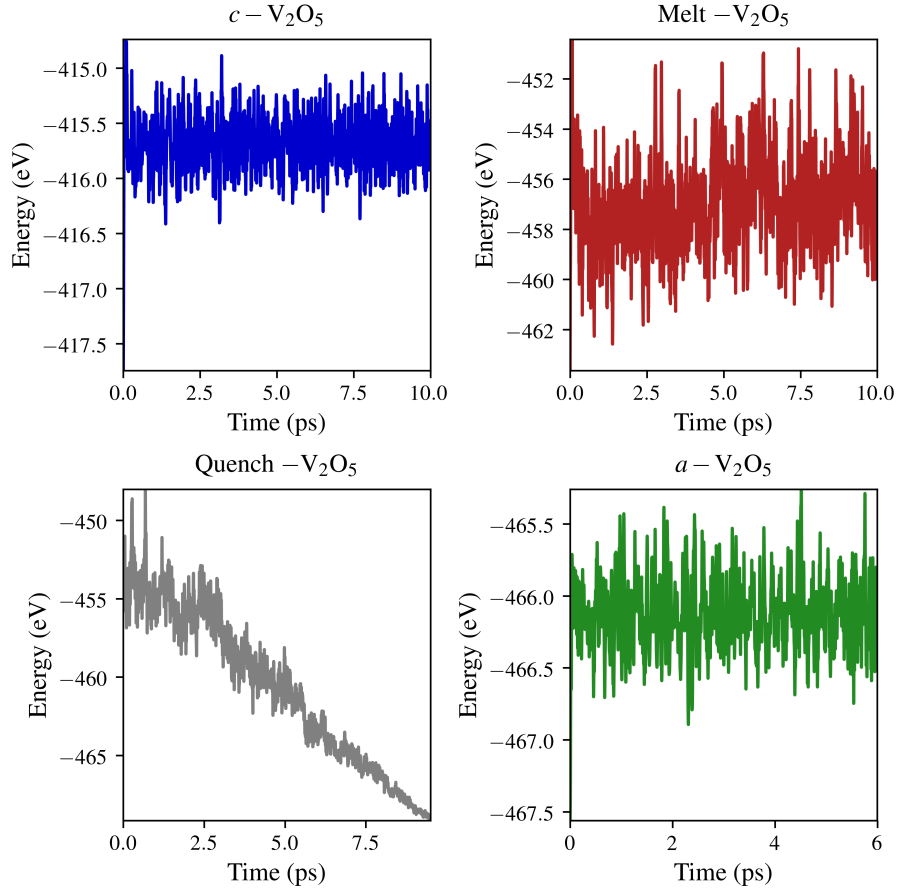
Figure 4.23: Total energy for each simulation of LQMD.

a-V_2O_5 in Figure 4.24. As expected, upon melting, we observe a large reduction in short-range order and a complete loss of medium-range order (above 3 Å.) Upon quenching and equilibration, the $g(r)$ of a-V_2O_5 regains short- and medium-range order, but the short-range order is very different from that of the crystal. In Figure 4.25, we show that the first two peaks of the c-V_2O_5 $g(r)$ correspond to the first shell of V-O (the VO_5 pyramid). We do not observe such double peaks in the $g(r)$ of a-V_2O_5.

In fact, in the a-V_2O_5 $g(r)$, we observe a transformation of all double peaks in the
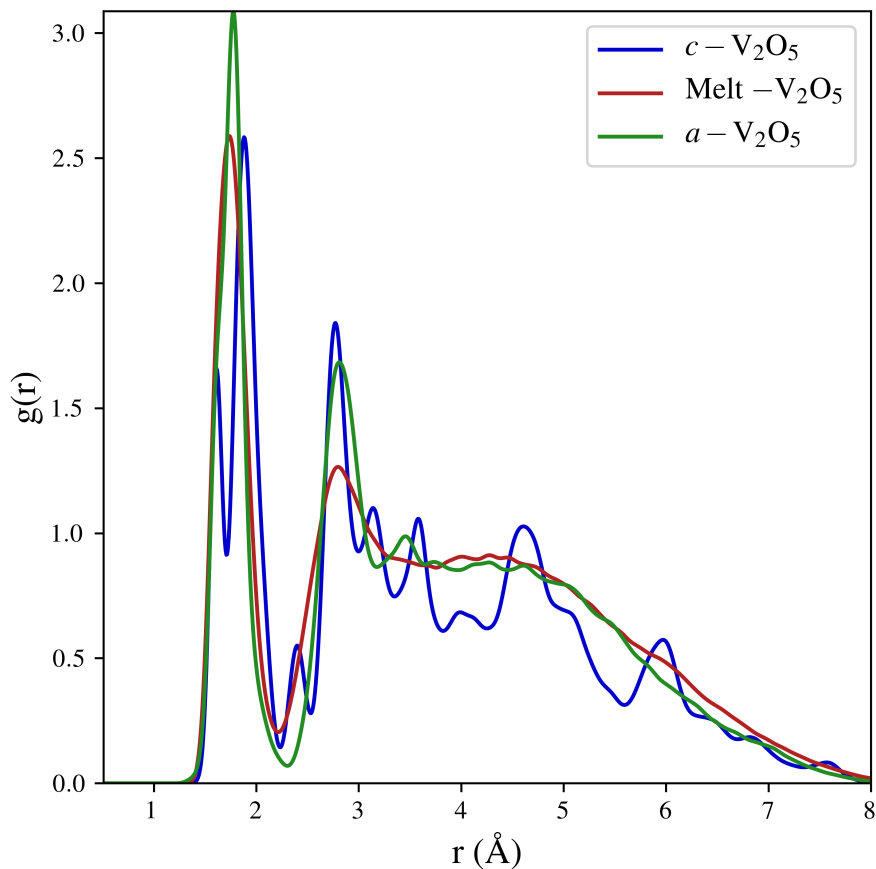
Figure 4.24: Pair distribution functions $g(r)$ for c-V_2O_5, $melt$-V_2O_5, and a-V_2O_5

crystal to broader single peaks. In addition, as shown in Table 4.2, we observe a significatn reduction in V-O coordination number, from 5 in c-V_2O_5 to 4 in a-V_2O_5. The presence of such a large number of coordination defects is also confirmed by the existence of only corner-sharing polyhedra in a-V_2O_5.

