Liquid-vapour phase space of refractory oxides

Timoteüs Bögels

Numéro National de Thèse : 2023ENSL0008

THESE

en vue de l'obtention du grade de Docteur, délivré par

l'ECOLE NORMALE SUPERIEURE DE LYON

Ecole Doctorale N° 52
Physique et Astrophysique de Lyon (PHAST)

Discipline : Sciences de la Terre

Soutenue publiquement le 24/02/2023, par :

Timoteüs BÖGELS

Liquid-vapour phase space of refractory oxides

Espace de phase liquide-vapeur des oxydes réfractaires

Devant le jury composé de :

SOSSI, Paolo	Professeur	ETH Zürich	Rapporteur
BRODHOLT, John	Professeur	University College London	Rapporteur
MICHAUT, Chloé	Professeure des universités	ENS de Lyon	Examinatrice
EVERAERS, Ralf	Professeur des universités	ENS de Lyon	Examinateur
CARACAS, Razvan	Directeur de recherche	IPGP	Directeur de these

Contents

CHAPTER 1

Introduction

Contents

This chapter gives an overview of lunar formation models and how new insights forced them to adapt. We will discuss past and present Moon formation models, their strengths and shortcomings. We will present the research topic and how it can contribute to improving lunar formation models.

1.1 Lunar Formation

The formation of our satellite, the Moon, has been on the minds of people for thousands of years. In ancient times, the existence of the Moon was attributed to a divine origin. As centuries passed and scientific knowledge grew, mankind's curiosity for the universe expanded. The earliest recorded non-divine explanation for the Moon's existence is attributed to the nebular hypothesis. The nebular hypothesis was developed in the late 18th century by Swedenborg (1734), Kant (1755), and Laplace

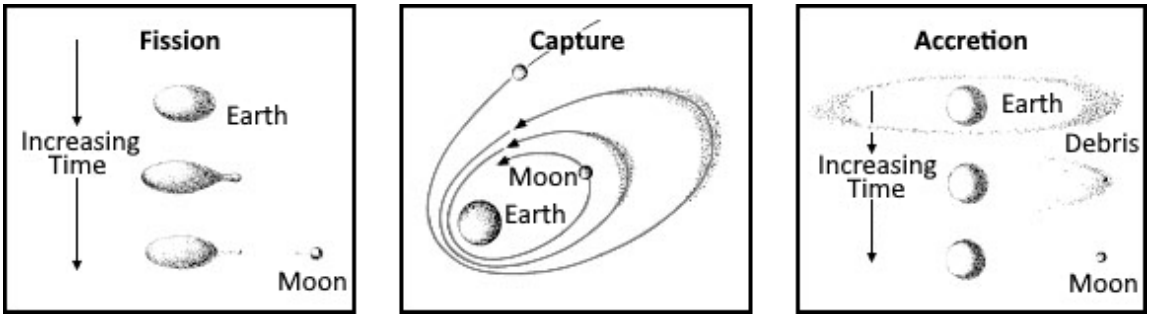


Figure 1.1: The early lunar formation models proposed. From left to right we have; The fission model, where a fast-spinning earth ejects material that forms into the moon. The capture model, where the Moon is formed elsewhere in the solar system to later on be captured by the Earth's gravity during one of its orbits. The accretion model, which suggests the Moon formed in a similar manner as the Earth from material in the protoplanetary disk. Figure from http://ase.tufts.edu/cosmos/ by Professor Kenneth R. Lang

(1796). At the time it was believed that the Moon formed from the same material and at the same time as the Earth, so-called binary accretion (or co-formation).

Later, other suggestions were proposed for the Moon's formation. The most popular of which were the capture model (Gerstenkorn (1955) and MacDonald (1964)) and the fission model (Darwin (1879)), the co-formation model was also still part of the discussion. In the capture model, the Moon is created around the same time as the other moons and planets in our solar system. However, at the time of its formation, its orbit is centred around the sun. Later on, during one of its orbits, the Moon gets captured by Earth, making it the Moon as we know it today. The fission model proposes that the Earth was spinning rapidly while still in its molten state. The model suggests that if the Earth was spinning sufficiently fast it could have ejected some material into an orbit around Earth, which cooled down to form the Moon. A summary of these models is shown in figure 1.1. The discussions around the Moon formation model took an enormous turn after the Apollo missions, which brought back material from the surface of the Moon in the form of soil and rock samples. More specifically, results obtained from isotopic analysis of the returned samples proved to be very difficult to reconcile with lunar formation models at the time.

1.2 Giant impact model

The giant impact hypothesis is currently the leading theory for the formation of the Moon. This hypothesis proposes that a Mars-sized object collided with the Earth early on in its formative history. Material ejected during this collision partially fell back into the Earth and the remainder coalesced to form the Moon of today. However, the models we apply today are the product of decades of iterative improvements.

Two groups proposed similar theories based on different assumptions. Hartmann and Davis (1975) suggested an impact model after constructing accretionary models based on previous work by Safronov, who explained the obliquity of Uranus with an impact. They proposed that a giant impact model has two advantages over the classical fission model; the energy source for ejecting the

lunar material is provided and a giant impact is not purely evolutionary. If the formation of a Moon was something that occurred purely in the evolution of a planetary body, we would expect to find similar relationships between planets and their satellites for other planets in our solar system. The other group by Cameron and Ward (1976) proposed their giant impact model on the basis of the angular momentum constraint of the Earth-Moon system. The angular momentum of the Earth-Moon system consists of three parts: the orbital angular momentum of the Moon with respect to the Earth and the rotational angular momentum of both the Earth and the Moon. In the case of the Earth-Moon system, this total angular momentum is abnormally high compared to other rocky planets in the solar system. Hartmann and Larson (1967) and Fish (1967) noticed how the angular momentum of the Earth-Moon system tends to fall more in line with asteroids and giant planets, when comparing the angular momentum density as a function of mass.

The giant impact hypothesis started to see widespread adoption after the "Conference on the Origin of the Moon" in 1984. Stevenson (1987) provides an excellent review of the dynamical and chemical constraints on the Earth-Moon system and the main propositions for the giant impact model.

1.3 Evolution of the giant impact model

Although there exist some alternative lunar formation models, like the georeactor model presented by Herndon (1992) and more recently assessed by Meijer et al. (2013), in this manuscript we will only focus on models involving an impact. The giant impact model henceforth referred to as the canonical impact model, has been further improved over the past decades. The results of the research done in the two decades following the conference in 1984 are summarized in a review article by Canup (2004a). Models constructed over the years mainly applied Smoothed-particle hydrodynamics (SPH). In SPH models a system is represented using many overlapping spherical particles, each containing a mass. The three-dimensional distribution of these particles is determined by the density weighting function and the smoothing length (the characteristic radial scale). During an impact SPH simulation, the particles' kinematic variables evolve due to changes in gravity, compressional heating, expansive cooling, and shock dissipation. SPH simulations rely on parameters which can be tuned in order to make the simulations match known results. An example of a SPH simulation of a giant impact by Canup (2004b) is shown in figure 1.2.

1.3.1 Isotopes as part of the puzzle

Isotopic analysis of lunar samples from the Apollo missions had a significant impact on the understanding of the evolution, formation, and interior of the Moon. One of the more important being the oxygen isotope values. Robert et al. (1992) showed how samples from a well-mixed planetary body fall on a perfect line of slope 0.5, demonstrating mass-dependant fractionation. Franchi et al. (1999) performed a similar analysis on martian meteorites and found that they also fell on a single line of the same slope, but with an offset of $\pm 0.32\%$ from the Earth one, figure 1.3. Samples from a distinct planetary body fall on their own distinct fractionation line. This demonstrates mass-independent fractionation implying that oxygen isotopes were distributed heterogeneously throughout the solar system (Clayton and Mayeda (1996)). The similarity in the oxygen isotopes between the Earth and Moon demonstrates a strong link between the Earth's mantle and the Moon's origin, providing a significant constraint on any lunar formation model that gets developed.

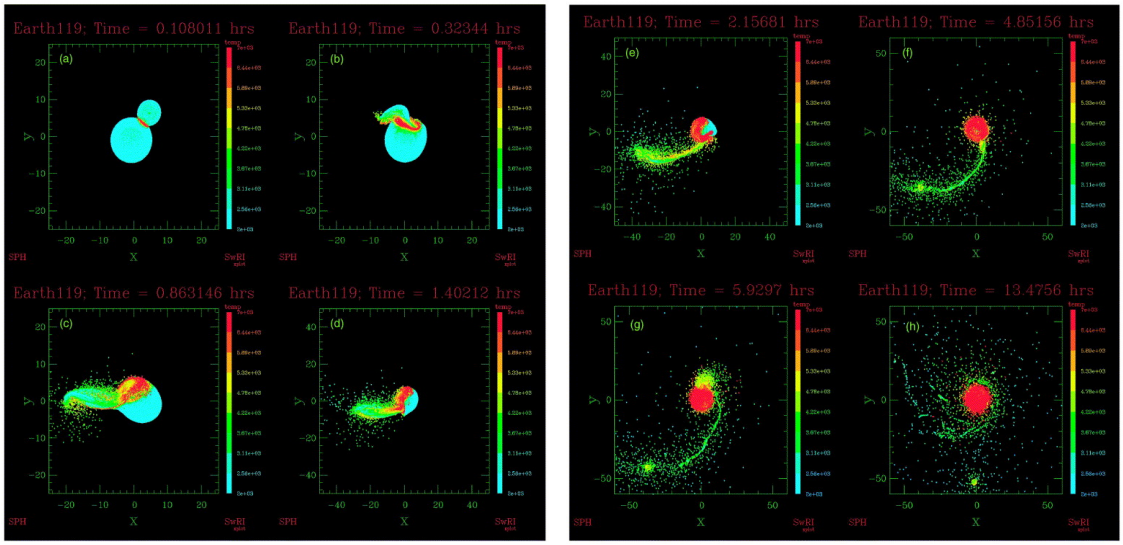


Figure 1.2: SPH simulation of an impact using 60000 particles. The times are shown in hours above each snapshot and the colour scales with particle temperature in K. Figure from Canup (2004b)

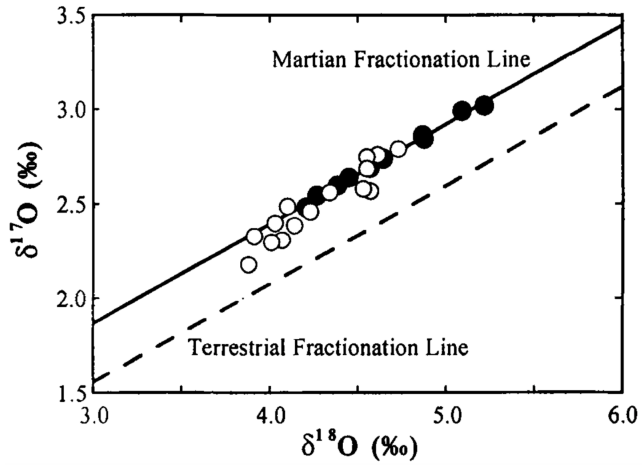


Figure 1.3: Oxygen three-isotope plot (($^{18}O/^{16}O$) versus ($^{17}O/^{16}O$)) of martian meteorites. The terrestrial line as described by Robert et al. (1992) is given as a reference. We can see how samples from a planetary body fall on their own perfect line. Figure from Franchi et al. (1999).

SPH simulations from the beginning of the millennium predicted the Moon to be derived mostly (∼80%) from the impactor. However, if this were the case we would expect to see a difference in the oxygen isotopic fractionation lines. Even a 50/50 mixture of impactor and Earth would make the Earth-Moon similarity difficult to explain. Most of these works are focused on resolving the relatively low amount of iron in the Moon and obtaining the relatively large mass of the Moon in comparison to the Earth. Pahlevan and Stevenson (2007) highlighted this issue and worked on creating a model to allow for equilibration between the Earth and the Moon. They propose diffusive equilibration after the giant impact as a method of increasing the isotopic similarity between the Earth and the Moon.

Oxygen isotopes were of course not the only isotopes to be analyzed and compared between the Earth and the Moon. Besides oxygen, the most often studied are the titanium (^{50}Ti/^{47}Ti) and the tungsten (^{182}W/^{184}W) ratios. Zhang et al. (2012) studied titanium isotopes, as they are distributed heterogeneously throughout the solar system (Trinquier et al. (2009)). Knowing this, we expect the impactor to have a different ^{50}Ti/^{47}Ti ratio than the proto-Earth's and in turn the Moon to reflect this difference. However, they find that this ratio is identical between the Earth and the Moon, within four parts per million. This level of homogeneity can only be explained if the Moon is formed from a significant portion of the proto-Earths' mantle, together with equilibrium mixing within the proto-lunar disk.

Both Kruijer et al. (2015) and Touboul et al. (2015) examined the ^{182}W/^{184}W ratio of lunar samples and terrestrial references. Contrary to the titanium ratio, the tungsten ratio of the modern Earth-Moon system is not homogeneous. They found an elevated ^{182}W/^{184}W ratio within the lunar samples compared to the terrestrial ones. Both studies hypothesize that the early Earth-Moon system was homogeneous. They suggest that the difference we find in the samples today originates from the late veneer. These findings resulted in a constraint on the lunar formation model by way of a homogeneous ^{182}W pre-late-veneer Earth and Moon. This in turn implies that the core of the impactor must have thoroughly stripped the Earth's mantle of highly siderophile elements on its way to merge with the terrestrial core. For an alternative explanation of the ^{182}W/^{184}W ratios Thiemens et al. (2019) propose two other scenarios. The first scenario is that the Moon-forming event happened while core-formation of the Earth was still ongoing. In the second scenario the growing lunar core gathered tungsten, increasing the Hf/W ratio of the Bulk Silicate Moon (BSM).

1.3.2 The angular momentum problem

In this section we will discuss resonance evection, which is the main mechanism proposed to remove angular momentum from the Earth-Moon system. The occurrence of resonance evection during the lunar orbit evolution was first discussed by Kaula and Yoder (1976). However, at the time co-accretion models were more prevalent than giant impact ones and co-accretion scenarios placed the Moon outside of the resonance. Two decades later this topic was picked up again by Touma and Wisdom (1998) who developed it further. A simplified explanation would go as followed; during the early stages of the lunar orbit after formation, tidal effects moved the Moon to a higher orbit. Afterwards, when the period of precession of the lunar perigee (point of the orbit that is the closest to the Earth) equals the period of the Earth's orbit, evection resonance occurs. During this process, the eccentricity of the lunar orbit increases drastically. Figure 1.4 shows a simple illustration of this entire process. The Earth's rotation is gradually decreased by tides raised on the Earth, this decrease of spin has an effect on the oblateness of Earth. The oblateness in turn controls the location of the evection resonance, so as the Earth's rotation is decreased this location is moved inwards. All

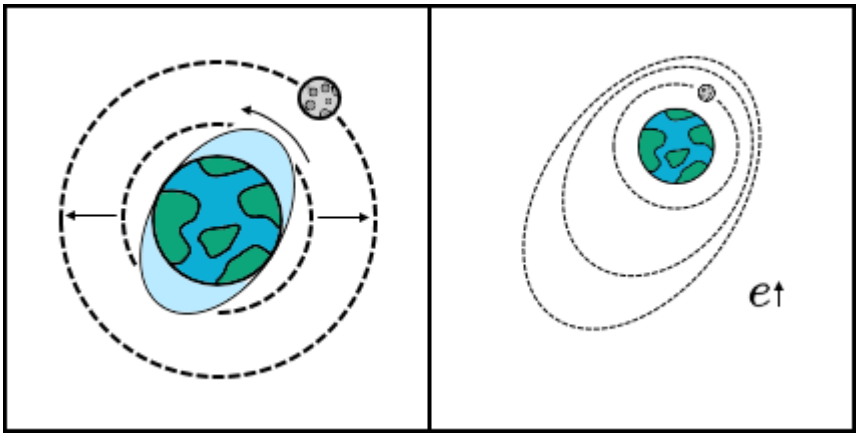


Figure 1.4: Illustration of the changes in the lunar orbit, shortly after its formation. The main mechanism for the removal of angular momentum in the high energy, high angular momentum impact models. Left: The tides on Earth, caused by the Moon, establish a transfer of energy from the Earth to the Moon. This transfer allows the Moon to reach a higher orbit, making the Moon recede from the Earth. Right: At a certain point in the lunar orbit, when the period of the lunar precession is equal to the Earth's orbit (1 year), evection resonance will take effect. During this process, the eccentricity (e) of the lunar orbit increases drastically (Touma and Wisdom (1998)).

the while, angular momentum is transferred from the Earth's spin to the Moon's orbit by tides on Earth. Angular momentum of the lunar orbit can be transferred to the Earth's heliocentric orbit by solar resonant perturbations. Due to the large variety of input variables, a considerable range of outcomes is possible. Examples of studies with varying starting parameters are Deng et al. (2019), Ćuk and Stewart (2012) and Rufu et al. (2017). It should be noted however that the effect of resonance evection may not be sufficient to drain all the required angular momentum, as described by Ward et al. (2020).

To conclude; a modern lunar formation model has to fulfil a set of constraints. The strongest of these constraints are the masses of the Earth and Moon, the angular momentum of the Earth-Moon system, a Moon with a low iron content of roughly 8% by mass, similar stable isotope ratios between the lunar and terrestrial mantle with a degree of lunar volatile depletion (Barr (2016)).

1.4 Modern giant impact scenarios

The overall consensus in the scientific community seems to be that there are currently two or three main stances of giant impact models. All of these models match most of the constraints, but none of them seems to perfectly match the required constraints.

One class of models is based on the canonical impact model, where a Mars-mass impactor collided with the proto-Earth at a low velocity. The key issue for this theory is reconciling the isotopic similarities we find in the modern Earth-Moon system. One way to resolve this issue is to assume the impactor formed with an Earth-like chemical composition. The other, more probable solution, requires intense chemical equilibration during the post-impact phase (e.g. Pahlevan and Stevenson (2007)). The structure of the proto-lunar disk consists of a magma disk in the centre

with a large silicate vapour atmosphere surrounding it. The liquid-vapour exchange between these two reservoirs would allow them to equilibrate.