Such an undercoordinated system can possibly host a large number of traps; however, there is nothing obvious in the structure that we can quickly identify as a trap as we could for In_2O_3. To generate more samples of a-V_2O_5, we utilize the ReaxFF poten-
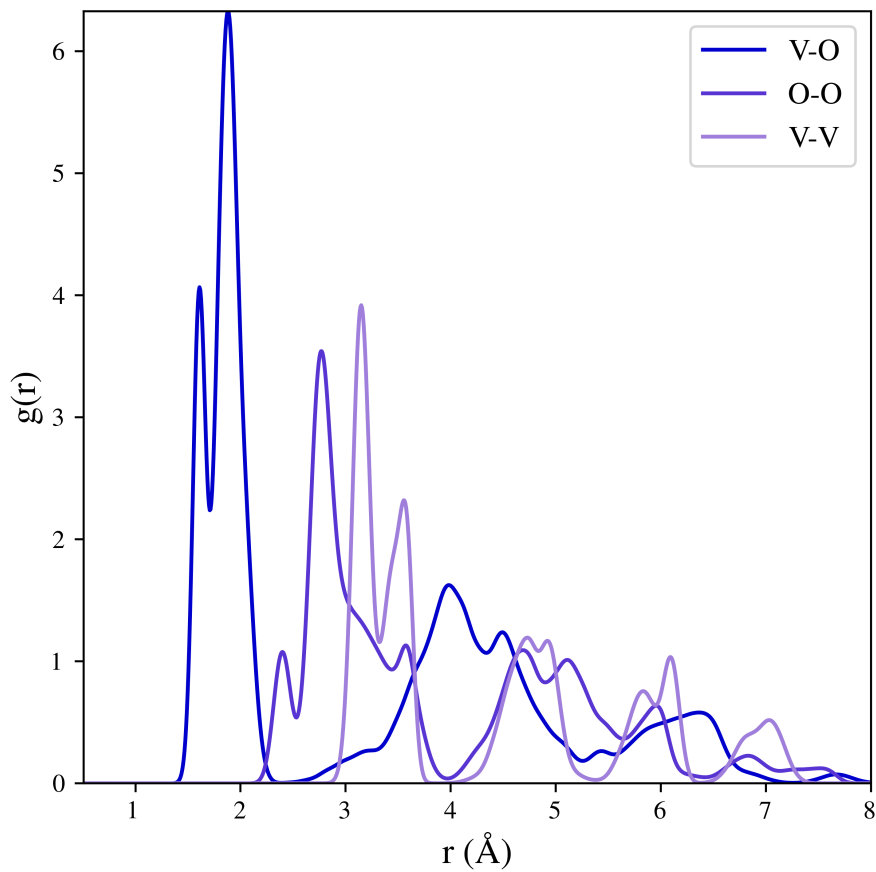
Figure 4.25: Pair distribution functions $g(r)$ for c-V_2O_5

tial to run a series of classical LQMD simulations as we did for In_2O_3. We sample these simulations and in Figure 4.27 show the distance of the HOMO and LUMO energies from the Fermi energy. We observe a wide range in these distances comparable to k_bT, signaling that outliers in these systems could serve as potential traps. However, further analysis is necessary.
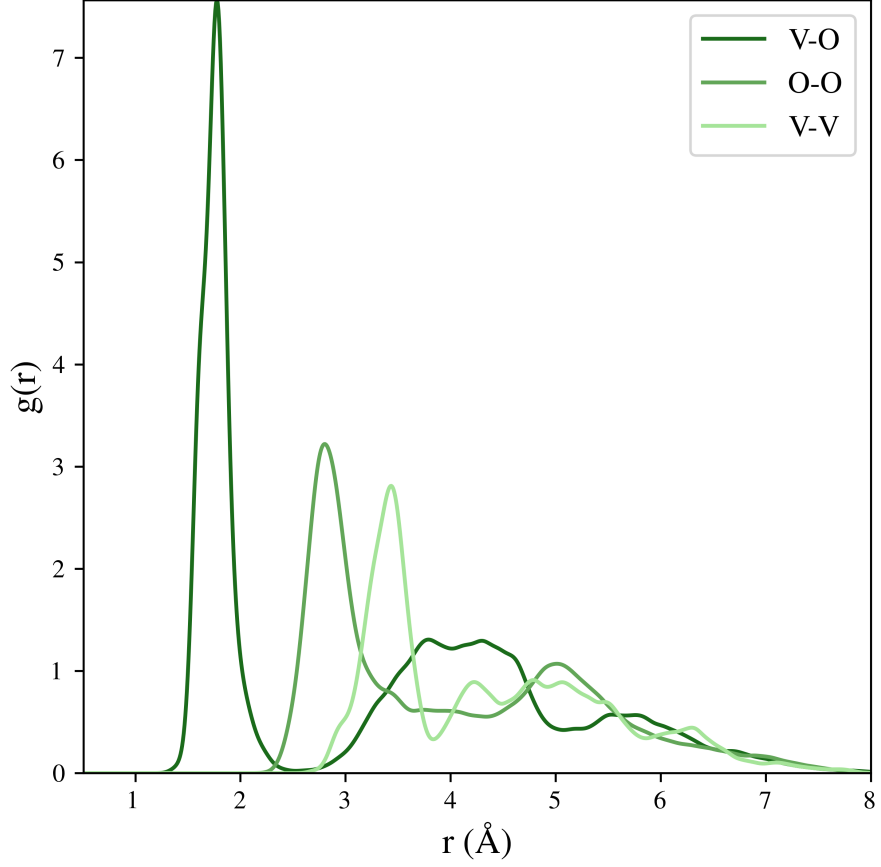
Figure 4.26: Pair distribution functions $g(r)$ for a-V_2O_5

	$N_{V\text{-}O}$	$R_{V\text{-}O}$ (Å)	$R_{V\text{-}V}$ (Å)	$R_{V\text{-}V^*}$ (Å)
c-V_2O_5	5.0	1.61, 1.88	3.15	3.56
$melt$-V_2O_5	4.2	1.74	3.4	NA
a-V_2O_5	4.2	1.78	3.44	4.22

Table 4.2: Coordination number and bond lengths in c-V_2O_5, $melt$-V_2O_5, and a-V_2O_5.
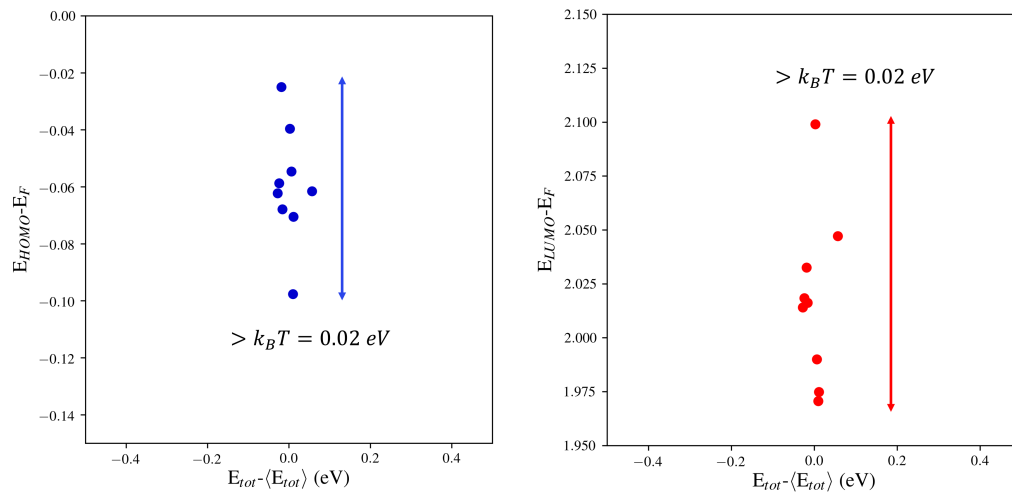
Figure 4.27: The distance of the HOMO and LUMO energies from the Fermi energy as an indicator of electron/hole traps.