The other family of impact models relies on a fast-spinning earth (Ćuk and Stewart (2012) and Lock et al. (2018)) and/or a high-energy impact. These models rely on an impact with an energy an order of magnitude higher than the canonical model. This, in turn, gives the proto-lunar disk a significantly different structure called a synestia. These models produce a rapidly rotating disk consisting mainly of vapour, which would allow for equilibration through convective/turbulent mixing. Figure 1.5 shows a schematic illustration of what lunar formation would look like with the synestia model. The defining feature of the synestia is the high-energy vapour structure that facilities chemical equilibration between the proto-Earth material and impactor material. The main complication of these models is the angular momentum, which is often predicted to be higher than what we have today. They rely on evection resonance, as described in Touma and Wisdom (1998), to remove angular momentum from the Earth-Moon system. This process is explained in more detail in section 1.4.1. Some other, less-discussed, impact models are relying on multiple impacts (Rufu et al. (2017)) and a hit-and-run impact (Reufer et al. (2012) and Deng et al. (2019)). The results and associated issues of these scenarios, together with others, are discussed at great length in an excellent review. Canup et al. (2020) provides an extensive up-to-date summary of lunar formation models and has been contributed to by many prominent scientists in the field.

1.4.1 Post-impact dynamics

In order to describe the dynamics post-impact and the evolution of the proto-lunar disk, we first have to decide what lunar formation models to consider. Two main structures for the proto-lunar disk are discussed in the current literature; the vertically stratified one and a majority-vapour well-mixed continuous structure. The former of the two is based on the canonical giant impact, the latter after a high energy, high angular momentum giant impact.

The vertically stratified proto-lunar disk consists of a liquid mid-plane surrounded above and below it by a vapour layer. The layers exchange mass and energy to maintain equilibrium. Radiative energy transport takes place between all three bodies, and between the vapour's photosphere and space.

The majority-vapour well-mixed continuous structure was referred to as the mantle atmosphere and disk (MAD) structure (Lock et al. (2016)). Nowadays the term *synestia* is applied to it; (Lock and Stewart (2017)) stemming from "syn-" meaning together and "Hestia", goddess of the civic and domestic hearth. A synestia can only form if the post-impact body exceeds the co-rotation limit (CoRoL). The CoRoL is defined as the limit for planetary bodies with constant angular velocity. The CoRoL forms a surface that depends on the angular momentum, thermal profile, mass, and compositional layering of the body (Lock and Stewart (2017)). In this structure, there is a co-rotating inner region and a disk-like outer region.

1.5 Refractory elements and the giant impact

Highly refractory elements play an important role in the proto-lunar disk. Refractory elements (or materials) are able to withstand extreme heat and wear, giving them an elevated melting point compared to other elements. The isotopic differences are one of the key topics of debate when reviewing lunar formation models. Volatiles and their refractory counterparts are often discussed when reviewing giant impact models.

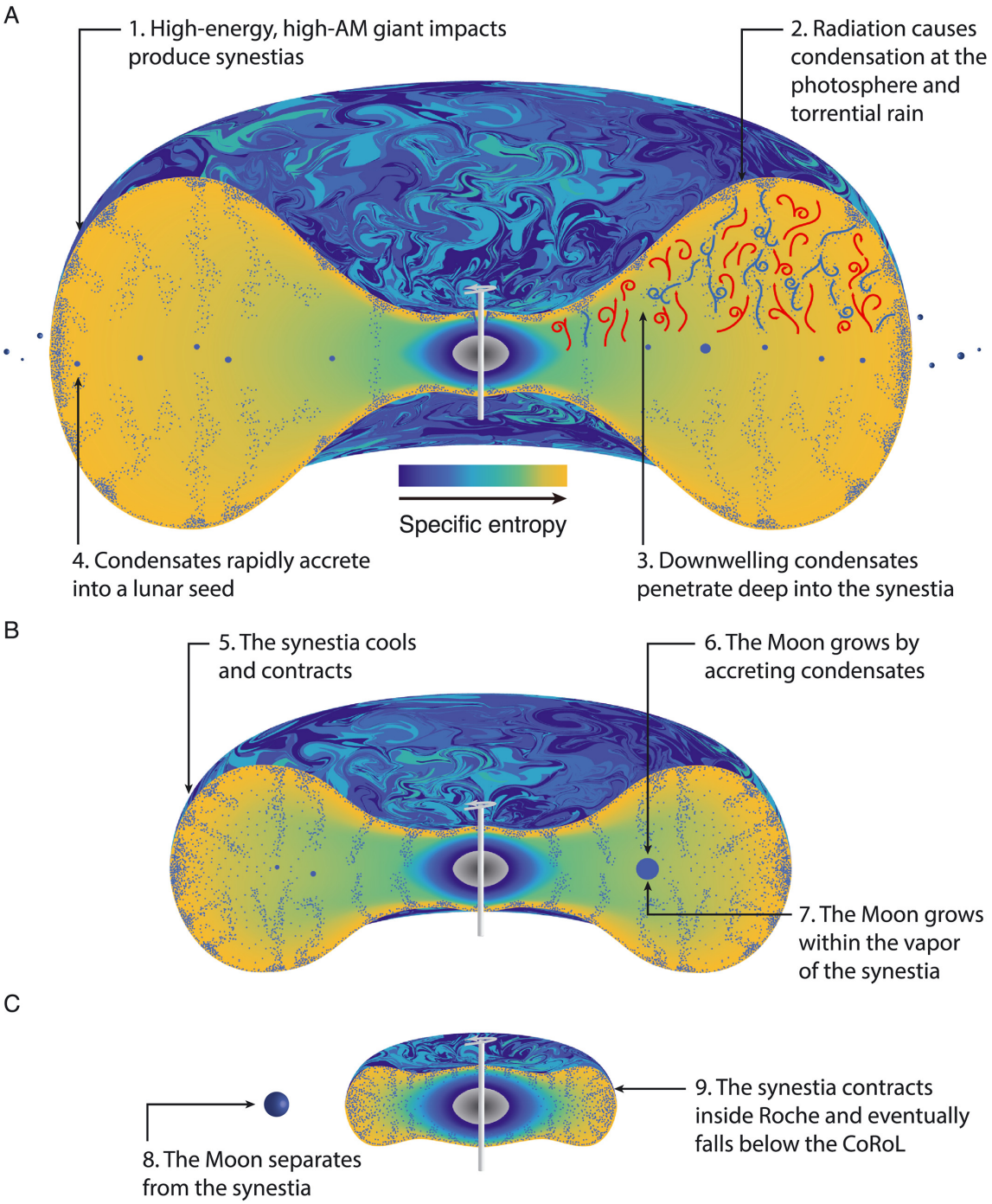


Figure 1.5: A schematic illustration of the synestia structure at three stages. Where A shows the initial formation of the high-energy, high-angular momentum synestia where condensates form the lunar seed material. In B the Moon is growing by accretion of condensates formed in the cooling synestia. At C the Moon is separated from the synestia and the synestia shrinks within the Roche limit, eventually falling below the corotation limit (CoRoL). The phase of the silicate material making up the disk varies with the pressure and specific entropy. Condensates are shown in darker, more blue, colours. Radiative cooling preferentially cools the photosphere creating a large amount of condensates in the region. Turbulent convection between the photosphere and the equatorial plane enables the displacement of condensates. Figure from Lock et al. 2018

The most-discussed refractory element in lunar formation models is Ti, for its $^{50}Ti/^{47}Ti$ ratio. Which, as mentioned in section 1.3.1, is found to be near identical throughout the Earth-Moon system. Other refractory elements like Mg and Ca, in the form of an oxide, are important building blocks of telluric materials. In planetary sciences the refractory nature of elements is often categorized based on the 50% condensation temperatures of the elements at 10^{-4} bar. Based on this classification Mg and Ca would fall into the 'moderately refractory' and 'refractory' categories respectively (Taylor (2001)). In this manuscript we choose to use the term refractory oxides to emphasize the fact that we are dealing with materials that have a higher resistance to temperature than other common geological oxides like silicate and magnesium silicates. Ringwood (1991) show estimates for MgO and CaO in pyrolite model compositions of about 38% and 3.5% respectively. Estimates for the content of these materials in the Moon are given by Wood (1986). Given ranges for MgO and CaO are 29.1-32.7% and 3.7-6.1% respectively. MgO also is one of the end-members of the fundamental MgO-SiO_2 system, making it a particularly interesting material for telluric planets. Planetary mantles of the Earth and Moon are both rich in these oxides. Together with Al_2O_3, refractory oxides account for almost half of both the Earth and the Moon.

As refractory materials are one of the last materials to melt and vaporize during the giant impact, they will also be one of the first to condense again (Lock et al. (2018)). This implies that the thermodynamic behaviour of any liquid inside of a proto-lunar disk, be it a canonical or synestia model, will be dominated by refractory liquids. Understanding the thermodynamic properties of refractory liquids should therefore prove useful when attempting to describe the liquid in a proto-lunar disk.

1.6 Phase diagrams of refractory materials

1.6.1 MgO

Starting with the solid phase of MgO, we can find it on the Earth's surface in its B1 form. The B1 form is also referred to as the rocksalt structure, the $Fm\bar{3}m$ space group. The B2 phase, also known as the caesium chloride structure and $Pm\bar{3}m$ space group, is a common high-pressure phase for many binary compounds. The B1-B2 phase transition was confirmed in an experimental setting by McWilliams et al. (2012), Coppari et al. (2013), and Root et al. (2015) and in a more recent numerical study by Musella et al. (2019). The high-pressure and high-density phase space was experimentally sampled by shock experiments by McWilliams et al. (2012), where the pressure-volume equation of state was measured up to 600 GPa and the temperature and optical reflectivity to beyond 1400 GPa and 50000 K. The experiments first identified a negative Clapeyron slope from about 440 ± 80 GPa to about 9000 ± 700 K, which was related to the solid-solid phase transition of B1 to B2. Next, a positive volume and entropy change around 650 ± 50 GPa and 14000 ± 1100 K was associated with melting. The first numerical results for the B1-B2 phase transition were produced in the late '70s (e.g. Cohen and Gordon (1976)) and in the late 80's the results started to converge significantly, starting with the prediction from Mehl et al. (1988). The phase transition to the more compact cubic B2 phase takes place around 500 GPa at 0 K (Soubiran and Militzer (2020)). The upper-bound limit of this transition is placed in the range of 250 GPa at 9000 K to 400 GPa and 10000 K. There is still some disagreement as to the steepness of the Clapeyron slope of this transition, especially at higher temperatures. These conditions, while not relevant for the Earth, can be reached inside super-Earth exoplanets, where this phase transition might induce further layering in the rocky mantles of these planets (Umemoto et al. (2017)).

The earliest, to our knowledge, reported value of the melting point of MgO can be traced back to Kanolt (1913) where it was calculated at 3073 K. This experiment was performed using a graphic resistance vacuum furnace at atmospheric pressure while measuring the temperature with an optical pyrometer. Some later notable experiments were done in the 60s and 70s where the melting point was recalculated to be slightly higher (R. N. Mcnally et al. (1961) and Leu et al. (1975)). The first numerical calculations (Jackson and Liebermann (1974) and Ohtani (1983)) performed on MgO tended to overestimate the melting temperatures. More recent molecular dynamics simulations predict the melting of the B1 phase to occur at 3100 K and 0 GPa and at 9400 K and 240 GPa (Taniuchi and Tsuchiya (2018)). Figure 1.6 summarizes the results by providing a schematic phase diagram of MgO. The boiling point of MgO is not studied often, the CRC Handbook of Chemistry and Physics (Haynes (2017)) puts it at 3870 K.

Going to the low-pressure side of the phase diagram; the liquid-vapour region, not much is known about the behaviour of MgO. This is due to the experimental difficulty of reaching these conditions, which would need to be done via laser shock-wave experiments. The difficulty comes from the requirement of the high-power laser, on the order of hundreds of GW in the case of McWilliams et al. (2012). However, we can obtain some clues on the vaporization from theoretical (Fegley et al. (2016)) and experimental studies, using molecular beam epitaxy (Vassent et al. (2000)) and a vacuum thermogravity apparatus (Jacobson et al. (2017)). These studies indicate that the vaporization of MgO is a congruent process, meaning that gas obtained from heating MgO crystals is a stoichiometric mixture of Mg and O atoms. In contrast, for a large majority of rock-forming minerals, the vaporization is incongruent meaning stoichiometry in the liquid and the vapour are not the same, examples being feldspars Kobsch and Caracas (2020) and Mg_2SiO_4 Townsend et al. (2020).

1.6.2 CaO

As expected with alkaline-oxides, CaO has a B1 to B2 phase transition at high pressures, just like MgO. CaO in its B1 phase at ambient conditions has the classical rock salt (NaCl) atomic structure. Jeanloz et al. (1979) provide a review of the discovery of the B1 to B2 phase transition in CaO. Both shock-wave experiments and x-ray diffraction analysis with a diamond anvil cell (DAC) were carried out, this gave a range for the B1-B2 transition between 60-70 GPa at room temperature. The B2 phase, just like with MgO, has a CsCl atomic structure. From this point on the phase space becomes less known. Various theoretical and experimental studies have been performed on the B1-B2 phase transition but nothing past that can be found. There seems to be no estimate for the B2 melting line like we do have for MgO.

Just like with MgO, one of the earliest works on the melting point of CaO is by Kanolt (1913). The reported melting temperature at the time was 2572 K, the modern estimate of the melting temperature lies between 2800 K and 3200 K. Manara et al. (2014) provide an overview of the reasoning behind these two very different temperatures, one group associates the higher temperatures with a liquid-vapour transition. The other group in favour of the higher melting temperatures suggests the lower melting temperatures correspond to the melting temperature of a composition that reacted with their containment. Manara et al. suggest CaO is hard to study at high temperatures in part due its optical properties in this regime. The sudden change in emissivity of CaO at temperatures between 2000 K and 2300 K, due to a phenomena called limelight of Drummond light, can make it difficult to measure stable thermograms with pyrometry. Manara et al. perform their own experiments and favor a melting temperature around 3200 K, they also estimate the boiling point at 3450

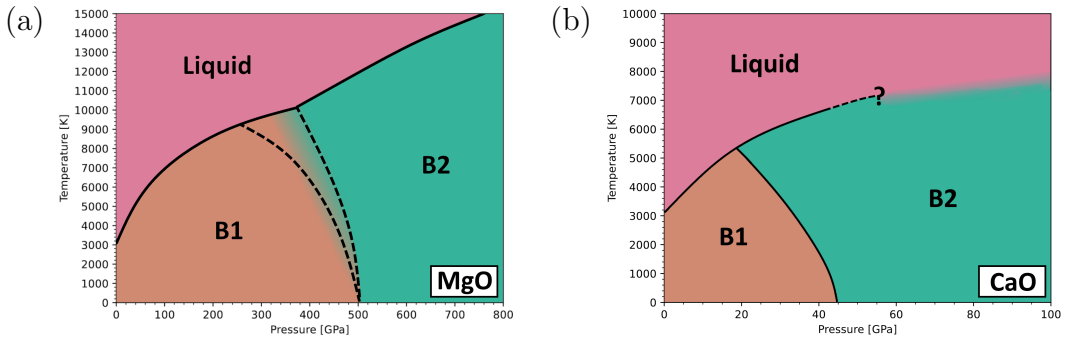


Figure 1.6: Schematic depiction of the phase diagrams of MgO (left) and CaO (right). The B1-B2 phase boundary of MgO is not very well defined at the highest temperatures and the melting line of CaO has only been determined up to 60 GPa. These phase diagrams were constructed using numerical approaches performed by other studies and should only serve as rough reference. For more details on the methodologies and accuracy refer to the original publications used to construct the phase diagrams: Soubiran and Militzer (2020) for MgO and Karki and Wentzcovitch (2003) and Sun et al. (2010) for CaO

K. This puts the melting temperature at around same value as for MgO. In contrast, the boiling point of MgO is estimated significantly higher than the one of CaO. CaO has a lot of similarities with MgO, both structurally and chemically. However, MgO has been studied more extensively in the geological context, as it is a more relevant material than CaO. Nonetheless, has the phase space of CaO been studied thoroughly, as can be seen in figure 1.6.

1.7 Research goals

The topics covered in this research focus on refractory materials. Experimentally it is possible to study the liquid-vapour space of materials using laser-induced shock waves, although this method is quite expensive and difficult to carry out. For this reason, we study MgO and CaO by applying Density Functional Theory (DFT) molecular dynamics (MD). This allows us to study the phase space in the liquid-vapour region using numerical methods. As mentioned earlier, another reason for studying MgO is because it is an end-member of the ubiquitous MgO-SiO_2 system. Using the knowledge of previously studied materials in this system and MgO, we determine how the MgO/SiO ratio influences important dynamical and structural features of the liquid/vapour system.

We also attempt to study the behaviour of ^{16}O and ^{18}O in MgO. The similarity of the oxygen isotope fractionation in the Earth-Moon system is one of the main constraints on any lunar formation model. Exploring any possible mass-dependant fractionation during the vaporization, or condensation, in the liquid-vapour phase space should contribute to future modelling attempts.

Lastly, we attempt to apply novel machine learning (ML) techniques in order to scale up our simulations. This allows us to take a much better look at processes that happen on larger time and length scales, such as vaporization and viscosity. While normally with DFT-MD our simulations are run up to a couple of tens of picoseconds, ML-MD is able to do nanosecond simulations in a fraction of the computational time. Liquid-vapor exchange plays a vital role in the canonical post-impact body, where a central liquid is allowed to exchange with the surrounding vapour body.

CHAPTER 2

Methodology

Contents

In this chapter, we will briefly go through the theoretical background of the methodology applied in this work. This can be split into two parts; the Density Functional Theory molecular dynamics part and the Machine Learning applied to classical MD section.

2.1 Molecular dynamics

The fundamentals of MD rely on Newtonian equations of motion to displace atoms. Integration of Newton's equations of motion enables us to derive thermodynamic properties, like structural and

transport behaviour. The specifics of this method come into play when discussing the exact driving forces of this displacement. The core of molecular dynamics is describing interatomic interactions. Interatomic interactions are described with potentials. Traditionally, these potentials were determined in advance, so-called empirical potentials. The interactions in a system are typically broken up into two-body, many-body, long and short-range, and (non-)electrostatic interactions. The earliest 'MD' simulations, like Alder and Wainwright (1957), consisted of hard spheres models where square well potentials described the interaction between the spheres. The most popular interatomic potential, that is still used to this day for many tests and examples, is the Lennard-Jones potential, also refered to as the 12-6 potential. First described by Lennard-Jones (1931), eq. 1.1 shows the expression of this potential, where χ is the depth of the potential well, r is the distance between the particles, and σ is the distance at which the particle-particle potential energy is zero.

$$V_{LJ}(r) = 4\chi \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{1.1}$$

We apply MD as it allows us to include a temperature component in the simulations, by way of introducing kinetic energy. This way, we are able to determine material behaviour at a given temperature. MD also enables us to study dynamics behavior, like diffusion.