4.4 Summary

In this chapter, we generated amorphous structures for two amorphous metal oxides, indium oxide and vanadium oxide, using both *ab initio* MD and classical MD. In both systems, we were able to confirm the loss of long-range order and decrease in medium-range order characteristic of amorphous systems and identified the types of coordination defects in these systems. Lastly, we identified potential traps by analyzing outliers in distance of the frontier orbitals from the Fermi energy of these systems.

Chapter 5

Water-in-Salt Electrolytes

The majority of our digital devices are now powered by lithium-ion batteries, a type of rechargeable battery that relies on oxidation reactions involving Li ions at the electrodes for charging and discharging. However, concerns surrounding their cost, safety, and environmental impact limits their use in electric vehicles and electrical energy storage (EES) on the grid. Most of these concerns stem from the use of non-aqueous electrolytes needed to withstand the high voltages of the battery electrochemistry [51], which contain flammable and reactive ester-based solvents [52], as well as a toxic and thermally unstable salt ($LiPF_6$) [53].

Among aqueous batteries, Zn batteries display the highest energy densities and are inherently safe, inexpensive, and sustainable. Unlike non-aqueous electrolytes, aqueous electrolytes are non-flammable; however, they are limited by the electrochemical stability window of water (1.23 V), which limits the operating voltage of the battery. In addition, the hydrogen evolution reaction (HER) results in parasitic reactions at the electrodes, causing deterioration and hindering reversibility. However, the recently discovered advantages of

73

super concentrated aqueous electrolytes, water-in-salt electrolytes (WiSE), have enabled unprecedented reversibility and performance in Zn-metal batteries, sparking a surge of interest in WiSE-based batteries for EES applications.

The motivation for this research is the experimentally observed stabilization of water in zinc chloride ($ZnCl_2$) WiSE upon addition of lithium chloride (LiCl), as demonstrated by the blue shift of water frequencies in the Raman spectra upon addition of LiCl shown in Figure 5.1. Interestingly, the efficiency of the battery is increased upon addition of 5 m LiCl, but decreases upon any further addition of LiCl, signaling the existence of an intricate interplay between the concentrations of each salt in the electrolyte. To shed light on the underlying mechanism of this synergistic effect, here we examine the effect of LiCl concentration on *i)* the hydrogen bonding network and *ii)* the coordination shell of the Zn cation in $ZnCl_2$ WiSE.
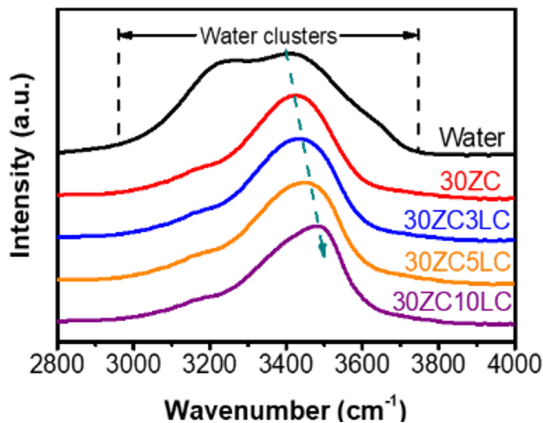

Figure 5.1: Raman spectra for bi-salt $ZnCl_2$/LiCl WiSE of varying concentrations and pure water in the high frequency region.

5.1 Hydrogen bonding network

In Figure 5.2, we plot the O-H pair distribution function $g(r)$ for the $ZnCl_2$ WiSE (30 m $ZnCl_2$) and the $ZnCl_2$ WiSE upon addition of 5 m LiCl and 10 m Li Cl. Upon examination of the first peak, we observe that the OH bond lengths R_{OH} in 30 m $ZnCl_2$ and 30 m $ZnCl_2$ + 5 m LiCl are the same - $R_{OH}^0 = R_{OH}^5 = 0.995$ Å, where the superscript denotes the concentration of Li Cl. However, addition of 10 m LiCl shifts the OH bond length to $R_{OH}^{10} = 1.01$ Å, which is close to the OH bond length in ice. However, as Figure 5.1 shows, the stability of water increases with increasing LiCl concentration, which could imply a reduction in OH bond length. To reconcile our computational results with the Raman spectra, we first note the increase in peak width in the Raman spectra upon addition of LiCl, implying that there is an increasing variety of OH bond lengths upon addition of LiCl. This is shown in Figure 5.3, where we plot R_{OH} for each water molecule in each concentration of simulated $ZnCl_2$/LiCl WiSE.

Next, we examine the second peak of the O-H $g(r)$, which corresponds to hydrogen bonds between water molecules. In the 30 m $ZnCl_2$ WiSE, we observe a peak at $R_{Hbond} = 1.54$ Å, where the subscript $Hbond$ signifies the hydrogen bond between water molecules. Upon addition of 5 m LiCl and 10 m LiCl, we observe decreases in this value to 1.46 and 1.42 Å, respectively. If we average R_{Hbond} and R_{OH}, we obtain 1.27 Å for the 30 m $ZnCl_2$ WiSE and values of 1.23 and 1.21 Å upon addition of 5 m LiCl and 10 m Li Cl, respectively. Therefore, in the Raman spectrum, the single peak observed in the Raman spectra of the WiSE may be an average of the first two peaks in the $g(r)$, explaining the
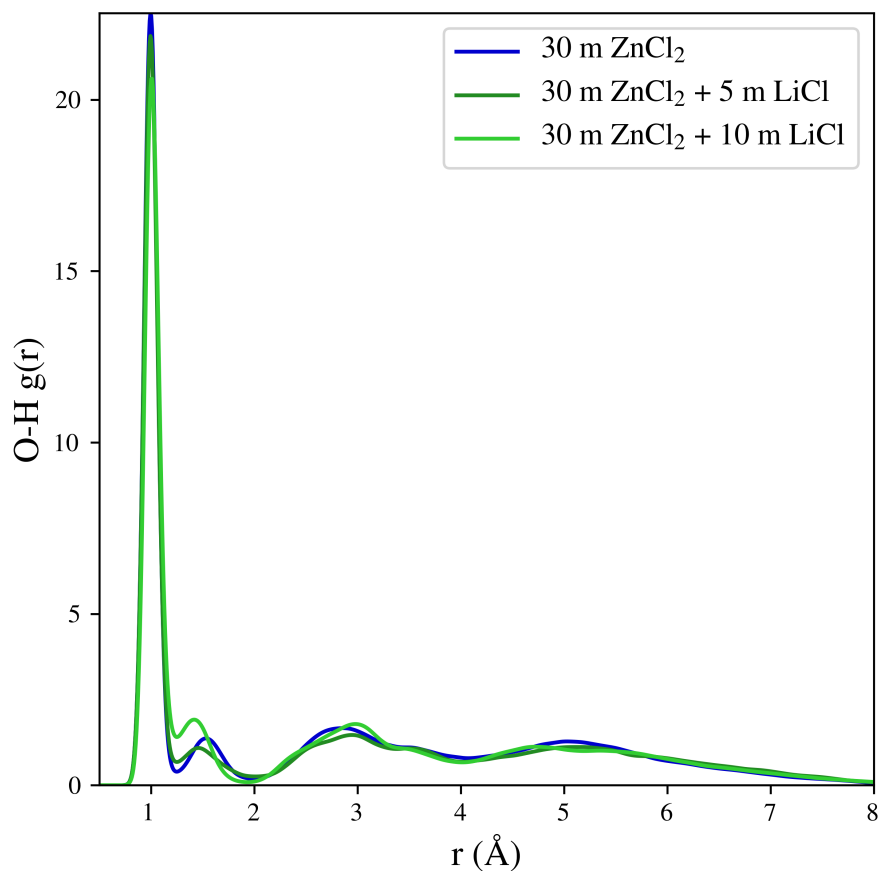
Figure 5.2: O-H pair distribution functions $g(r)$ for bi-salt $ZnCl_2$/LiCl WiSE of varying concentrations.