For decades, empirical potentials proved very useful and a substantial amount of research was performed with them and is still done to this day. However, they have some drawbacks, notably limitations when it comes to the chemical complexity of the system and their single-application nature. Parameters of empirical potentials are optimized to reproduce some specific observable or quantity and are in general not flexible enough to accurately match multiple properties at once. Because we want to look at a large variety of thermodynamic properties we decide to apply this method.

Due to these shortcomings, new methods were developed to improve the flexibility of molecular dynamics. The main goal of these methods are to determine the interatomic interaction "on the fly". These methods can efficiently and accurately describe electrons; one of these new methods is Density Functional Theory (DFT). While a Lennard-Jones potential could give reasonable results, in this work we opt to use the more complex DFT framework which allows us to explicitly describe the electronic structure of the system and derive more realistic results.

2.2 Density Functional Theory

DFT molecular dynamics, sometimes wrongly referred to as first principle molecular dynamics, is an adaptation of classical MD. Combining the chemical accuracy of density functional theory (DFT) with the dynamical properties of MD.

Several popular frameworks exist for DFT-MD, we will however only discuss the Born-Oppenheimer framework in this work as this is the one applied here. For more details on other frameworks like the Ehrenfest or Car-Parrinello ones, I direct the reader to the book by Marx and Hutter (2009) or the original article by Car and Parrinello (1985).

In order to give an overview of the theory we start with the time-independent schrödinger equation, eq 2.2. Where H is the Hamiltonian, E is the energy of the system, and ψ the total wavefunction.

$$\hat{H}\psi(\vec{R},\vec{r}) = E\psi(\vec{R},\vec{r}) \tag{2.2}$$

The Hamiltonian is the sum of the kinetic (T) and potential (V) energy of all the particles in the system, as described in eq. 2.3. Where the kinetic energy can be divided into the contributions of the electrons, T_e, and the ions, T_i. The can be done for the potential energy, as described by Coulomb interactions between electrons, V_{e-e}, electrons and ions, V_{e-i}, and lastly ions and ions, V_{i-i}, equation 2.4. Here m_e and $-e$ denote the electron mass and charge, ∇^2 is the Laplacian, $\hbar$ the reduced Planck constant, M_I and Z_I are mass and atomic number of the Ith nucleus, $|r_i - R_I|$, $|r_i - r_j|$, $|R_I - R_J|$ are electron-ion, electron-electron, and ion-ion distances.

$$H = T + V \tag{2.3}$$

$$\begin{aligned} H = &-\frac{\hbar^2}{2m_e}\sum_i \nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I}\nabla_I^2 \\ &- \sum_{i,I} \frac{Z_I e^2}{|r_i - R_I|} \\ &+ \frac{1}{2}\sum_{i \neq j} \frac{e^2}{|r_i - r_j|} \\ &+ \frac{1}{2}\sum_{I \neq J} \frac{Z_I Z_J e^2}{|R_I - R_J|} \\ &= T_e + T_i + V_{e-i} + V_{e-e} + V_{i-i} \end{aligned} \tag{2.4}$$

2.2.1 Approximations of Schrödinger's equation

Born-Oppenheimer

One major issue with Schrödinger's equation is that it becomes prohibitively expensive to solve exactly as the number of particles in the system increases. An exact solution of the time-dependant Schrödinger equation, without approximations, is only known for the hydrogen atom. For this reason, we have to rely on some approximations or simplifications. The first important simplification comes in the form of the Born-Oppenheimer approach (Born and Oppenheimer (1927)), which postulates that the fast electronic motion can be separated from the slow nuclear motion, due to the large difference in mass ($M_I >> m_e$). In practice, this means the Hamiltonian can be divided into two parts, the electronic (eq. 2.5) and the ionic (eq. 2.6). This means that the coupled wavefunction of the system can be decoupled as the product of an electronic wavefunction and a nuclear wavefunction (eq. 2.7).

$$H_{electronic} = T_e + V_{e-e} + V_{e-i} \tag{2.5}$$

$$H_{ionic} = T_i + V_{i-i} \tag{2.6}$$

$$\psi(\vec{R}, \vec{r}) = \psi(\vec{R})\psi(\vec{r}) \tag{2.7}$$

Hohenberg-Kohn

The next crucial approximation came from Hohenberg and Kohn (1964) and the following year Kohn and Sham (1965). Hohenberg and Kohn propose that the total energy ($\varepsilon[n(r)]$) is a unique functional (function of a function) of the electron density. In practice, this looks like eq. 2.8. Where T is again the kinetic energy, E_H is the Hartree energy (analytically know part of electronic interactions), E_{ext} is the external energy, which includes electron-ion and ion-ion interactions and external field effects. The last term, E_{xc}, is the exchange-correlation energy and contains all unknown terms, it is discussed more in detail later in section 2.2.2.

$$\varepsilon[n(r)] = T[n(r)] + E_H[n(r)] + E_{ext}[n(r)] + E_{xc}[n(r)] \tag{2.8}$$

Hohenberg and Kohn also put forward the theorem that the functional that gives the lowest energy, is the functional that gives the ground state energy. This is only true if the input electron density is in its true ground state density. In practice, this means that the Hamiltonian reaches its minimum when the input density is in the ground state.

Kohn-Sham

The last important approximation to discuss came a year after Hohenberg and Kohn's, from Kohn and Sham (1965). Kohn and Sham developed a set of self-consistent equations based on the work of Hohenberg and Kohn. These equations enable us to determine the electron density self-consistently, and subsequently the energy. The Kohn-Sham set of self-consistency equations goes as follows: beginning with an assumed electron density $n(r)$, construct V_{eff}, part of the Kohn-Sham hamiltonian that is known,using eq. 2.9 and μ_{XC}, exchange correlation part of the potential, and find a new $n(r)$ with eq. 2.10 and eq. 2.11. Here E_{XC} and ϵ_i are exchange-correlation energy and Kohn-Sham eigenvalue. ψ_i Here is the wavefunction of Kohn-Sham orbital i described by a Slater determinant.

$$V_{eff} = V_{ext}(r) + \int \frac{n(r')}{|r - r'|} dr' + \mu_{XC}(n) \tag{2.9}$$

$$[-\frac{1}{2}\nabla^2 + V_{eff}]\psi_i(r) = \epsilon_i \psi_i(r) \tag{2.10}$$

$$n(r) = \sum_{i=1}^{N} |\psi_i(r)|^2 \tag{2.11}$$

Now, let us see what all of this looks like in practice. What does an actual ab initio molecular dynamical simulation look like for our case? We start out by initializing the system with the atomic positions and introduce the pseudopotentials relevant to the system. By applying Bloch's theorem we are able to describe the Kohn-Sham wavefunction as an expansion of a plane waves basis set, with coefficients (or weights). Using the pseudopotentials and atomic positions we can obtain an initial value for the plane wave coefficients in the plane waves expansion. From these coefficients, using Kohn-Sham equations, we get an initial guess for the electronic energy. Next, we make a guess for the plane wave coefficients and perform the same Kohn-Sham equations and obtain another electronic energy. We compare the initial electronic energy and the newly obtained electronic energy to get a degree of convergence. If the level of convergence is not within the desired range we choose a new set of plane wave coefficients and recalculate the electronic energy, compare

etc. This loop of energy minimization is called the self-consistency (SC) loop and continues to cycle until a desired level of convergence in the energy is achieved. The minimization algorithms applied in this work are the residual minimization method direct inversion in the iterative subspace (RMM-DIIS) (Wood and Zunger (1985) and Pulay (1980) and the Blocked-Davidson algorithm.

2.2.2 Exchange & correlation

The remaining unknown quantity is the exchange-correlation energy, E_{xc}. This energy arises from electron-electron interactions and consists of two parts as the name implies, exchange and correlation. The exchange part originates from the anti-symmetry of the wave function of a many-electron system. This anti-symmetry causes a spatial separation between electrons of the same spin, which in turn reduces the Coulomb energy of the electronic system. The correlation energy is the difference between the many-body energy of an electronic system and the energy of the system as determined with the Hartree-Fock approximation, as described by Payne et al. (1992).

It is only possible for a limited number of systems to have an exact description of the electron-electron interactions mentioned above, such as a homogeneous electron gas. For this reason, various functionals have been developed to replace this E_{xc}, the most popular of which are the Local Density Approximation (LDA) and Generalized Gradient Approximation (GGA). The LDA is an approximation of the exchange and correlation-based purely on the electron density of the coordinate where this energy is evaluated. In practice, this takes the form of eq. 2.12. In the LDA the exchange energy is underestimated, while the correlation energy is overestimated, which leads to a beneficial error cancellation. The LDA is based on a uniform electron density. On the other hand, GGA takes into account the gradient of the electron density together with the point electron eq. 2.13. Over time numerous other exchange-correlation functionals have been constructed. Examples include the meta-GGA, which includes the kinetic energy density and the second derivative of the electron density, or hybrid functionals. As we strive to reach chemical accuracy, the computational cost to implement these exchange-correlation functionals increases drastically. What is the the best functional depends on the implementation, an often referred to scheme is Jacob's ladder of exchange-correlation functionals. In our study chemical accuracy is not of the utmost importance and we need to run relatively larger systems over many picoseconds. We are also running systems at low densities, meaning lower degrees of coordination within the system. Both LDA and GGA would be reasonable choices for us in this work. We decide to use GGA as Alfè (2005) shows that this gives results closer to experimental values than LDA.

$$E_{XC}^{LDA}[n] = \int \epsilon_{XC}(n) n(r) d^3r \tag{2.12}$$

$$E_{XC}^{GGA}[n] = \int \epsilon_{XC}(n, \nabla n) n(r) d^3r \tag{2.13}$$

2.2.3 Pseudopotentials

One last tool applied to DFT is the pseudopotential. In our low pressure systems, core electrons do not contribute to the chemical behaviour of atoms. At higher pressures different pseudopotentials should be used that treat some of the core electrons as valence electrons. What this pressure is, depends on the material and should be verified by running simulations with different potentials. For our range of pressure we did a quick check with a 'harder' (more valence electrons) potential

and found similar results as the ones we obtained with the standard potential. By simplifying the description of the electron wavefunction close to the atomic centre we can decrease the computational time drastically. A pseudopotential contains a simplified description of the core electrons while maintaining an accurate description of the valence electrons at a distance r onwards. The core electrons are of course still taken into account when determining the total energy and charge distribution.

All combined, these theorems and approximations form the fundamentals for our DFT-MD simulations. Applying them in order to perform calculations is done using a Kohn-Sham scheme.

2.2.4 Simulation details

Our first-principle molecular dynamics simulations were performed using the Vienne Ab initio Software Package (VASP) (Kresse and Furthmüller (1996a), Kresse and Furthmüller (1996b), Kresse and Hafner (1993), and Kresse and Joubert (1999). The projector augmented-waveform (PAW) (Blöchl (1994)) interpretation of DFT was applied here. PAW apply harmonic oscillator plane waves to sample the valence electrons close to the nuclei, where the normal plane waves would not be sufficient. Calculations were done with a version of the GGA functional, the Perdew-Burk-Enzerhof (PBE) (Perdew et al. (1996)). The cut-off energy used was different per system, for MgO and $MgSiO_3$ we used 550 eV and for CaO we used 600 eV and in all cases we use only the gamma point to sample the k-space. All simulations were done as a canonical ensemble (NVT), meaning a constant number of atoms, volume, and temperature. Temperature is regulated using a Nosé-Hoover thermostat (Nosé (1984)). A thermostat regulates the temperature by having the system of interest interacting with a heat bath. In the Nosé-Hoover this heat bath is treated as an extra degree of freedom. The heat bath is described as a single fictitious particle that is constrained. In VASP this constrain is placed using a specific mass, the Nosé-mass. The VASP manual states that the Nosé-mass should be set such that the temperature fluctuations have approximately the same frequencies as the typical phonon-frequencies. In practice, this means we used a Nosé-mass of 5 for MgO and $MgSiO_3$, which was not adapted with density. Later on, for CaO and some MgO simulations, we opt to use an automatically updating Nosé-mass which adapts to the density, as this is more widely used in the literature. For the level of energy convergence, we use 10^{-4} eV. For the general DFT-MD simulations, we apply Fermi-Dirac smearing to describe the partial occupancies for each orbital. The smearing width is adapted to the temperature of the simulation. The number of available electronic bands is increased as necessary, and a buffer of around 50 bands at 0 occupancy, in shown digits, is applied to all simulations. In order to give the reader a sense of what the energy, temperature, and pressure behaviour looks like in a simulation, we include their progression over time for a representative simulation in figure 2.1.

2.2.5 Convergence testing and error estimation

Like most modelling methods there is a certain trade-off between accuracy and (computational) expense. In order to determine what level of accuracy we can reach with reasonable computational usage, we perform convergence tests on various input parameters.

The most commonly described, and one of the most important, convergence tests is the convergence with respect to the cut-off energy. This energy describes the limit of the kinetic energy of the plane waves in the plane wave basis set expansion. Starting out, we compare cutoff energies for two MgO systems, one at ambient density (3.59 g/cm^3) and one at 0.52 g/cm^3, both at 5000

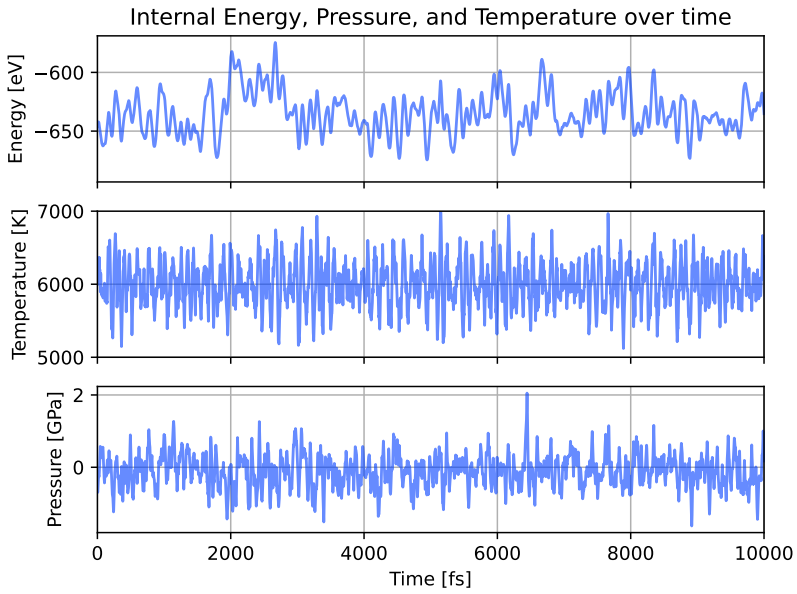


Figure 2.1: Internal energy, pressure, and temperature as a function of time for a representative simulation. The system used for this simulation is at 6000K and 0.90 g/cm^3.

K. We perform static calculations between cutoff energies ranging from 200 to 1400 eV, with steps of 100 eV. We compare the total energies for all cut-off energies with the most accurate test, 1400 eV. The normalized results are shown in figure 2.2. We see that the very lowest energy cutoffs give a very different (normalized) total energy and we see a great level of convergence starting from 900 eV. Related to this we look at the effect of the cut-off energy for the plane waves in the PAW part. Results show us that it has very little effect on the overall total energy.

Because we run our MgO simulations at 550 eV (due to computational time limitations) but convergence is reached at 900 eV we look at the difference in accuracy between these two cut-offs. For this look at the effect of the cut-off energy on the pressure, starting out by looking at a single system in detail. The MgO system of 0.52 g/cm^3 at 8000 K is split into 20 even parts and each part is analyzed at cut-off energies of 550 and 900 eV. Results from this test are shown in figure 2.3. We find that the difference (for static calculations) in the pressure between 550 and 900 eV is more or less constant. Meaning that for a single system, the effect of the compromise of 550 eV cut-off on the pressure can be accounted for with a single value. Knowing this we, calculate this pressure difference for all systems analyzed. These results are shown in table 2.1. The correction factor of $\pm 1\%$ we see in fig 2.3 is varied depending on the density and temperature, but hovers around 0.5-2%. We did this check for MgO, we assume other studied materials here like CaO and $MgSiO_3$ behave similar. Although it is good to know that these pressure differences exist, we find that they have little to no effect on the final results of the simulation, nor on any of the conclusions made.

There are a couple different ways to go about calculating the statistical error in the simulations.

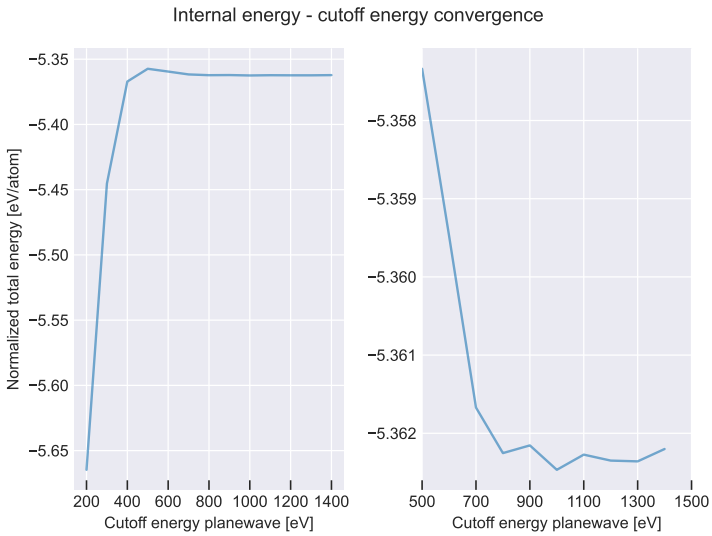


Figure 2.2: Convergence tests of the normalized total energy of a system at 5000 K and 3.59 g/cm^3. On the scale of all tests, convergence appears to be reached around 500 eV. Taking a closer look at only the higher cutoff values we can see that a convergence of 1 meV/atom is reached after 900 eV.