overall observed blue shift of the peaks in both 30 m $ZnCl_2$ + 5 m LiCl and 30 m $ZnCl_2$ + 10 m LiCl. Furthermore, counting up the number of water molecules contributing to the second peak of the O-H $g(r)$, we observe that upon addition of 5 m LiCl, the number of water molecules participating in hydrogen bonds with other water molecules decreases, from an average of 17.7 during the simulation of 30 m $ZnCl_2$ to 16.4 in 30 m $ZnCl_2$ + 5 m LiCl. However, upon addition of 10 m LiCl, this number increases to 19.1. Therefore, the hydrogen bonding network seems to play the biggest role in the interplay between LiCl concentraion and $ZnCl_2$, the nature of which may not be directly observable from Raman
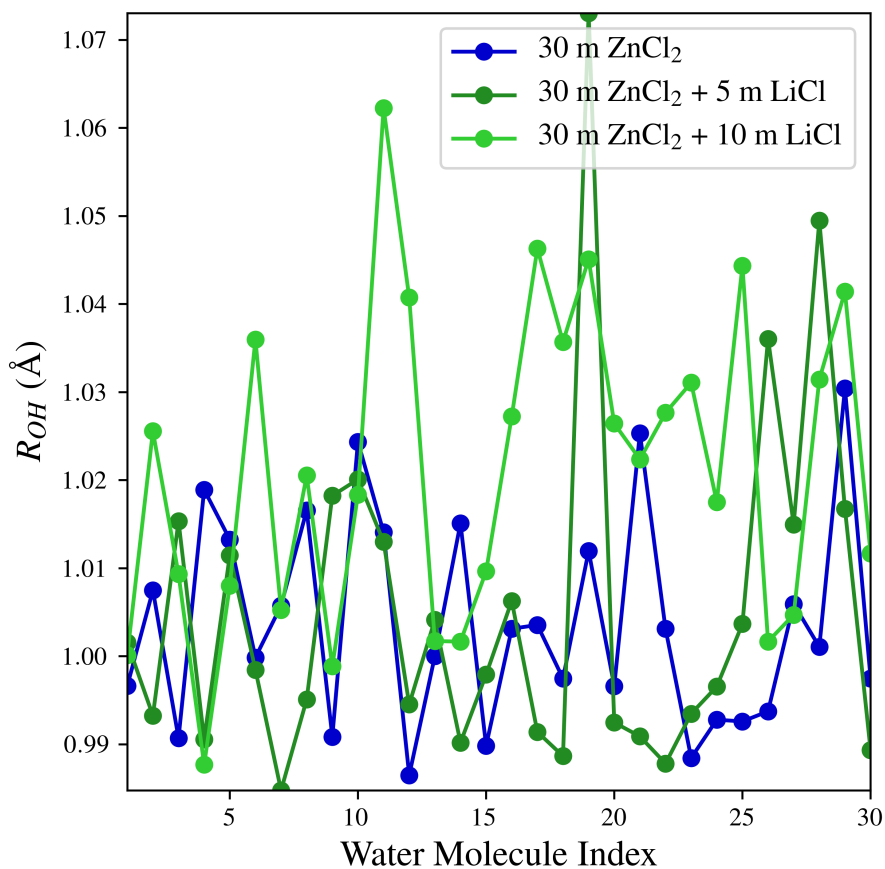
Figure 5.3: The average OH bond length in each water molecule of bi-salt $ZnCl_2$/LiCl WiSE of varying concentrations.

spectra.

Lastly, in Figure we validate our simulation results by comparing the power spectrum of the OH bonds in our 30 m $ZnCl_2$ + 10 m LiCl simulation to the corresponding curve (purple) in the Raman spectrum. As expected, we observe a range of frequencies which is correlated to the OH bond length, where shorter bonds exhibit higher frequencies.
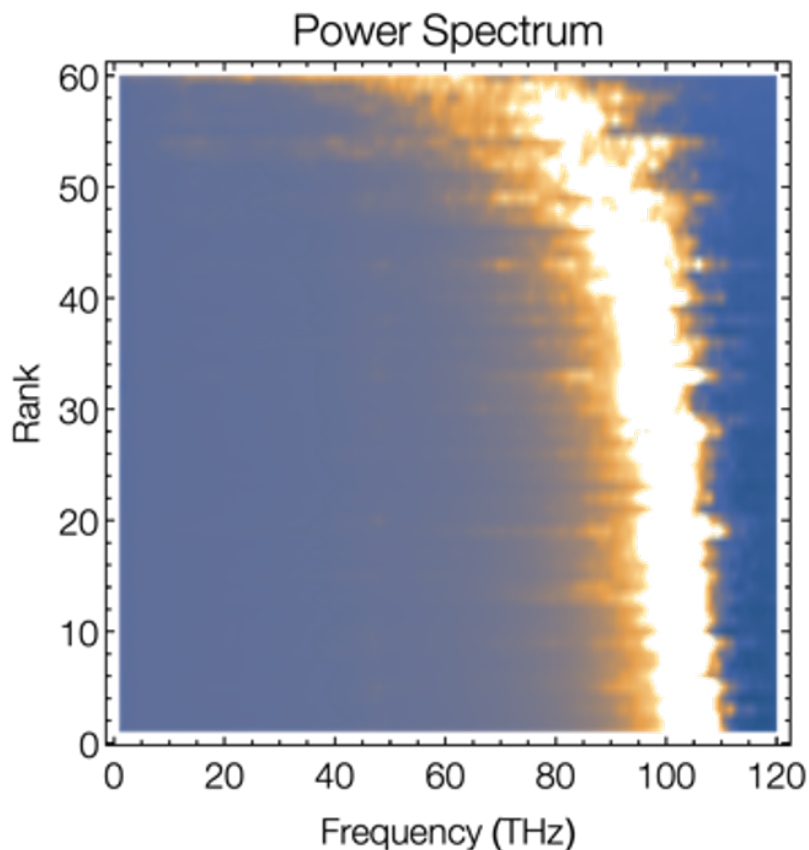
Figure 5.4: Power spectrum derived from the velocity autocorrelation function of the OH bonds in each water molecule of simulated 30 m $ZnCl_2$ + 10 m LiCl. We sort the OH bonds by length from shortest bond lengths to longest bond lengths.