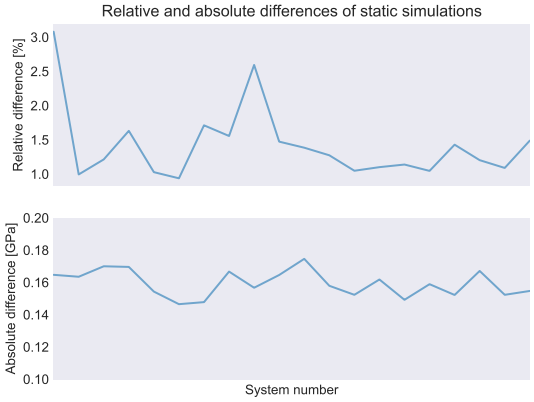


Figure 2.3: Results from 20 static calculations along a single MD simulation. The upper graph shows the % difference comparing a cut-off energy of 550 with 900 eV, the bottom graph shows this difference in absolute values (in GPa). Tests are performed on systems at 8000 K and 0.52 g/cm^3

	Pressure [GPa]						
Density [g/cm^3]	**4000 K**	**5000 K**	**6000 K**	**7000 K**	**8000 K**	**9000 K**	**10000 K**
2.63	0.09349	0.08976	0.0968	0.0891	0.0862	0.08670	0.03940
1.76	0.06077	0.05702	0.06178	0.06229	0.05949	0.05862	0.05410
1.24	0.04243	0.04147	0.04725	0.03907	0.03674	0.04135	0.04053
0.90	0.03053	0.03046	0.02908	0.02950	0.02591	0.02550	0.02529
0.78	0.02642	0.02734	0.02493	0.02452	0.02139	0.02246	0.02322
0.68		0.02223	0.02202	0.02090	0.01985	0.02128	0.02191
0.59		0.01985	0.01859	0.01852	0.01735	0.01710	0.01709
0.52			0.01753	0.015571	0.01547	0.01765	0.014098

Table 2.1: Absolute differences in pressure (in GPa) between static calculations of 550 and 900 eV cut-off energy

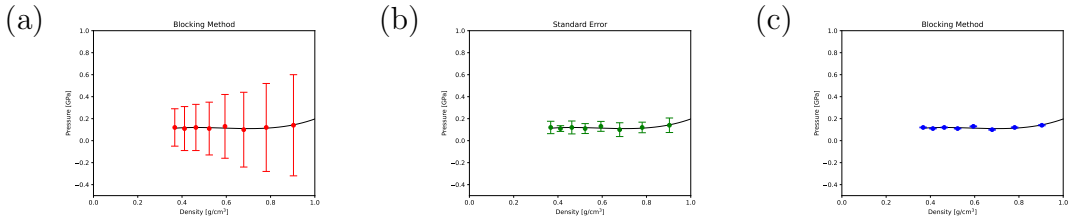


Figure 2.4: Examples of different approaches to calculate the statistical error. From left to right: standard deviation, standard error, and blocking method. Statistical errors here are shown with respect to average values obtained along a 6500 K isotherm of MgO.

Here we will present three: standard deviation, standard error, and a blocking method. In figure x we show the magnitude of these various methods of error estimation for an average MD simulation. The standard deviation is calculated using eq. 2.14, where x is the value (in this case pressure) and n is the number of samples. For the standard error we propose the following; we use the decorrelation time of the stress in the system as measurement for each sample length. So, if we have a simulation of 10 ps and a decorrelation time of 200 fs, this means that we have 50 individual samples (n^*) in the simulation. We then calculate the average pressure of each sample and calculated the standard deviation over those averages ($\sigma^*_{std.dev.}$) and use eq. 2.15 to calculate a standard error. The last method, the blocking method, was proposed in by Flyvbjerg and Petersen 1989 and is a renormalization group method especially designed for correlated data. We recommend reading the source material or Dr. Anaïs Kobsch's manuscript for its implementation in the UMD package.

$$\sigma_{std.dev.} = \sqrt{\frac{\sum(x_i - x_{avg})^2}{n-1}} \tag{2.14}$$

$$\sigma_{std.err.} = \frac{\sigma^*_{std.dev.}}{\sqrt{n^*}} \tag{2.15}$$

The Nosé-mass is estimated based on the fundamental phonon frequencies of the material. In the case of MgO, we perform a set of simulations with Nosé-masses of 0, 2, 4, 8, and 16. From these tests, we estimate a Nosé-mass of 5 to make the Nosé-frequency match the frequency of

dominant transverse optical (TO) and transverse acoustic (TA) phonons. The actual effect of the Nosé-mass is difficult to see, as it only has influence on the dynamic behavior of the system, not the thermodynamic averages. Later on during this project, and after feedback from a referee, we adapted the Nosé-mass to be set to 0 in VASP. This means that the software automatically chooses a Nosé-mass to keep the period constant at 40 time steps. After this conclusion, we went back to our MgO simulations and made sure the periods for each system were not too far off with a Nosé-mass of 5, which was the case. An unsuitable Nosé-mass can lead to decoupling of the thermostat, we do not see this in any of our simulations.

System size plays another important role in the simulation in many ways, the finite-size effect can be seen in many of the observables. We run tests for systems of sizes 64, 216, and 512 atoms. The resulting pressures obtained from these simulations are all within the required accuracy. Due to interest in speciation analysis on the one hand and computational limitations on the other, we opt to run with systems of 216 atoms for MgO and CaO, and 320 atoms for $MgSiO_3$. System size has a big effect on the temperature of the system over time, decreasing system size leads to higher deviation from the instantaneous temperature. In our simulations we see the instantaneous temperature deviate about 10% from the input temperature, the average temperature over the entire simulation does match the temperature set with the thermostat. Because we are working with correlated systems, error calculations can be done by using the blocking method (Flyvbjerg and Petersen (1989)).

In order to cover our bases, we also looked at possible London dispersion effects (also known as van der Waals forces) on extreme vaporous systems. For this, we looked at five systems of 0.11 g/cm^3 and 5000 K and compared them with five systems of the same starting configuration with a Van der Waals corrections enabled. We chose this density as the London dispersion effects, if there are any, should be the most noticeable here. For Van der Waals correction we applied the DFT-D3 method from Grimme et al. (2010). We run these simulations for only 150 steps as this region is prohibitively expensive to sample and we do not expect any significant influence from the correction. This implies that the results are not all-conclusive and only serve as a notion of whether it is worthwhile to investigate further. With this disclaimer out of the way, let us first discuss the average pressures obtained. All of the pressures fall very close to 0 GPa as expected, overall the simulations without Van der Waals corrections have increased pressures. However, the difference between the pressures obtained falls easily within the standard deviation which is comparatively huge. We are talking here about pressures ranging from -0.002 to 0.011 GPa for all systems while the standard deviation hovers around 0.063 GPa. A more interesting statistic to look at might be the polymer speciation, in order to look at the distribution of the vaporous species. Looking at O, O_2, Mg, and MgO and comparing the percentage contribution to the total lifetime, there does not appear to be any significant correlation between the Van der Waals correction and vapour species. The only thing that is consistent is in the simulations without Van der Waals correction MgO is slightly more present across all five systems, but this is only a difference of about two per cent on average. One interesting feature though is that the number of vapour species that pass a thresh hold of 5% is slightly higher for the simulations with Van der Waals correction. Sadly, there does not seem to be any specific kind of species that is more common due to this.

2.3 Machine learning potentials

As artificial intelligence and machine learning have become the new buzzwords in most fields of research and industry, so has molecular dynamics started applying these methodologies to increase

its performance. Recent developments in ML have exponentially increased interest and new ML schemes seem to be published every year. In this section, I will briefly give a general introduction to ML potentials, and then I will talk more in detail about the ML potential I have worked with. Namely, the Gaussian Approximation Potential (GAP). This is a very exciting field, and will most likely define the future of DFT-MD for some time to come.

2.3.1 Fundamentals

Although recent industry development may lead you to believe that ML is a quite new development, the contrary is true. Due to recent leaps in computational power and the vast increase in data, ML has been put at the forefront of many industries and sciences. ML is considered a part of the larger field of Artificial Intelligence (AI). In AI, the system perceives its environment and is able to make decisions based on the information it receives, in order to achieve its goal. ML on the other hand relies on past observations to make predictions about the future, there is no aspect of decision-making. ML is considered an interpolation method, as future predictions are made based on past data it is difficult to make predictions outside of that data range. The main reason for applying ML in the field of MD is that it allows great decreases in the computational cost involved with running simulations. This allows us to perform simulations with system sizes of two magnitudes higher than before, within a reasonable time frame. Besides system size we can of course also choose to run a modest system size over a much longer period than we could before within a reasonable amount of time, meaning that simulations of a nanosecond are easily within our grasp.

2.3.2 Gaussian Process Regression

Before discussing the GAP, let us discuss the underlying mathematical concept that is the Gaussian Process Regression (GPR). GPR is a form of non-linear regression and it is the inference of continuous values based on a Gaussian process prior. The Gaussian process is a generalisation of multi-variate Gaussian distribution to infinitely many variables. In other words, it is a collection of random variables, any finite number of which have consistent Gaussian distributions. Whereas a multivariate Gaussian distribution can be fully described by a mean vector and covariance matrix σ. A Gaussian process is described by a mean function $m(x)$ and covariance function $K(x, x')$, this last term is also often referred to as a kernel function.

2.3.3 Machine learning applied to molecular dynamics

Running MD simulations gives us a tremendous amount of data, so it was only a matter of time before ML would be implemented, one way or another. One of the earliest mentions of applying ML is in the paper of Blank et al. (1995) where they use a neural network (NN), a sub-field of ML, to model global properties of potential energy surfaces. This was later further developed by Behler and Parrinello (2007). The ML method applied in this work is based on Gaussian Process Regression (GPR) and was first applied to modelling a potential energy landscape by Bartók et al. (2010). Whereas a neural network relies on weights attributed to each node, GPR relies on hyperparameters. The choice of covariance function has a big influence on the outcome of your GPR. As described by Bartók et al. (2018) (and in more detail in Bartók and Csányi (2015)) the total GAP model energy for a system is a sum of a predefined pair potential and a many-body term given by the linear sum over kernel basis functions, eq. 3.16.

$$E = \sum_{i<j} V^{(2)}(r_{ij}) + \sum_{i} \sum_{s}^{M} \alpha_s K(\mathcal{R}_i, \mathcal{R}_s) \tag{3.16}$$

Here i and j range over the number of atoms in the system, $V^{(2)}$ is the pair potential, r_{ij} is the distance between atoms i and j, $K(\mathcal{R}_i, \mathcal{R}_s)$ is the covariance between the environment of atom i and the environments of the representative atoms s in the database. The last summation runs over a set of M representative atoms, so s is a selected atom of a relevant neighbourhood. The algorithm used to select a set of representative neighbourhoods is done with a matrix reconstruction technique called CUR matrix decomposition (Mahoney and Drineas (2009)). I will leave it to any invested reader to examine the source material themselves, the important thing to know is that this decomposition leads to a low-rank approximation of the full kernel matrix while using only a part of its columns and rows.

As mentioned earlier, the choice of covariance function plays a big role in GPR. In the case of GAP, they make use of a so-called Smooth Overlap of Atomic Positions (SOAP) (figure 2.5). They describe a SOAP kernel with a SOAP descriptor. The neighbourhood, $\mathcal{R}_i$, is described by a neighbour density, ρ_i, according to eq. 3.17.

$$\rho_i(r) = \sum_{i'} f_{cut}(r_{ii'} e^{-(r-r_{ii'})/2\sigma^2_{atom}} \tag{3.17}$$

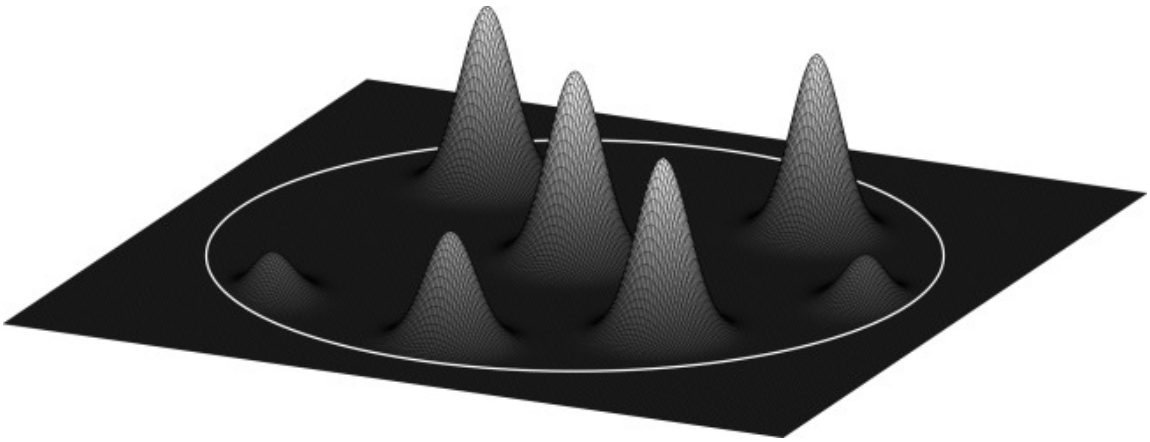


Figure 2.5: Two-dimensional illustration of the atomic neighbour density as described by the Smooth Overlap of Atomic Positions (SOAP). The white circle represents the cut-off radius, r_{cut}. From Bartók and Csányi (2015).

Where i' are then neighbours of i, this includes i itself, f_{cut} is a cutoff function that smoothly goes to zero beyond a cut-off radius (it sets a limit of the neighbour environment considered) r_{cut}, and σ_{atom} is a smearing parameter. This neighbour density can be expanded in a basis of spherical harmonics and radial functions which allows the method to construct a spherical power spectrum which is uniquely rationally and permutationally invariant, like a descriptor should be. The power of the GPR is that in order to make a prediction, all you require are observations and the covariance

function. This method is inherently interpolating in nature, but we have had some successes doing extrapolation with respect to temperature.

2.4 Details on post-processing

All post-processing was done using the Universal Molecular Dynamics (UMD) package produced by **Caracas**. UMD is a python-based software package that attempts to convert the output of various ab initio software, currently supporting VASP, into a single format output. It then offers various post-processing methods to apply to the universal output. Some of the more note-worthy post-processing include determining diffusion coefficients and speciation.

The natural progression when post-processing is parsing the VASP output (OUTCAR) to the UMD format. One can then decide to calculate a pair distribution function, which in turn can be used for determining speciation. Some of the possibilities can be seen in figure 2.7.

2.4.1 Constructing the critical point

Let us first discuss how we determine the critical point and how we construct the isotherms and define spinodal points. Figure 2.6 shows an example of what the phase stability field of material might look like in the liquid-vapour region. Solid lines describe isotherms and we can see that some of the isotherms reach negative pressures as the density of the system is decreased, which would indicate some hydrostatic tension in the system. Classical nucleation theory (Karthika et al. (2016)) states that a first-order phase transition, like liquid to vapour, is regulated by an energy barrier which is dependent on the surface energy of the interface between the two phases. In our case, a system with negative pressures finds itself unable to overcome this energy barrier, due to the interfacial effect. Therefore, this system will be thermodynamically metastable but mechanically stable. Once the system passes the minimum, which marks the liquid spinodal point (black crosses), it increases again until it reaches a maximum, the gas spinodal point (grey crosses). All densities between the two spinodal points of an isotherm describe the region where a two-phase mixture coexists. What this implies is that if we follow an isotherm and we do not encounter any maxima nor any minima (i.e. spinodal points), we do not encounter the first-order phase transition at that temperature. What this implies is that this isotherm is above the critical temperature, and the system is continuously in a supercritical fluid state. Knowing this, we can make our way downwards to lower temperature isotherms until we find these minima and maxima again. At that point, we will have obtained a range for the critical temperature. Due to technical limitations, it is not justifiable to determine the gas spinodal points far below the critical temperature. Theoretically it is possible to determine the separation line between the metastable phases and the single states, these lines are the binodal lines (also refered to as coexistense curves) and are similar to the ones we would expect to find on a standard phase diagram. A Maxwell construction is shown in figure 2.6 with areas containing diagonal lines. The difficulty of performing this Maxwell construction lies in the fact that you need to accurately describe your isotherm up to very low densities. We determine the critical point for all of our materials studied as this point is a keystone feature within the liquid-vapour region. The critical point acts like a reference point, by constructing it we are able to determine in what part of phase space each of our configurations are located.

As the choice of fit for the isotherm can have a significant effect on the spinodal points and curve constructed from it, we employ a cubic function to fit the pressure-density points along any given isotherm. The choice of a cubic function was made on the basis of both the van der Waals

equation of state and the Maxwell construction having a cubic shape. This method has been applied successfully in other studies (Li et al. (2020), Vasisht et al. (2011), Townsend et al. (2020), Kobsch and Caracas (2020), and Faussurier et al. (2009)).

2.4.2 Pair distribution function

The pair distribution function (PDF) or pair correlation function (PCF), oftentimes just referred to as $g(r)$, is an important part of the post-processing. The PDF contains a lot of data on the structure of the system, it is calculated using eq. 4.18.

$$g_{ab}(r) = \frac{N_a - 1}{4\pi\rho^2} \langle \delta(r - r_{ij}) \rangle \tag{4.18}$$

2.4.3 Mean-squared displacement

One of the dynamical properties we can analyze for our systems, is the mean-squared displacement (MSD). The MSD is the square of the average distance an atom or cluster of atoms travels as a function of time. We are particularly interested in the MSD as it can be used to determine the self-diffusion coefficient. We calculate the MSD using eq 4.19. Here, N_α is the number of atoms of type α, and τ is the time window.

$$MSD(\tau)_\alpha = \left\langle \frac{1}{N_\alpha} \sum_{i=1}^{N_\alpha} (r_{\alpha,i}(\tau + t) - r_{\alpha,i}(t))^2 \right\rangle \tag{4.19}$$

Due to the correlative nature of the equation, we can calculate an MSD up until half the simulation length. As our simulations vary from 5 to 20 ps the MSDs are calculated up to 2.5 to 10 ps. Because all of our simulations are in either a pure liquid or liquid with a vapour phase the MSD continuously increases. For solids, this would not be the case. A plateau would form where the MSD stays constant. The MSD contains a ballistic part at the very start, described by a small curvature at the outset of the MSD. After the ballistic part comes the diffusion part, which is the part we will use later on to determine the slope.

Using the MSD, we can determine the self-diffusion coefficient using eq. 4.20. Where n is equal to 2,4 or 6 for one, two or three dimensions, respectively. Self-diffusion coefficients were also determined using the velocity auto-correlation function (VAF) which is calculated using eq. 4.21, where $v_\alpha(t)$ is the velocity of a particle type α at time t. From this, we can get the diffusion using eq. 4.22.

$$D_\alpha = \lim_{\tau \to \infty} \frac{1}{nt} MSD_\alpha \tag{4.20}$$

$$\phi_\alpha(\tau) = \frac{\langle v_\alpha(0) v_\alpha(\tau) \rangle}{\langle v_\alpha(0) v_\alpha(0) \rangle} \tag{4.21}$$

$$D_\alpha = \lim_{t \to \infty} \frac{1}{3} \int_0^t d\tau \, \langle v_\alpha(0) v_\alpha(\tau) \rangle \tag{4.22}$$

We will not go into all of the post-processing methods in detail, one can refer to the source material if so inclined. However, one result we want to go more into detail about is the speciation,

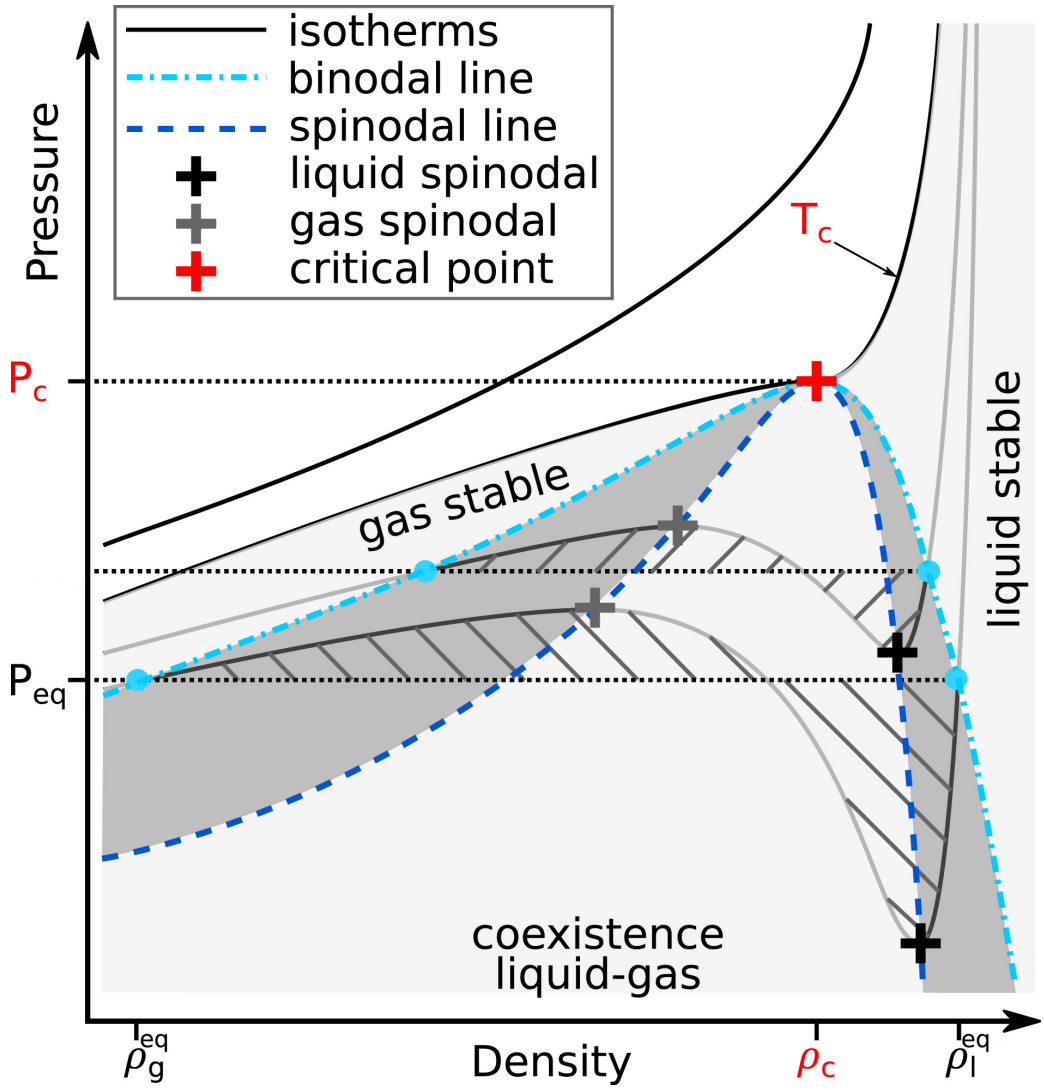


Figure 2.6: Determining the critical point and the phase stability fields using the constructed isotherms. The minima and maxima found along an isotherm describe the liquid and gas spinodal points, respectively. The spinodal points can be connected with a spinodal line. A Maxwell construction is able to determine the binodal points, which can be connected with a binodal line. Both the spinodal and the binodal lines culminate into the critical point at the top of the liquid vapour dome. The areas between the spinodal and binodal lines are metastable regions. The area within the spinodal line describes the liquid-vapour coexistence region. As higher temperature isotherms are constructed they will lose the maximum and minimum, which means the material is in a supercritical state at said temperature. Figure adapted from Kobsch and Caracas (2020)

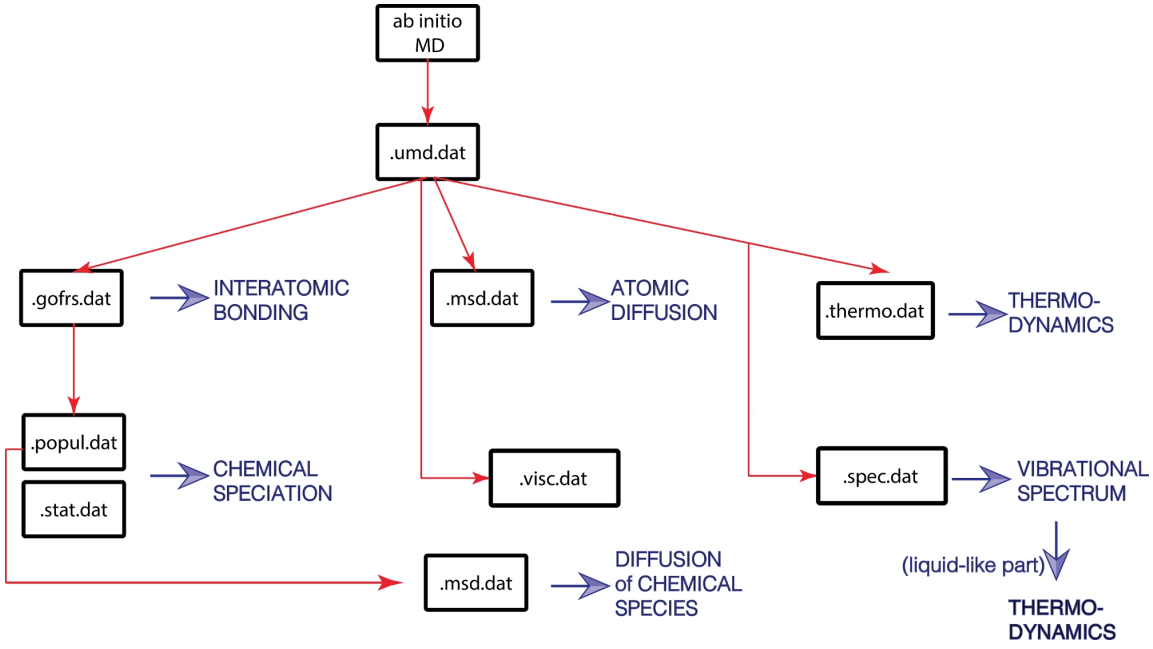


Figure 2.7: Post-processing modules from the Universal Molecular Dynamics (UMD) package. Red arrows indicate python scripts are run and blue arrows show physical properties that are output. Adapted from Caracas et al. (2021).

as these results are often of great interest to geo-chemists and they speak to the imagination of most non-physicists. The basic principle of the method is determining an average bond length between atom types by looking at the pair distribution function. The first minimum, after the first peak, describes the maximum radius of the first coordination sphere. One can consider that atom pairs, of the types examined in the pair distribution function, are 'bonded' with each other if they fall within this radius. We can use this method to create a connectivity matrix for every step in our simulation. With this connectivity matrix we can determine the distribution of species in the system. The speciation script does this calculation for the systems you give it and for the atom pairs you want to look at. There is also an option to look at the degree of polymerization of the system. Here the script does not just look at every single atom and its surroundings, but it can connect a string of atoms as long as the distance between them is within the maximum radius. Figure 2.8 shows an illustration of this, where the first coordination sphere is the red circles and the polymerization method is the blue connecting line. This method is particularly useful when working with liquid-vapour systems, where all of the liquid will be included in the polymer, giving one big species. The species left (in figure 2.8 the one with the green sphere) are the vapour and can be easily examined this way.

This sounds all well and good but the fact we want to highlight is that this is referring purely to geometry, it says nothing directly about chemistry. Although they are very closely linked, this is a very important distinction to make. Of course, when we see two oxygen atoms at a distance of some 1.8 Åin a vacuum, we would be right to assume they form an oxygen molecule together. But in some cases disregarding this distinction can lead to some flawed conclusions. One can have

for example MgO and look at the pair distribution function of oxygen-oxygen and note two peaks, one related to molecular oxygen, and one related to double the Mg-O bond distance and the bond angles. Of course, this second peak is not related to any actual oxygen-oxygen bonding, but if the first peak is not there it is possible one confuses it for some sort of bonding and feeds it into the speciation script. The script does not know any better and will happily spit out new results with this characteristic bond length. In general, we think the results obtained from the speciation script are still useful indications of chemical speciation, assuming the choice of bond length within the pair distribution function is correct. One should not hold these output values as a ground truth but merely as indications and for comparison between themselves.

One second important thing we want to emphasize is that the speciation seems to be quite dependent on the configuration of the system. It is not uncommon for a system to fall into a certain configuration, examples being slab or cylindrical vapour spaces in the liquid, and to maintain this configuration for extended periods of time. The configuration of the system has a non-negligible effect on the outcome speciation. For this reason, one can run two identical systems (number of atoms, temperature, and volume) and obtain variations in the speciation between them. If one wants to get a more accurate speciation, one should run, for example, four simulations of 5 ps with different starting configurations, instead of a single 20 ps simulation. We go into more detail regarding this topic in section 4.2.2.

The last topic in post-processing we briefly want to mention is the Bader charge analysis (Bader et al. (1979)) as formalized by the Henkelman group (Henkelman et al. (2006), Tang et al. (2009), Sanville et al. (2007), and Yu and Trinkle (2011). The Bader charge analysis is a method to quantify the charge of each atom in the system. This is especially useful when studying reducing and/or oxidizing reactions. This method works by dividing the system into so-called Bader regions. the surfaces of the Bader regions are defined by the minima of the charge density. By integrating the charge density in the Bader regions one can get a sense of the total atomic charge for each atom.

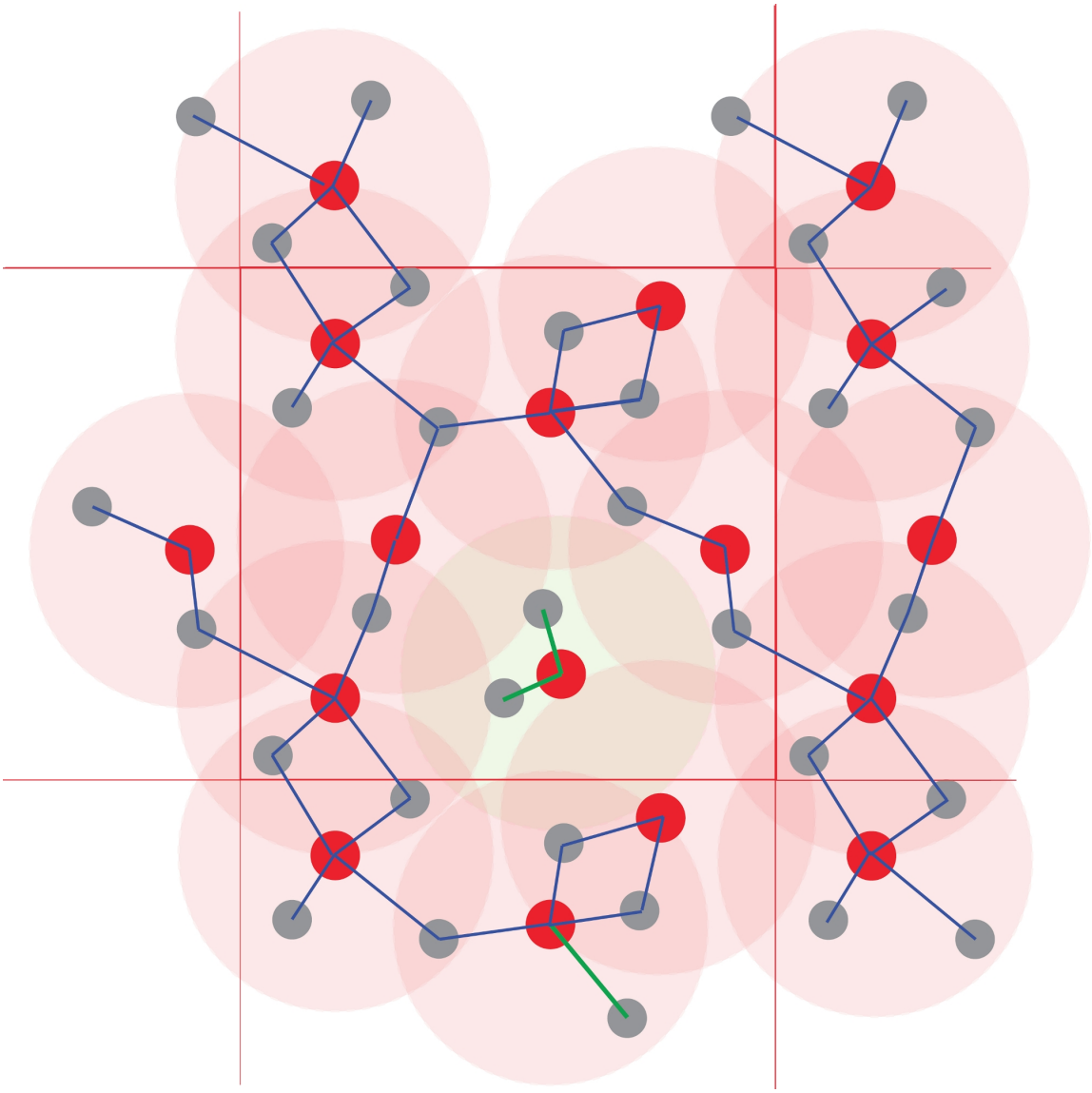


Figure 2.8: Illustration of coordination spheres in a system of red and grey atoms. The red and green spheres indicate the regions where speciation is calculated if polymerization is disabled. The blue line shows what the polymerization would look like, atoms overlapped by two spheres act as a connecting bridge. This polarization method shows clearly which clusters are not part of the main system, e.g. vapour species, in this case the cluster within the green circle.

CHAPTER 3

MgO Density Functional Theory Molecular Dynamics

Contents

The first material researched during this project is magnesium oxide, MgO. MgO is a fundamental material of telluric planets and an important part of a large variety of silicate minerals. It adds up to roughly 39% (Ringwood (1991)) of the Earth's and Moon's compositions. Besides the planetary scientific interest, MgO is an archetype for the numerous ionic AB diatomic compounds. It is a material of significant geological and technological importance.

For these reasons, it has proven to be an excellent starting point for the project. Naturally, MgO has been researched extensively both experimentally and numerically over a large range of pressures and temperatures, as discussed in section 1.6.1.

In the following chapter, we will discuss all of the results obtained for MgO; we will describe how they are obtained and what information we can obtain from them.

3.1 The critical point

In order to put all of the coming results into a framework we want to start by discussing the phase diagram of MgO. One of the main goals of this part of the manuscript is to determine the critical point of MgO. The critical point is especially interesting as it describes the top of the liquid vapour

dome. Although it is quite difficult to accurately describe the exact boundary between a liquid and liquid-vapour co-existence, we rely on spinodal points to construct the stable part of the liquid-vapour dome. In order to construct this phase diagram we run simulations ranging from 4000 K to 10000 K, making sure we are comfortably above the melting line and ensuring we capture the critical temperature.

Figure 3.1 shows the isotherm we construct with our MgO simulations. Looking at the isotherms we are able to define liquid spinodal points for temperatures between 4000 and 6500 K. We obtain gas spinodal points merely at 6000 and 6500 K, due to the technical limitations of obtaining gas spinodals. At lower densities the simulations become computationally more expensive as the electrons occupy more bands, so more bands have to be included in the calculation. This gives us an estimate for the critical temperature between 6500 and 7000 K. Which makes it higher than the critical temperatures found for SiO_2, $MgSiO_3$, Mg_2SiO_4, and K and Na-feldspar (Green et al. (2018), Xiao and Stixrude (2018), Townsend et al. (2020), and Kobsch and Caracas (2020)). Keeping in the mind the statistical errors with respect to the pressures it is not unthinkable to have spinodal points along the 7000 K isotherm. However, with the fit we present here this is not the case. All of the critical densities of the aforementioned materials lie below 1 g/cm^3 and MgO is no exception on that front. We estimate the critical density between 0.45 and 0.6 g/cm^3. As we approach the critical point the relative fluctuations in pressure with respect to density increase significantly, meaning it is quite challenging to pinpoint the critical point exactly.