5.2 Cation coordination shells

We next examine the effects of the hydrogen bonding network on the coordination of water to Zn and to Li, starting with Zn. In Figure 5.5, We first observe a decrease in ZnO bond length, R_{ZnO}, upon addition of LiCl from 2.05 Å to 2.02 Å to 2.01 Å in 30 m $ZnCl_2$, 30 m $ZnCl_2$ + 5 m LiCl, and 30 m $ZnCl_2$ + 10 m LiCl, respectively.

Counting up the number of water molecules coordinated to each Zn atom, we observe that upon addition of 5 m LiCl, the number of water molecules coordinated to Zn
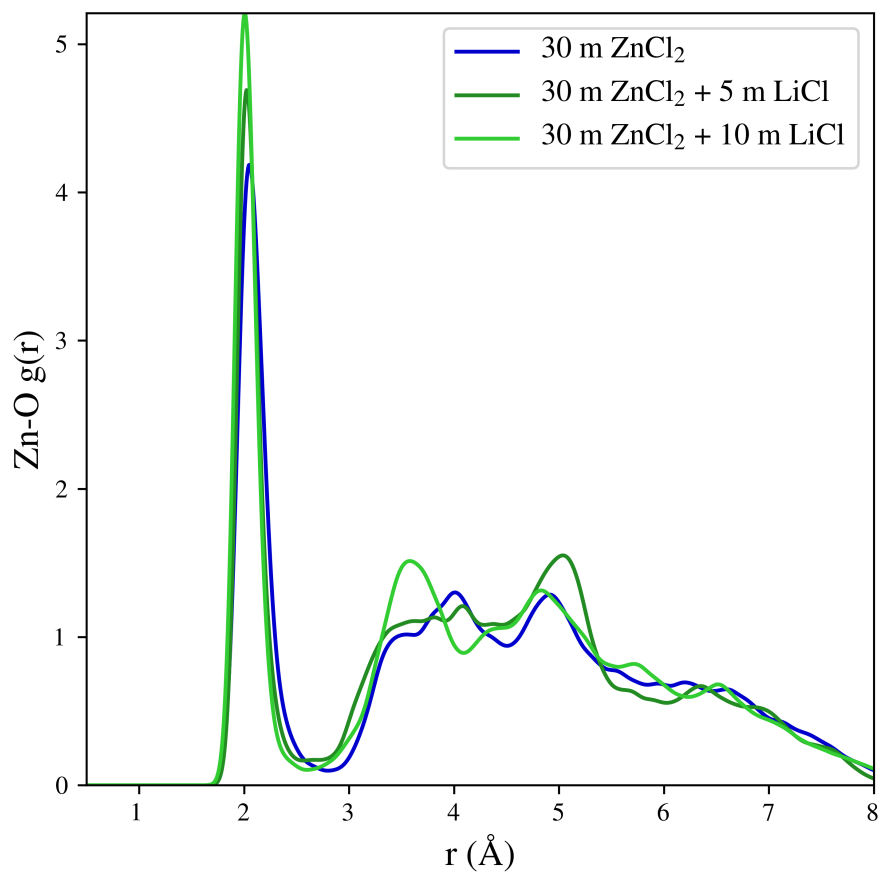
Figure 5.5: Zn-O pair distribution functions $g(r)$ for bi-salt $ZnCl_2$/LiCl WiSE of varying concentrations.

atoms decreases to 26, while uppon addition of 10 m LiCl, the number of water molecules coordinated to Zn atoms increases to 29. This is confirmed upon examination of Li-O $g(r)$, shown in Figure 5.6, where we observe a very large decrease in the first peak, indicative of a reduction in coordination number of O to Li.
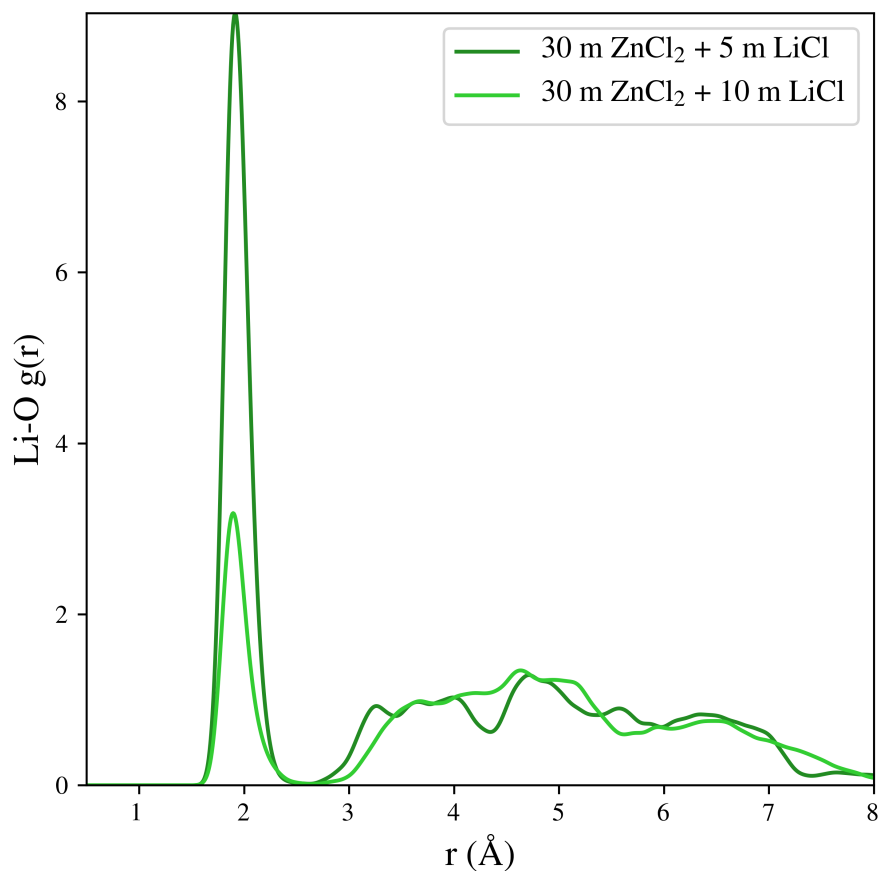
Figure 5.6: Li-O pair distribution functions $g(r)$ for bi-salt $ZnCl_2$/LiCl WiSE of varying concentrations.

5.3 Summary

In conclusion, we observe a delicate interplay in the hydrogen bonding network and the cation coordination shells of bi-salt $ZnCl_2$/LiCl WiSE that validates experimental Raman spectra of these systems. More importantly, we help elucidate the mechanism underlying the initial increase in efficiency upon addition of 5 m LiCl and subsequent decrease in efficiency upon addition of 10 m LiCl. Specifically, we observe that initially, addition of LiCl reduces the number of water molecules participating in hydrogen bonding with other

water molecules, while maintaining the stability of the OH water bond. However, increasing LiCl concentration beyond 10 m LiCl increases the number of water molecules participating in the hydrogen bonding network, accompanied by an increase in OH bond length, possibly leading to instability. Furthermore, the number of water molecules participating in the hydrogen bonding network is correlated with the fraction of water molecules coordinating Zn atoms. As Zn is more electronegative than Li, we expect that charge transfer between Zn and coordinating waters is directly related to the observed experimental results, where 30 m $ZnCl_2$ + 5 m LiCl strikes the correct balance for reducing charge transfer between Zn and water.