3.2 Structure analysis

3.2.1 Pair distribution function

Figure 3.3 shows the PDF calculated for three isotherms at the whole range of densities, 3.4 also shows the PDF but at three densities over all the iostherms. Starting out with Mg-O the start of the PDF stays more or less the same for all temperatures and densities. The exclusion radius, the distance to which repulsive forces are too great for two atoms to be next to each other, stays around 1.5 Å. The location of the first peak occurs at around 1.90 Å at 4000 K, decreasing ever-so-slightly to 1.85 Å at 10000 K, which is a negligible amount. In our simulations the density seems to have very little effect on the peak location. A bigger difference between both the temperatures and the densities can be found for the location of the first minimum, which we use to define the limit of the first coordination sphere. At 4000 K, the location of the minimum is 2.97 Å for 3.29 g/cm^3 and 3.06 Å for 0.78 g/cm^3. At the other extreme of the isotherms, 10000 K, the minimum of 3.29 g/cm^3 is located at 3.00 Å. For the lowest density of this isotherm at 0.52 g/cm^3 it becomes a bit trickier to define a minimum, as the PDF seems to flatten out completely after the first peak. We can estimate it to be at least further than the higher density ones, somewhere past 3.50 Å. We summarize both the density and the temperature effects on the peak location and first minimum in figure 3.2. Solomatova and Caracas (2021) show that even at pressures at the base of the magma oceans the their Si-O bond distance is reduced by no more than 10 %. Looking at all the PDF on the whole we can say that the first peak tends to increase in asymmetry and width with increasing temperatures and decreasing density.

Now, for the O-O PDF, we can see a peak in the 6000 K and 10000 K isotherms that we can barely see in the PDF of 4000 K. This peak, with its location around 1.5 Å, is related to molecular oxygen. Knowing this, it is not surprising that it is more common at higher temperatures and lower densities. Molecular oxygen has been found in all studies of silicates and oxides and it is a major

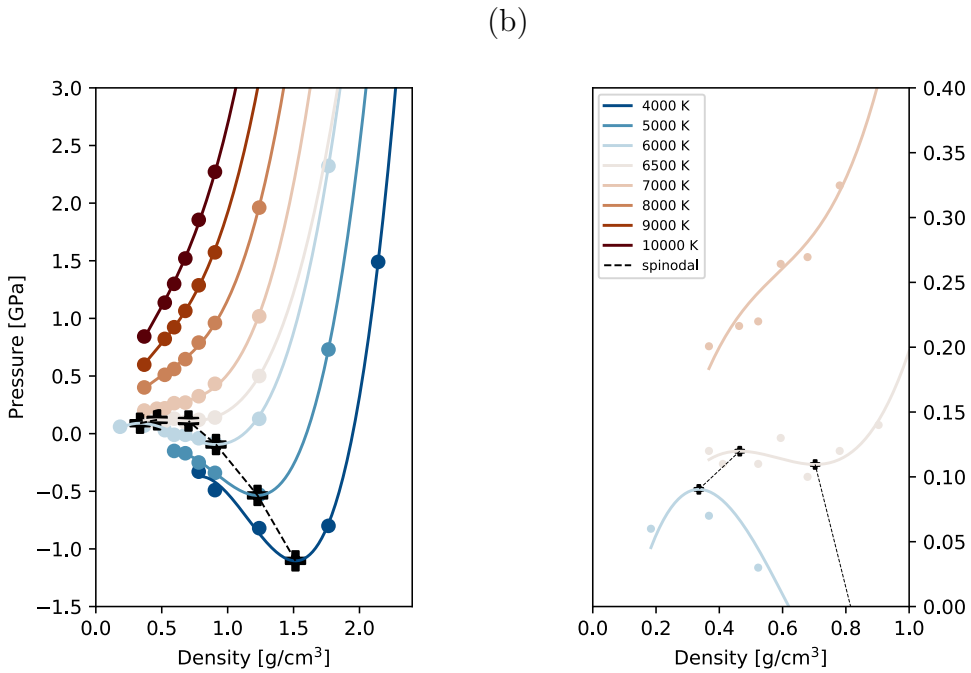
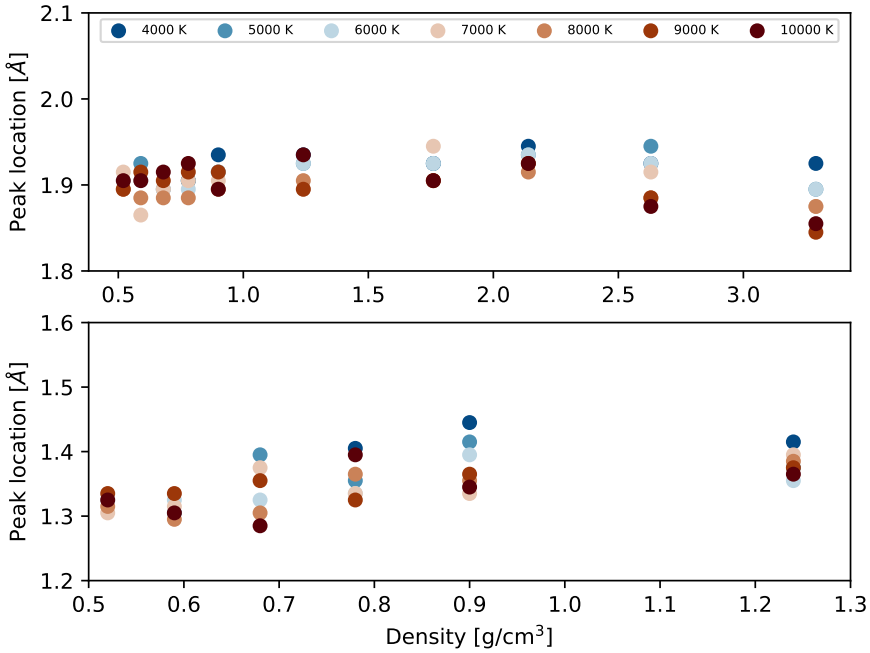


Figure 3.1: Variation of the pressure as a function of density for various isotherms (a). The solid lines are cubic function fits. Their local minima and maxima yield respectively the liquid and the gas spinodal points, represented with thick black crosses. The spinodal lines are represented with thin dashed black lines. The critical point lies between the liquid and gas spinodals in density, and between the last isotherm where the pressure still reveals local extremes, and the first isotherm where the pressure is monotonously decreasing the decreasing density. For MgO, this places the critical point in the 0.45-0.6 g/cm^3 in the density range and between 6500 K and 7000 K in the temperature range. The corresponding pressures are on the order of 0.1 - 0.2 GPa (b).

part of the vapour phase. The exclusion radius for the oxygen molecule in this system is around 1.0 to 1.1 Å, increasing slightly at higher densities. The first minimum at 6000 K and 2.63 g/cm^3 is located around 1.8 Å, for this density it is difficult to pinpoint exactly. At the other side of the spectrum, at 0.52 g/cm^3, the minimum is found at 1.85 Å. These values are nearly identical for 10000 K. One final thing to note it that in 3.4f we note that the height of the first peak is lower at 10000 K than at 9000 K, which does not follow the expected trend. This change could be related to a decrease in molecular oxygen and an increase in isolated oxygen atoms as this supercritical fluid gets gradually more ionized at these temperatures.

In both the Mg-O and O-O PDFs we note that a second peak, which describes the second coordination sphere, is present. In both cases, it describes the distance between the atomic pairs in a bulk (fluid) setting. In both pairs we can a shift in the location of the second peak, the tendency is to increase the location with decreasing density. This makes sense intuitively; as the liquid is given more space it expands and this allows the distance to increase. Because we use periodic boundary conditions, the range of the PDF is limited to half of the cell length. The integral of the pair distribution stays constant. So as the cell size is increased to decrease the density, the peak height will increase to maintain this integral. This also explains the interesting feature of the second peak in the O-O PDFs. Where we can see a decrease in the height followed by an increase again, all the while the peak moves to the right. This reduction in peak height is caused by an increase in width, the following increase is caused by the aforementioned effect.

(a)



(b)

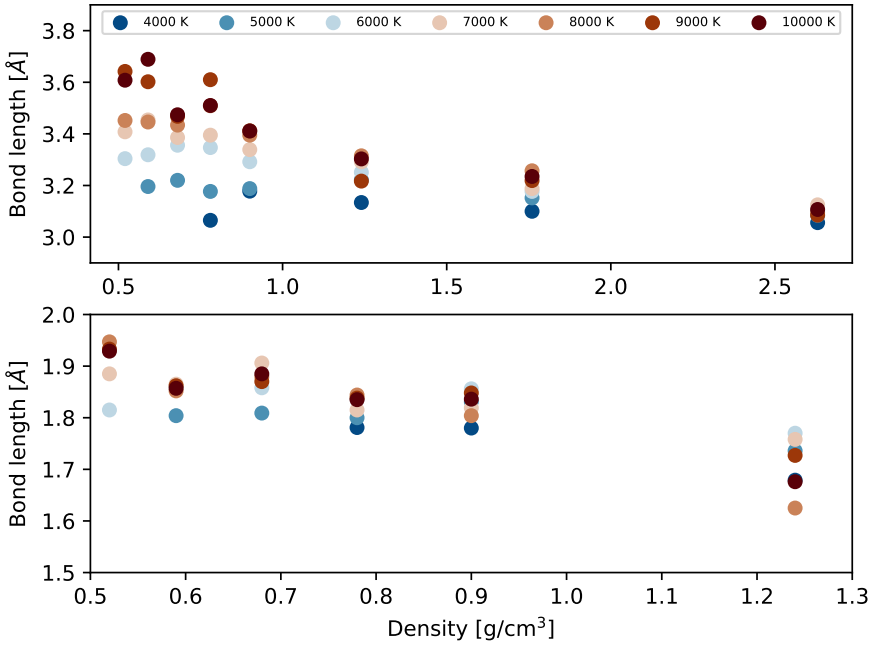


Figure 3.2: (a): peak location of the Mg-O and O-O PDF, the top part describes Mg-O and the bottom O-O. We see a slight temperature effect, with increasing temperatures the bond distance is decreased. Increasing the density somewhat decreases the peak location of Mg-O, but increases the peak location of O-O. (b): bond length as determined by the first minimum after the peak, again the top describes Mg-O and the bottom O-O. The temperature and density effects on the bond length are much stronger than on the peak location. For both Mg-O and O-O we see decreasing the temperature decreases the bond length and increasing the density decreases the bond length as well.

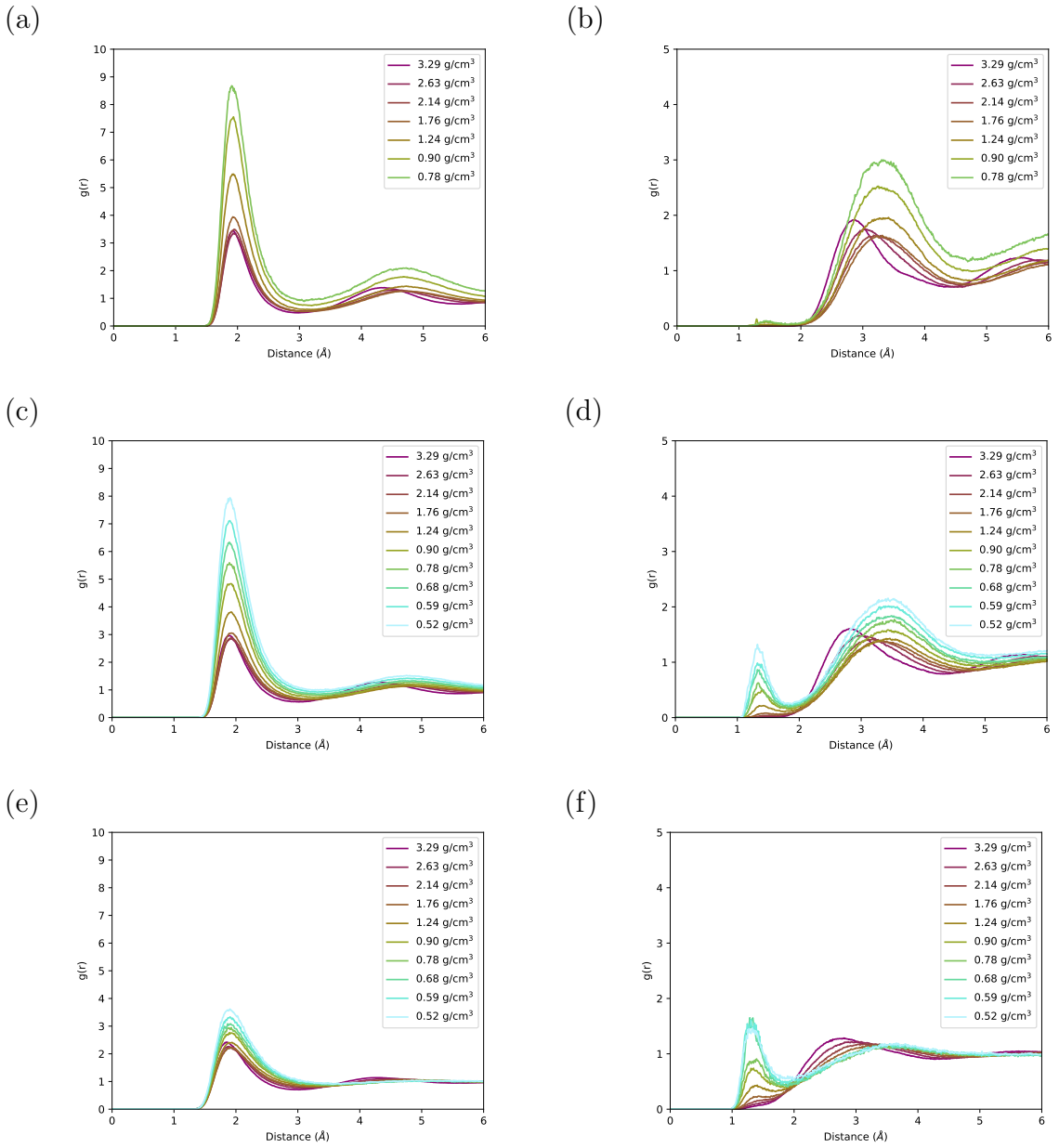


Figure 3.3: The pair distribution function for Mg-O (a, c, e) and O-O (b, d, f) at three isotherms, 4000 (a,b), 6000 (c,d), and 10000 (e,f) K, and densities ranging from 3.29 to 0.52 g/cm^3. The first maxima yield a good approximation of the average bond distances. The first minima yield the radius of the first coordination sphere. As mentioned in section 2.4 we use this value to establish interatomic bonding in the post-processing

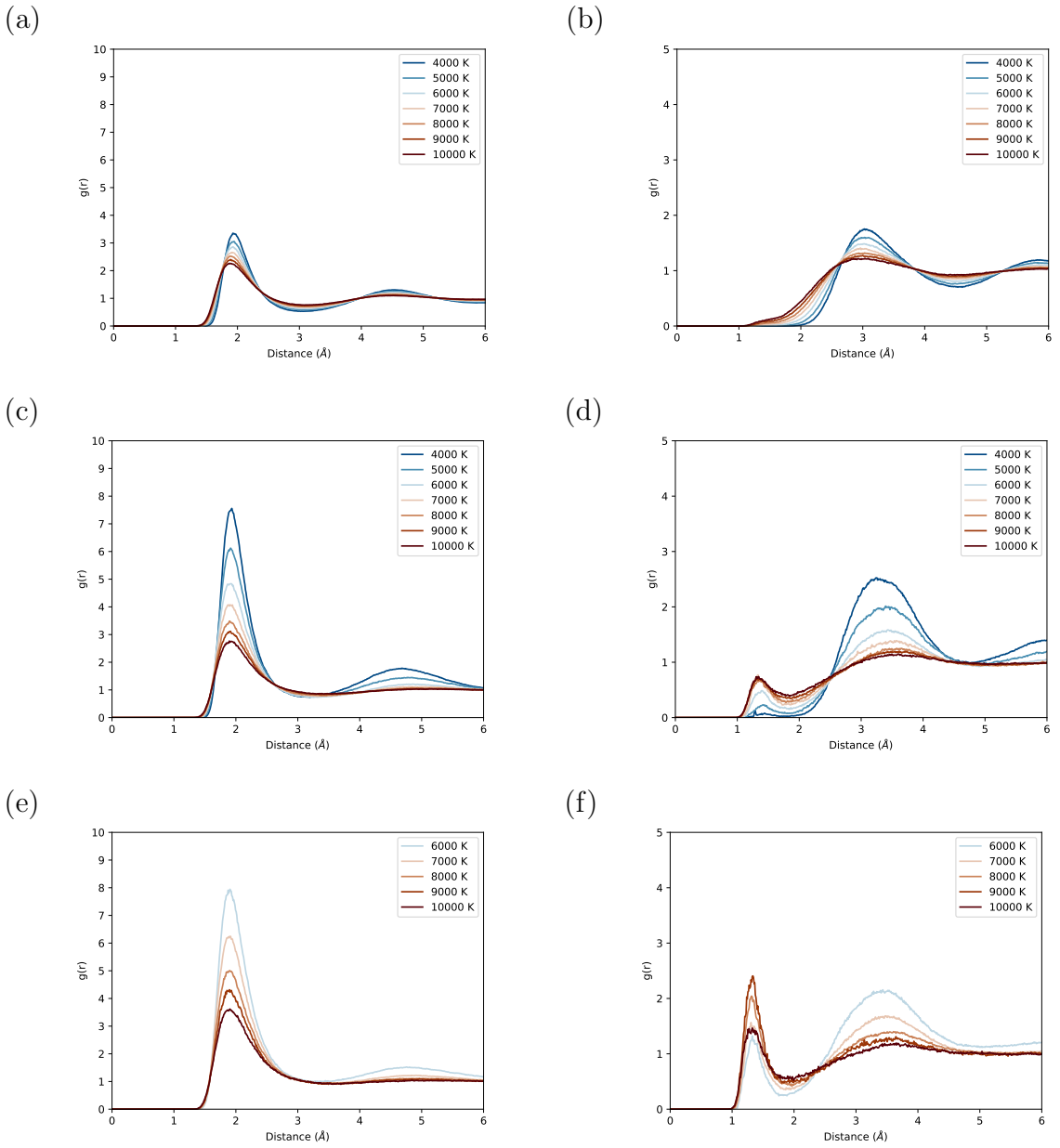


Figure 3.4: The pair distribution function for Mg-O (a, c, e) and O-O (b, d, f) at three densities, 2.63 g/cm^3 (a,b), 0.90 g/cm^3 (c,d), and 0.52 g/cm^3 (e,f) , and all their isotherms. The first maxima yield a good approximation of the average bond distances. The first minima yield the radius of the first coordination sphere. As mentioned in section 2.4 we use this value to establish interatomic bonding in the post-processing

3.2.2 Speciation

In order to follow the natural progression described in figure 2.7, we will now discuss the speciation of MgO. Using the bond lengths obtained from the PDFs discussed in the previous section, we can determine connectivity matrices that describe the speciation of the system.

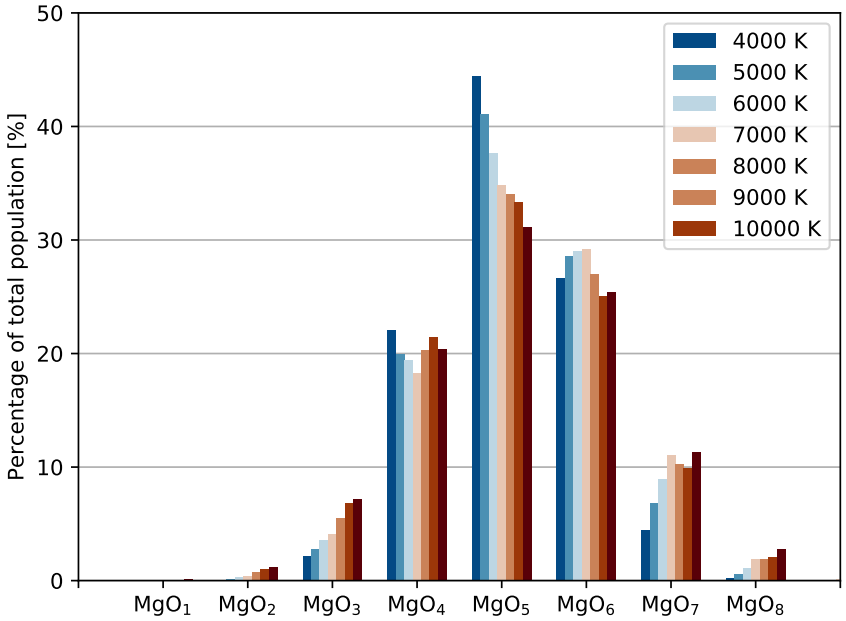
On the whole, we see a shift towards a lower order of coordination as the temperature is increased and density decreases. Both increasing temperatures and decreasing densities favor open, less dense liquid structures. When decreasing densities this is easily achieved as the simulation box volume is increased. For increasing temperatures this means that any void spaces get filled as the liquid attempts to take up more volume. At 2.63 g/cm^3 and 4000 to 6000 K we can see that the MgO liquid consists predominantly of MgO_5 followed by MgO_6 and MgO_4 in order of presence. Supercritical fluid MgO at these densities follows roughly the same distribution, MgO_5 and MgO_6 become almost equivalent and more extreme polyhedra (both large and small) become more common. At the other density extreme, 0.52 g/cm^3, we note a dominance of MgO_2 and MgO_3 at sub-critical temperatures. At super-critical temperatures, MgO_1 becomes much more prevalent and MgO_2 passes MgO_3 in presence. The speciation displayed in figure 3.6 is of the system as a whole, there is no separation between the liquid and vapour phases. For example, at lower densities, it is not unlikely that a significant part of the MgO_1 occurrence is vaporous. Note that MgO_0 (i.e. Mg) is not included in the figure but is most definitely in the system, this species is discussed in the next section.

3.3 Vaporization

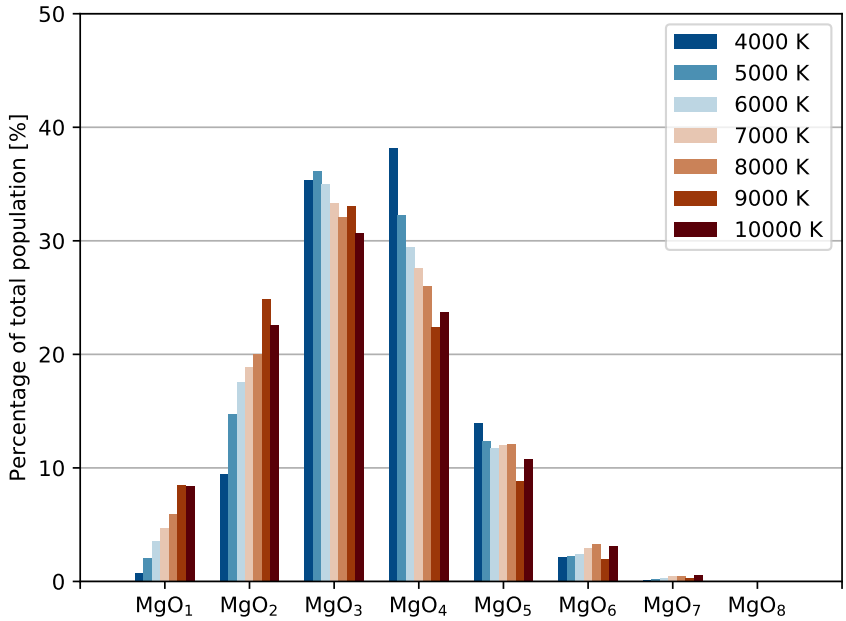
If we want to study the nature of the vapour, we apply the polymer method of the speciation, as described in section 2.4. In an ideal situation, this would describe all of the liquid as a single cluster, while all the remaining small clusters compose the vapour. Figure 3.7 illustrates what this polymer method looks like in a simulation, Tables 3.1 to 3.6 below show the results of the polymerization speciation at 5000 and 6000 K. The percentage value given displays how much of the total lifetime of all clusters together the given cluster provides. The percentages do not add up to 100% as a big contribution to the total lifetime of all clusters is the remaining bulk liquid. Say for one simulation step we might have a vapour consisting of five Mg and four O atoms, the remaining bulk liquid consisting of 103 Mg and 104 O will also have a lifetime. Of course the percentages can be normalized to the total lifetime contribution of all species with less than five total atoms, for example. We also include the average lifetime for each cluster and the count of clusters. Figure 3.8 shows the percentage vapor distribution per isotherm. We choose to only look at clusters of a maximum size of five, this is an arbitrary value that we feel should be sufficient for describing the vapour. A couple of important things to keep in mind when looking at these results, especially the average lifetime and the count, is that not all simulations are of equal length. This means that one can not reliably compare the count results of one system with another, it can only be used for comparison within the same system. The way that clusters are calculated also means that if a cluster is 'absorbed' by the bulk liquid for just a single time-step and released in the next step it is counted as two 'unique' clusters. As mentioned many times before, this speciation is based purely on distance, it is quite possible a MgO_3 is counted whilst it is just an O_2 flying too close to a MgO cluster.

Now, keeping all the aforementioned caveats in mind, we can look at the results and draw some conclusions. We see the presence of molecular oxygen increasing as density is decreased. O_2 is also more present at 6000 K than 5000 K. Looking a both Mg and O ions we can see a similar story; both

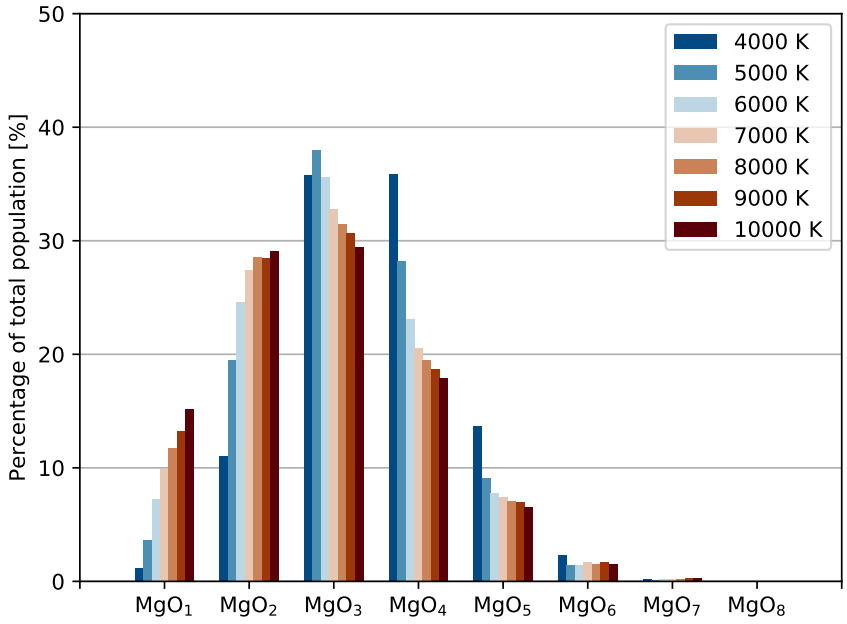
(a)


Percentage of total population [%]
0
10
20
30
40
50
MgO1
MgO2
MgO3
MgO4
MgO5
MgO6
MgO7
MgO8
4000 K
5000 K
6000 K
7000 K
8000 K
9000 K
10000 K

(b)


Percentage of total population [%]
0
10
20
30
40
50
MgO1
MgO2
MgO3
MgO4
MgO5
MgO6
MgO7
MgO8
4000 K
5000 K
6000 K
7000 K
8000 K
9000 K
10000 K

(c)



(d)

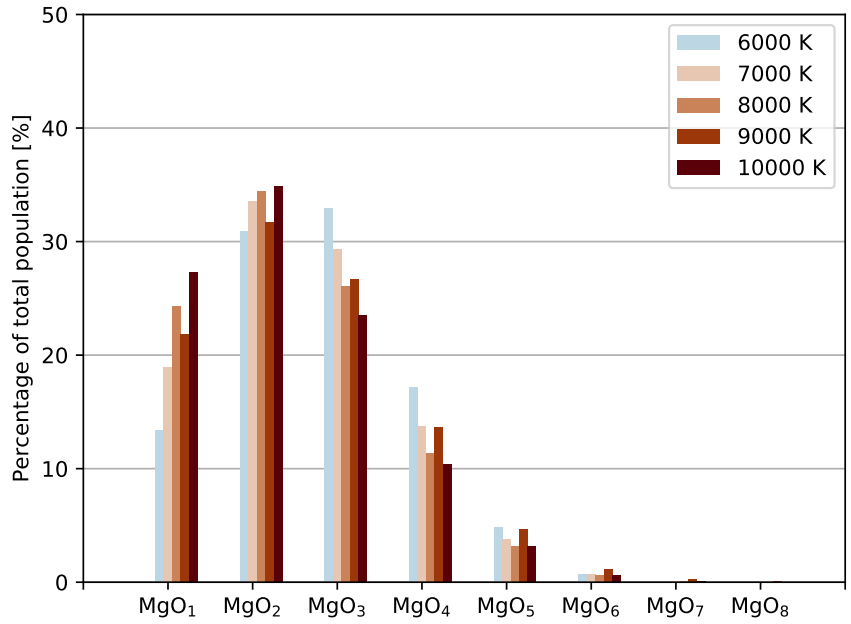


Figure 3.6: The speciation of MgO_x polyhedra. The coordination polyhedra around each atom are obtained using the analysis of the pair distribution function (Figures 3.3 and 3.4). The MgO fluid is dominated by MgO_5 and MgO_6 at high densities. The coordination decreases sharply as the density decreases towards and passes the spinodal density. a, b, c, and d correspond respectively to 2.63, 1.24, 0.90, and 0.52 g/cm^3 densities, respectively.

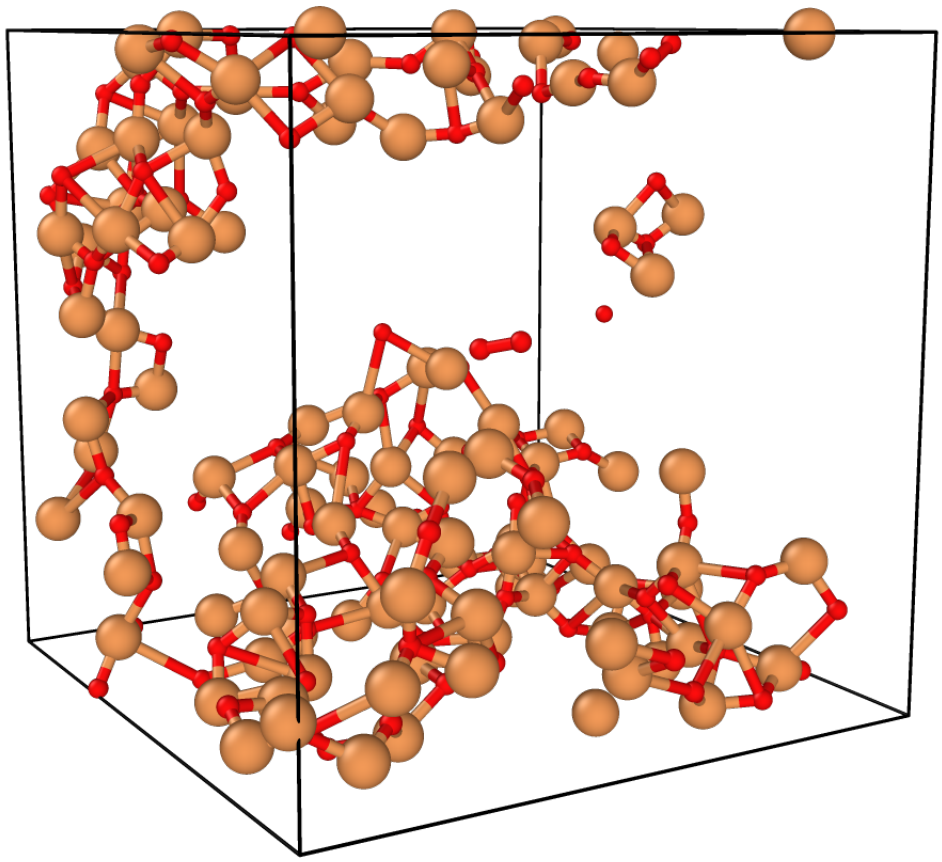


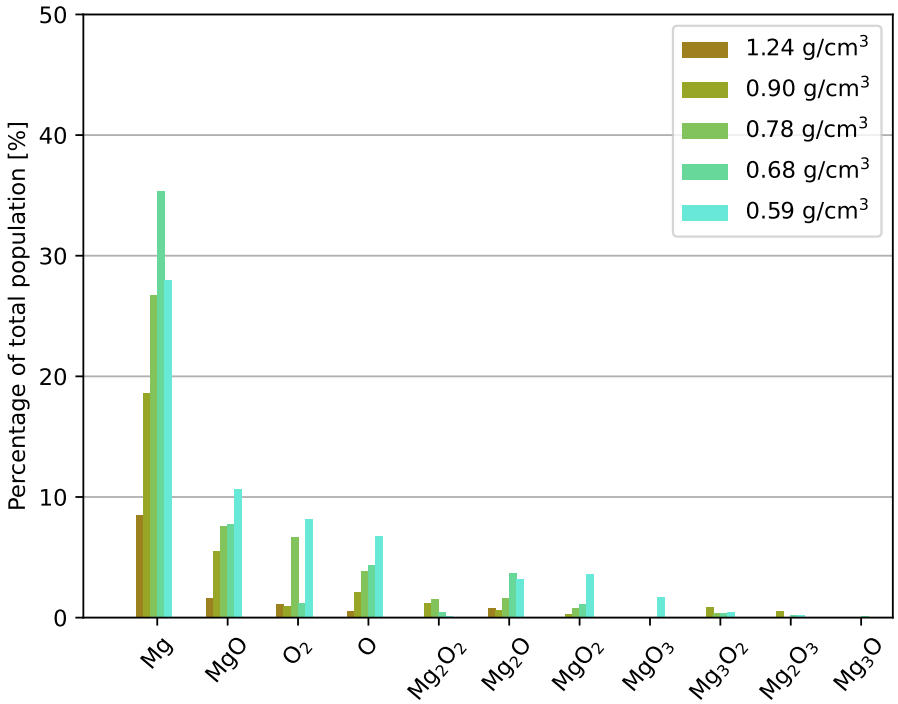
Figure 3.7: The polymer method applied in a simulation at 6000 K and 0.59 g/cm^3. We can see how the majority of the system is connected, forming a bulk liquid. At the same time we can have several isolated atoms or molecules that are not included in the polymer. In this case we choose a frame where we can see an oxygen molecule, two bonded red atoms in the middle of the figure, that has been released from the liquid and is now free-floating in the simulation cell.

increase in presence as density is decreased. This increase is more drastic at the higher densities and the percentages stay near-constant at the lower densities. We note that the vapour consists almost solely of simple clusters like O, O_2, Mg, and MgO at the liquid side of the liquid-vapour dome. Only when we move deeper into the liquid-vapour dome do we start seeing an appreciable amount of more complex clusters consisting of more than two atoms. The fact that the presence of associated species is augmented with decreasing density might seem counter-intuitive at first. Associated species have a lower molar volume than a mono-atomic gas made up from its parts (Mg_2O_2 has a lower molar volume than 2Mg+2O). We could speculate that collision theory plays some role here. As the vaporous species obtain more space, more void space is created, it becomes easier for associated species to survive as they are less likely to interact with other vaporous species or the bulk liquid.

Now let us examine the average lifetimes. One interesting feature that stands out is that the average lifetimes of both O_2 and Mg are considerably larger than the average lifetime of O when looking at the lower densities. We also note that in the 5000 K systems some more complex clusters like MgO_2 and MgO_3 have quite large average lifetimes. However, for MgO_3 this is purely based on a single cluster, which means this average lifetime is unreliable. Except for the aforementioned MgO_2, we see that the average lifetimes for simple clusters of less than or equal to two atoms, are two to three times higher than for the complex ones across all systems.

The last thing we can look at is how the system vaporizes. As discussed in section 1.6.1 the literature states MgO vaporizes congruently. Looking at our results it is quite difficult to draw a similar conclusion. In order to make this a bit easier we look at the stoichiometry of the remaining liquid, instead of the vapour species. On the whole, we see that the stoichiometry stays relatively close to 1:1. We find a maximum variation of three difference between the number of Mg and O in the liquid when looking at liquid of a percentage $> 1\%$. This gives us more confidence to say that MgO vaporizes congruently in our simulations.