Chapter 6

Conclusions

Elucidating the structure-property relationships in amorphous systems is extremely difficult, due to the lack of accurate experimental techniques and a consistent model of their atomic structure. However, in this thesis, we have demonstrated that computational modeling plays a *crucial* role in not only rationalizing experiment, but also guiding them. In particular, we have successfully generated atomic structure models of two novel amorphous systems, amorphous metal oxides and water-in-salt electrolytes, and validated our results with experiments when available. We also demonstrated the importance of generating a variety of different samples of amorphous structures for a single system, as the timescales and system sizes feasible in simulations biases the final amorphous structure.

Bibliography

[1] J. M. Griffin, A. C. Forse, W.-Y.Tsai, P.-L. Taberna, P. Simon, and C. P. Grey. In situ NMR and electrochemical quartz crystal microbalance techniques reveal the structure of the electrical double layer in supercapacitors. *Nature Materials*, 14:812–819, 2015.

[2] O. Pecher, J. Carretero-Gonzalez, K. J. Griffith, and C. P. Grey. Materials' methods: NMR in battery research. *Chemistry of Materials*, 29:213–242, 2017.

[3] P. Biswas, R. Atta-Fynn, and D. A. Drabold. Reverse Monte Carlo modeling of amorphous silicon. *Physical Review B*, 69:195207, 2004.

[4] M. J. Cliffe, M. T. Dove, D. A. Drabold, and A. L. Goodwin. Structure determination of disordered materials from diffraction data. *Physical Review Letters*, 104:125501, 2010.

[5] D. Igram, B. Bhattarai, P. Biswas, and D. A. Drabold. Large and realistic models of amorphous silicon. *Journal of Non-Crystalline Solids*, 492:27–32, 2018.

[6] R. Timilsina and P. Biswas. A study of hydrogen microstructure in amorphous silicon via inversion of nuclear magnetic resonance spectra. *Journal of Physics: Condensed Matter*, 25(16):165801, 2013.

[7] B. Meredig and C. Wolverton. A hybrid computational-experimental approach for automated crystal structure solution. *Nature Materials*, 12(2):123–127, 2013.

[8] M. J. Cliffe, M. T. Dove, D. A. Drabold, and A. L. Goodwin. Structure determination of disordered materials from diffraction data. *Physical Review Letters*, 104:125501, 2010.

[9] G. Opletal, T. C. Petersen, D. G. McCulloch, I. K. Snook, and I. Yarovsky. The structure of disordered carbon solids studied using a hybrid reverse monte carlo algorithm. *Journal of Physics: Condensed Matter*, 17(17):2605, 2005.

[10] C. J. Pickard and R. J. Needs. Ab initio random structure searching. *Journal of Physics: Condensed Matter*, 23(5):053201, 2011.

[11] A. K. Soper. Tests of the empirical potential structure refinement method and a new method of application to neutron diffraction data on water. *Molecular Physics*, 99(17):1503–1516, 2001.

[12] P. Biswas, R. Atta-Fynn, and D. Drabold. The inclusion of experimental information in first principles modelling of materials. *Journal of Physics: Condensed Matter*, 16(44):S5173, 2004.

[13] P. Biswas, D. Tafen, and D. A. Drabold. Experimentally constrained molecular relaxation: the case of gese2. *Physical Review B*, 71:054204, 2005.

[14] K. Prasai, P. Biswas, and D. A. Drabold. Sculpting the band gap: a computational approach. *Scientific Reports*, 5:15522, 2015.

[15] A. Pandey, P. Biswas, and D. A. Drabold. Inversion of diffraction data for amorphous materials. *Scientific Reports*, 6:33731, 2016.

[16] E. Fortunato, P. Barquinha, and R. Martins. Oxide semiconductor thin-film transistors: A review of recent advances. *Advanced Materials*, 24:2945, 2012.

[17] H. Shima, F. Takano, H. Muramatsu, H. Akinaga, I. H. Inoue, and H. Takagi. Control of resistance switching voltages in rectifying pttioxpt trilayer. *Applied Physics Letters*, 92:043510, 2008.

[18] S. Clima, Y. Y. Chen, R. Degraeve, M. Mees, K. Sankaran, B. Govoreanu, M. Jurczak, S. De Gendt, and G. Pourtois. First-principles simulation of oxygen diffusion in hfox: Role in the resistive switching mechanism. *Applied Physics Letters*, 100:133102, 2012.

[19] K. G. Crawford, J. D. Weil, P. B. Shah, D. A. Ruzmetov, M. R. Neupane, K. Kingkeo, A. G. Birdwell, and T. G. Ivanov. Diamond field-effect transistors with v2o5-induced transfer doping: Scaling to 50-nm gate length. *IEEE Transactions on Electron Devices*, 67(6):2270–2275, 2020.

[20] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.

[21] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[22] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[23] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[24] V.L. Deringer and G. Csányi. Machine learning based interatomic potential for amorphous carbon. *Physical Review B*, 95(9):94203, 2017.

[25] S. Fujikake, V.L. Deringer, T. H. Lee, M. Krynski, S. R. Elliott, and G. Csányi. Gaussian approximation potential modeling of lithium intercalation in carbon nanostructures. *The Journal of Chemical Physics*, 148(24):241714, 2018.

[26] F. C. Mocanu, K. Konstantinou, T. H. Lee, N. Bernstein, V. L. Deringer, G. Csányi, and S. R. Elliott. Modeling the phase-change memory material, ge2sb2te5, with a machine-learned interatomic potential. *The Journal of Physical Chemistry B*, 122(38):8998–9006, 2018.

[27] J. Behler. Representing potential energy surfaces by high-dimensional neural network potentials. *Journal of Physics: Condensed Matter*, 26(18):183001, 2014.

[28] J. Behler. Constructing high-dimensional neural network potentials: A tutorial review. *International Journal of Quantum Chemistry*, 115(16):1032–1050, 2015.

[29] M. Behler, J.; Parrinello. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Physical Review Letters*, 98(14):146401, 2007.

[30] J. R. Boes and J. R. Kitchin. Modeling segregation on aupd(111) surfaces with density functional theory and monte carlo simulations. *The Journal of Physical Chemistry C*, 121(6):3479–3487, 2017.

[31] V. Quaranta, M. Hellström, and J. Behler. Proton-transfer mechanisms at the water–zno interface: The role of presolvation. *The Journal of Physical Chemistry Letters*, 8(7):1476–1483, 2017.

[32] W. Li, Y. Ando, E. Minamitani, and S. Watanabe. Study of li atom diffusion in amorphous li3po4 with neural network potential. *The Journal of Chemical Physics*, 147(21):214106, 2017.

[33] G. C. Sosso, G. Miceli, S. Caravati, J. Behler, and M. Bernasconi. Neural network interatomic potential for the phase change material gete. *Physical Review B*, 85(17):174103, 2012.

[34] N. Artrith, A. Urban, and G. Ceder. Constructing first-principles phase diagrams of amorphous lixsi using machine-learning-assisted sampling with an evolutionary algorithm. *The Journal of Chemical Physics*, 148(24):241711, 2018.

[35] B. Onat, E. D. Cubuk, B. D. Malone, and E. Kaxiras. Implanted neural network potentials: Application to li-si alloys. *Physical Review B*, 97(9):94106, 2018.

[36] N. Artrith, T. Morawietz, and J. Behler. High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide. *Physical Review B*, 83(15):153101, 2011.

[37] J. Behler. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *The Journal of Chemical Physics*, 134(7):074106, 2011.

[38] K. Lee, D. Yoo, W. Jeong, and S. Han. Simple-nn: An efficient package for training and executing neural-network interatomic potentials. *Computer Physics Communications*, 242:2019, 95-103.

[39] T. Blank, S. Brown, A. Cahoun, and D. Doren. Neural network models of potential energy surfaces. *The Journal of Chemical Physics*, 103:4129, 1995.

[40] S. Hobday, R. Smith, and J. Belbruno. Applications of neural networks to fitting interatomic potential functions. *Modelling and Simulation in Materials Science and Engineering*, 7:397, 1999.

[41] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Physical Review B*, 54:11169, 1996.

[42] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Physical Review B*, 59:1758, 1999.

[43] S. Plimpton. Fast parallel algorithms for short-range molecular dynamics. *Journal of Computational Physics*, 117(1):1 – 19, 1995.

[44] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard. Reaxff: A reactive force field for hydrocarbons. *The Journal of Physical Chemistry A*, 105(41):9396–9409, 2001.

[45] B. Radha Krishna, T.K. Subramanyam, B. Srinivasulu Naidu, and S. Uthanna. Effect of substrate temperature on the electrical and optical properties of dc reactive magnetron sputtered indium oxide films. *Optical Materials*, 15(3):217 – 224, 2000.

[46] R. L. Weiher and R. P. Ley. Optical properties of indium oxide. *Journal of Applied Physics*, 37(1):299–302, 1966.

[47] R. L. Weiher. Electrical properties of single crystals of indium oxide. *Journal of Applied Physics*, 33(9):2834–2839, 1962.

[48] L. Wang, Y. Yang, T. J. Marks, Z. Liu, and S. Ho. Near-infrared transparent electrodes for precision teng–man electro-optic measurements: In2o3 thin-film electrodes with tunable near-infrared transparency. *Applied Physics Letters*, 87(16):161107, 2005.

[49] C. Århammar, A. Pietzsch, N. Bock, E. Holmström, C. M. Araujo, J. Gråsjö, S. Zhao, S. Green, T. Peery, F. Hennies, S. Amerioun, A. Föhlisch, J. Schlappa, T. Schmitt, V. N. Strocov, G. A. Niklasson, D. C. Wallace, J. Rubensson, B. Johansson, and R. Ahuja. Unveiling the complex electronic structure of amorphous metal oxides. *Proceedings of the National Academy of Sciences*, 108(16):6355–6360, 2011.

[50] E. E. Chain. Optical properties of vanadium dioxide and vanadium pentoxide thin films. *Applied Optics*, 30(19):2782–2787, Jul 1991.

[51] K. Xu. Nonaqueous liquid electrolytes for lithium-based rechargeable batteries. *Chemical Reviews*, 104:4303–4418, 2004.

[52] L. Hu and K. Xu. Nonflammable electrolyte enhances battery safety. *Proceedings of the National Academy of Sciences*, 111:3205–3206, 2014.

[53] A. Hammami, N. Raymond, and M. Armand. Lithium-ion batteries: Runaway risk of forming toxic compounds. *Nature*, 424:635–636, 2003.

[54] R. Car and M. Parrinello. Structural, dynamical and electronic properties of amorphous silicon: An ab initio molecular dynamics study. *Physical Review Letters*, 60(3):204–207, 1988.

[55] I. Štich, R. Car, and M. Parrinello. Amorphous silicon studied by ab initio molecular dynamics: Preparation, structure, and properties. *Physical Review B*, 44:11092, 1991.

[56] N. C. Cooper, C. M. Goringe, and D. R. McKenzie. Density functional theory modelling of amorphous silicon. *Computational Materials Science*, 17:1–6, 2000.

[57] N. Bernstein, J. L. Feldman, and M. Fornari. Structural model of amorphous silicon annealed with tight binding. *Physical Review B*, 74:205202, 2006.

[58] A. Pedersen, L. Pizzagalli, and H. Jónsson. Optimal atomic structure of amorphous silicon obtained from density functional theory calculations. *New Journal of Physics*, 19:063018, 2017.

Appendix A

Code for pair distribution function

```python
import numpy as np

from ase import atoms

from ase.io import read


class GofR():
    ''' This is a class for computing the pair distribution function of a
    ↪    given
    structure.


    args:
      traj = (ASE atoms object) trajectory/structure
    '''
    def __init__( self, traj ):
```

```python
        self.traj = traj

        self.elements = self.traj[0].get_chemical_symbols()

        self.species = list(set(self.elements))

        self.spec_indices = {}

        for s in self.species:

            self.spec_indices[s] = [i for i,v in enumerate(self.elements)
            ↪    if v == s]

        self.N = 0



    def volume( self, LCC ):

        ''' Computes the volume of the simulation box

        from the matrix of lattice vectors.



        args:

          LCC = (3x3 array) lattice vector matrix [a,b,c]



        returns:

          V = (scalar) the volume of the simulation box'''



        return np.dot(LCC[0],np.cross(LCC[1],LCC[2]))


    def rho_0( self, N, V ):
```

```
    ''' Computes the average density, rho_0,

    from number of atoms, N, and the volume, V,

    of the computational system.


    args:

      N = (scalar) number of atoms in simulation box

      V = (scalar) volume of simulation box

    returns:

      rho_0 = (scalar) the average density'''


    return N/V


def r_ij( self, LCC, positions, indices1, indices2 ):

    ''' Computes r_ij for a given timestep


    args:

      LCC = (3x3 array) lattice vector matrix [a,b,c]

      positions = (n_atoms x 3 array) array of atomic coordinates


    returns:

      r_ij = (array) array of the shortest distances (across

             periodic boundaries) of all pairs of atoms'''
```

```python
        r = positions


        if indices1 != None and indices2 != None:

            v_ij = [ np.abs(np.subtract(r[j], r[i])) for i in indices1 for
            ↪  j in indices2 if j != i ]

        else:

            n_atoms = len(positions)          # number of atoms

            v_ij = [ np.abs(np.subtract(r[j], r[i])) for i in
            ↪  range(n_atoms) for j in range(i+1, n_atoms) ]


        LCC_inv = np.linalg.inv(LCC)     # inverse of lattice vec matrix

        vp_ij = [ np.mod( ( np.dot(LCC_inv,v) + 0.5), 1 ) - 0.5 for v in
        ↪  v_ij ]

        vp_ij = [ np.matmul( LCC, vp) for vp in vp_ij ]


        r_ij = np.linalg.norm(vp_ij,axis = 1)


        return r_ij


    def f( self, r_grid, r_ij, sigma ):

        ''' Computes f, the atomic contributions to the radial density.
```

```
args:

    r_grid = (1D array) the grid over which the Gaussian is computed

    r_ij = (array) the shortest distance(s) between the

          ith and jth atoms, accounting for PBC

    sigma = (scalar) width of Gaussian


returns:

    f = (array) the atomic contribution(s) to the radial density'''


pi_3_2 = np.power( np.sqrt(np.pi), 3)

a = 1 / ( 4 * pi_3_2 * sigma * np.square(r_ij) )

b = - 1 / (sigma * sigma)

return [a[i]*np.exp(b*(r_grid-r)**2) for i,r in enumerate(r_ij)]


def rho( self, cell, positions, ngrid, sigma, indices1=None,
↪   indices2=None ):
    ''' Computes rho, the radial density, as a function of r.


    args:

        N = (scalar) number of atoms in simulation box
```

```
            r_grid = (1D array) the grid over which the Gaussian is
↪   computed

            r_ij = (array) the shortest distance(s) between the

                    ith and jth atoms, accounting for PBC

            sigma = (scalar) width of Gaussian


        returns:

            [r, rho], where

            r = (array) the grid over which the Gaussian is computed

            rho = (array) the radial density as a function of r'''


        curr_r_ij = self.r_ij(cell,positions,indices1,indices2)

        r_grid = np.linspace(0.5,8.0,ngrid)

        if indices1 == None or indices2 == None:

            return [r_grid, np.divide( 2*np.sum( self.f(r_grid, curr_r_ij,
                ↪   sigma), axis = 0 ), self.N )]

        else:

            N1 = len(indices1)

            N2 = len(indices2)

            return [r_grid, np.divide( self.N*np.sum( self.f(r_grid,
                ↪   curr_r_ij, sigma), axis = 0 ), (N1*N2) )]
```

```python
def little_g( self, step, sigma=0.15, ngrid = 100, el1=None, el2=None
    ↪  ):
        ''' Computes R(r), the radial distribution function, and

        g(r), the pair distribution function.


        args:

            sigma = (scalar) width of Gaussian


        returns:

            [r, g], where

            r = (array) the grid over which the Gaussian is computed

            g = (array) the pair distribution function for all element
    ↪  types

            R = (array) the radial distribution function for all element
    ↪  types'''


        positions = self.traj[step].get_positions()

        cell = self.traj[step].get_cell()

        self.N = len( positions )


        V = self.volume(cell)
```

```python
if el1 != None and el2 != None:

    indices1 = self.spec_indices[el1]

    indices2 = self.spec_indices[el2]

    pair = el1+"-"+el2

    coeff = self.N / (len(indices1)*len(indices2))

    rho_r = self.rho(cell,positions,ngrid,sigma,indices1,indices2)


else:

    coeff = self.N

    rho_r = self.rho(cell,positions,ngrid,sigma)

r = rho_r[0]

g = rho_r[1]/self.rho_0(self.N,V)

return np.transpose([r, g])
```

Appendix B

Code for polyhedra network

```python
from ase.io import read

import numpy as np

import pandas as pd


flatten = lambda l: [item for sublist in l for item in sublist]


class Polyhedra():
    ''' This is a class for computing the polyhedra network of

    a material. The methods in this class provide the number of

    each type of polyhedra and the average distance between them.


    args:

        traj = (ASE atoms object) trajectory/structure
```

```python
'''

def __init__(self, traj):

    self.traj = traj

    self.elements = self.traj[0].get_chemical_symbols()

    self.species = list(set(self.elements))

    self.spec_indices = {}

    for s in self.species:

        self.spec_indices[s] = [i for i,v in enumerate(self.elements)
        ↪   if v == s]

def get_shell(self,step, el1, el2, cutoff1, cutoff2):

    indices1 = self.spec_indices[el1] #indices for the central element

    indices2 = self.spec_indices[el2] #indices for the coordinating
    ↪   element


    r = self.traj[step].get_positions()

    LCC = self.traj[step].get_cell()

    n_atoms = len(r)          # number of atoms

    LCC_inv = np.linalg.inv(LCC)     # inverse of lattice vec matrix

    ij = [ [i] + [j] for i in indices1 for j in indices2 ]

    v_ij = [ [np.abs(np.subtract(r[j], r[i]))] for i in indices1 for j
    ↪   in indices2 ]
```

```python
    vp_ij = [ np.mod( ( np.dot(LCC_inv,np.transpose(v)) + 0.5), 1 ) -
↪   0.5 for v in v_ij ]

    vp_ij = [ np.matmul( LCC, vp) for vp in vp_ij ]


    r_ij = np.linalg.norm(vp_ij,axis = 1)

    sorted_indices=np.argsort(np.array(r_ij), axis=None)

    r_ij=np.sort(r_ij,axis=None)

    shell_indices = pd.DataFrame([ ij[sorted_indices[i]] for i,d in
↪   enumerate(r_ij) if d >= cutoff1 and d <= cutoff2 ],

                columns = ['col1','col2'])

    shell_distances = [ d for d in r_ij if d > cutoff1 and d < cutoff2
↪   ]


    shell_indices = shell_indices.groupby('col1')['col2'].apply(list)

    return [shell_indices, shell_distances]


def get_polyhedra(self,step,el1,el2,cutoff1,cutoff2,cutoff3):
    ''' Computes the number of each type of polyhedra (corner-, edge-,
↪   and face-sharing)
    for each atom of type el1.


    args:
```

```
        step = (scalar) timestep to compute polyhedra for

        el1 = (string) element type (e.g. "In") for which to compute
↪    polyhedra

        el2 = (string) element type (e.g. "O") in first coord. shell of
↪    el1

        cutoff1 = (scalar) lower cutoff distance for first shell (usually
↪    0.0)

        cutoff2 = (scalar) upper cutoff distance for first shell (can be
↪    obtained as first minimum

                    in el1-el2 g(r))

        cutoff3 = (scalar) upper cutoff for computing polyhedra - e.g.
↪    for In2O3, corresponds

                    to the largest distance in the third shell around In
↪    (can be obtained as second

                    minimum in el1-el1 g(r))'''


    # first step is to get first shell of el

    first_shell = self.get_shell(step, el1, el2, cutoff1, cutoff2)[0]

    in_shells = self.get_shell(step, el1, el1, cutoff2, cutoff3)[0]


    n_el1 = len(self.spec_indices[el1])

    corner = np.zeros( n_el1 )
```

```python
edge = np.zeros( n_el1 )

face = np.zeros( n_el1 )

# loop through each In atom

for i in range(n_el1):

    O_atoms = first_shell[i]

    for j in in_shells[i]:

        shared_O_atoms = [ v for v in O_atoms if v in
        ↪  first_shell[j] ]

        if len(shared_O_atoms) == 1:

            corner[i] += 1

        if len(shared_O_atoms) == 2:

            edge[i] += 1

        if len(shared_O_atoms) == 3:

            face[i] += 1

return [corner, edge, face]
```