(a)



(b)

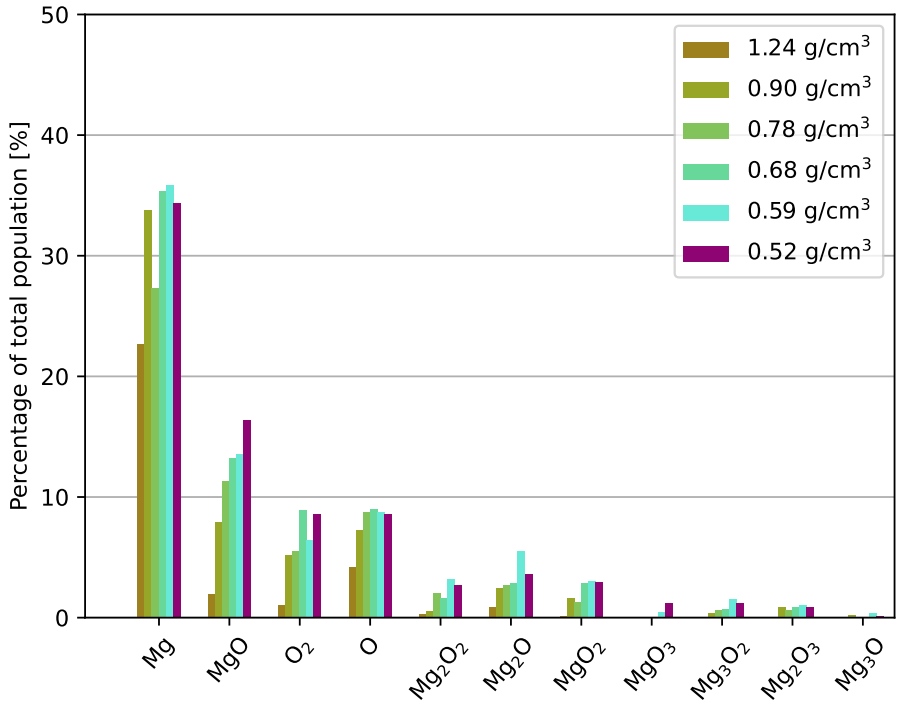


Figure 3.8: Polymer speciation analysis of the 5000 K (top) and 6000 K (bottom) isotherms. Vapor species are put in order of occurrence, determined at 1.24 g/cm^3.

Species \ Density [g/cm^3]	Cluster percentage [%]				
	1.24	**0.90**	**0.78**	**0.68**	**0.59**
O	0.52	2.15	3.87	4.38	6.78
O_2	1.11	0.95	6.66	1.21	8.19
Mg	8.50	18.62	26.71	35.32	27.95
MgO	1.63	5.55	7.59	7.74	10.62
MgO_2	-	0.27	0.77	1.10	3.62
MgO_3	-	-	-	-	1.68
MgO_4	-	-	-	-	-
Mg_2O	0.81	0.62	1.60	3.71	3.18
Mg_3O	-	-	-	0.11	-
Mg_4O	-	-	-	-	-
Mg_2O_2	-	1.24	1.57	0.48	0.12
Mg_2O_3	-	0.51	-	0.20	0.20
Mg_3O_2	-	0.84	0.38	0.35	0.42

Table 3.1: Cluster lifetime percentage contribution to the total lifetime of all clusters in the system for every system on the 5000 K isotherm.

Species \ Density [g/cm^3]	Average lifetime [fs]				
	1.24	**0.90**	**0.78**	**0.68**	**0.59**
O	23.60	45.43	59.61	57.85	87.36
O_2	84.33	27.20	117.95	41.85	206.25
Mg	41.23	95.55	98.46	155.61	143.53
MgO	37.40	48.18	74.14	82.40	84.79
MgO_2	0	13.33	34.00	69.71	140.08
MgO_3	0	251.00	0	0	158.00
MgO_4	0	0	0	31.00	0
Mg_2O	90.50	23.00	46.58	97.82	83.74
Mg_3O	0	0	0	51.00	11.50
Mg_4O	0	0	0	0	0
Mg_2O_2	0	61.66	91.33	53.00	12.50
Mg_2O_3	12.00	73.00	0	30.33	25.50
Mg_3O_2	0	62.00	33.75	31.20	69.67

Table 3.2: Average lifetimes in fs of clusters along the 5000 K isotherm.

Species \ Density [g/cm^3]	**1.24**	**0.90**	**0.78**	**0.68**	**0.59**
	Cluster count				
O	5	7	23	34	39
O_2	3	5	20	13	20
Mg	47	29	96	102	98
MgO	10	17	36	42	63
MgO_2	0	3	8	7	13
MgO_3	0	1	0	0	1
MgO_4	0	0	0	1	0
Mg_2O	2	4	12	17	19
Mg_3O	0	0	0	1	2
Mg_4O	0	0	0	0	0
Mg_2O_2	0	3	6	4	4
Mg_2O_3	1	1	0	3	4
Mg_3O_2	0	2	4	5	3

Table 3.3: Cluster count per system along the 5000 K isotherm. Keep in mind that not all simulation lengths are equal, so direct comparison between systems is not possible.

Species \ Density [g/cm^3]	**1.24**	**0.90**	**0.78**	**0.68**	**0.59**	**0.52**
	Cluster percentage [%]					
O	4.16	7.24	8.73	9.00	8.76	8.62
O_2	1.03	5.17	5.49	8.95	6.39	8.58
Mg	22.68	33.79	27.32	35.32	35.83	34.36
MgO	1.96	7.94	11.33	13.23	13.52	16.41
MgO_2	0.14	1.61	1.27	2.87	3.04	2.97
MgO_3	-	-	-	-	0.44	1.22
MgO_4	-	-	-	-	-	-
Mg_2O	0.85	2.45	2.70	2.87	5.51	3.58
Mg_3O	-	0.23	-	-	0.39	0.11
Mg_4O	-	-	-	-	-	-
Mg_2O_2	0.32	0.58	2.07	1.58	3.15	2.66
Mg_2O_3	-	0.89	0.62	0.84	1.04	0.85
Mg_3O_2	-	0.39	0.64	0.70	1.51	1.18

Table 3.4: Cluster lifetime percentage contribution to the total lifetime of all clusters in the system for every system on the 6000 K isotherm.

Average lifetime [fs]

Species \ Density [g/cm^3]	**1.24**	**0.90**	**0.78**	**0.68**	**0.59**	**0.52**
O	37.22	40.50	46.36	73.48	75.58	83.94
O_2	33.11	66.31	98.75	116.71	128.11	110.48
Mg	44.92	67.27	62.92	93.94	101.86	113.91
MgO	22.71	37.83	59.72	70.63	67.34	84.32
MgO_2	21.00	42.35	57.13	69.22	64.11	51.80
MgO_3	-	39.00	-	16.00	50.40	74.20
MgO_4	-	-	-	-	89.00	35.50
Mg_2O	24.80	29.51	50.74	50.16	57.53	47.97
Mg_3O	-	26.00	18.00	10.00	38.91	24.60
Mg_4O	-	-	-	-	-	-
Mg_2O_2	29.33	21.83	52.29	35.90	45.73	56.96
Mg_2O_3	-	36.09	56.25	41.22	56.43	54.42
Mg_3O_2	-	38.50	38.33	36.82	47.64	41.62

Table 3.5: Average lifetimes in fs of clusters along the 6000 K isotherm.

Cluster count

Species \ Density [g/cm^3]	**1.24**	**0.90**	**0.78**	**0.68**	**0.59**	**0.52**
O	32	80	67	110	132	125
O_2	9	35	20	69	57	95
Mg	145	226	156	338	402	370
MgO	24	93	68	168	229	238
MgO_2	2	17	8	37	54	70
MgO_3	0	1	0	3	10	20
MgO_4	0	0	0	0	1	2
Mg_2O	10	37	19	51	109	91
Mg_3O	0	4	1	3	11	5
Mg_4O	0	0	0	0	0	0
Mg_2O_2	3	12	14	39	78	57
Mg_2O_3	0	11	4	18	21	19
Mg_3O_2	0	4	6	17	36	34

Table 3.6: Cluster count per system along the 6000 K isotherm. Keep in mind that not all simulation lengths are equal, so direct comparison between systems is not possible.

3.4 Dynamical properties

Now that we have looked at the structural properties, it is time to take a look at dynamical properties, the first one being the MSD. The calculated MSDs of four densities are shown in figure 3.9. Predictably, on the whole, we can see the MSD increases with increasing temperatures and decreasing density. Using the MSD we calculate the self-diffusion coefficient, as mentioned in section 2.4.3. Values of both the VAF method and the slope method are presented in table 3.7.

We see across most systems that the oxygen atoms diffuse slightly faster than the magnesium atoms, which is most likely related to their lower mass. A more interesting feature can be seen in figure 3.10; for both species, we see that the self-diffusion coefficient follows a linear trend (in logarithmic scale) at high densities. However, at lower densities around the transition into the liquid-vapour coexistence region, we note a sudden change of slope for this linear trend, in the systems below the critical temperature. The linear trend seems more or less continuous for the systems above the critical temperature. Systems that fall in the liquid-vapour region display near-constant self-diffusion coefficient values. A possible explanation for this phenomenon is that the diffusion coefficient is controlled by the liquid phase. When the system reaches the liquid-vapour region, the liquid is no longer able to take up all of the space it is given. Instead, void spaces form where some individual gaseous particles are allowed to travel through. As nothing really changes about the liquid then, regarding temperature and liquid volume, the diffusion within the liquid stays more or less constant. Of course, we do not see this exactly, likely due to statistical errors and contributions of gaseous particles to the MSDs. The beauty of this method, is that this gives an additional indication of the critical temperature, without being related to the P-T-ρ plots discussed earlier.

	Diffusion coefficient (m^2/s)			
Temperature	Mg (VAF)	O (VAF)	Mg (MSD)	O (MSD)
4000	$4.87e^{-8}$	$4.34e^{-8}$	$3.22e^{-8}$	$3.25e^{-8}$
5000	$6.36e^{-8}$	$8.59e^{-8}$	$6.69e^{-8}$	$7.43e^{-8}$
6000	$1.51e^{-7}$	$1.35e^{-7}$	$1.39e^{-7}$	$1.29e^{-7}$
7000	$2.24e^{-7}$	$2.69e^{-7}$	$1.85e^{-7}$	$2.26e^{-7}$
8000	$2.83e^{-7}$	$3.32e^{-7}$	$2.37e^{-7}$	$2.50e^{-7}$
9000	$2.92e^{-7}$	$3.65e^{-7}$	$2.52e^{-7}$	$3.46e^{-7}$
10000	$2.98e^{-7}$	$4.09e^{-7}$	$2.86e^{-7}$	$3.85e^{-7}$

Table 3.7: Comparison between the diffusion coefficients at $0.78 g/cm^3$, as a function of temperature, estimated from the slope of the mean square displacements as a function of time (MSD) and obtained from the velocity auto-correlation function (VAF).

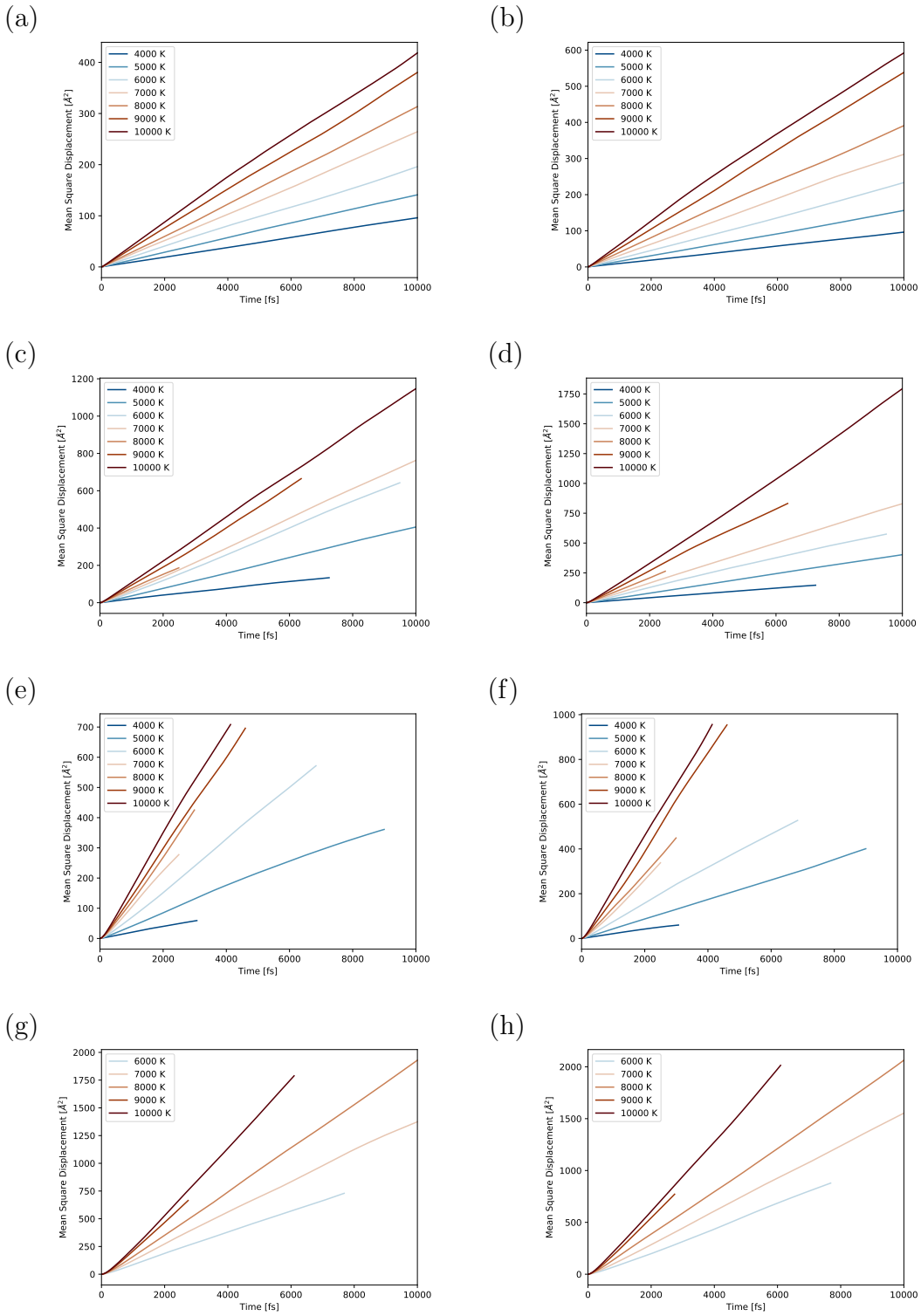
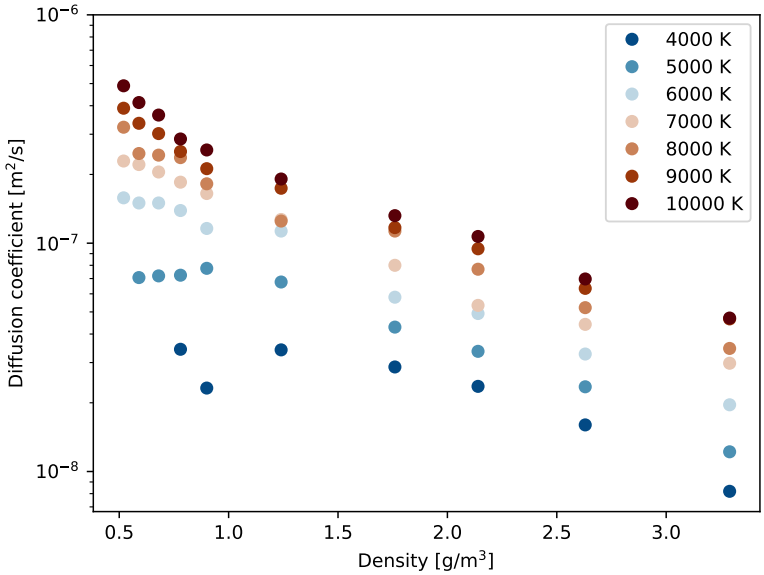


Figure 3.9: Mean square displacements calculated at four isochores. a, b: 2.63 g/cm^3, c, d: 1.24 g/cm^3, e, f: 0.78 g/cm^3, g, h:0.52 g/cm^3. The differences in lengths are the result of differences in the length of simulations, the shortest being 5 ps and the longest 20 ps. MgO is fluid at all conditions studied here.

(a)



(b)

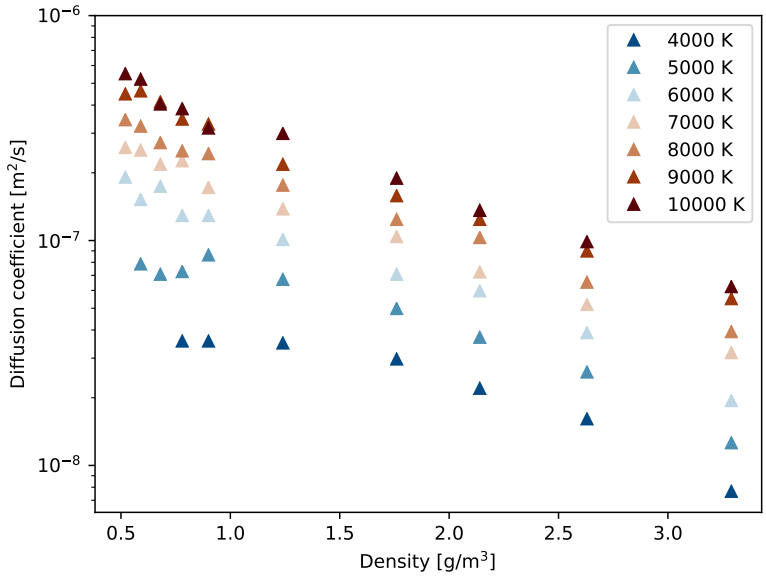


Figure 3.10: Self-diffusion coefficients for Mg atoms (a) and O atoms (b) as calculated from the slope of the mean-squared displacement (MSD) plotted in log scale against the density. Below the critical temperature, the diffusion coefficients exhibit a change of slope corresponding to the transition inside the liquid-vapour dome. Above the critical temperature, the diffusion coefficient changes monotonously with density.

Using the aforementioned VAC function we can obtain a vibrational spectrum by keeping the real part of the Fourier, eq. 4.1. The resulting spectra are shown in figure 3.11. All the vibration spectra shown here are calculated in liquid, vapour, and supercritical fluid states, which means that comparison with known vibrational modes of crystalline MgO is not as straightforward. Both experimental and theoretical work (e.g. Jasperse et al. (1966), Han et al. (2008), and Calandrini et al. (2021)) indicate that B1 MgO is infrared active in the region of 400 to 700 cm^{-1}, where the lower frequency part is related to the transverse optical (TO) component and the higher frequency to the longitudinal optical (LO) component. This range is also varied depending on pressure and temperature conditions. We also note that as the temperature is increased, the frequencies tend to decrease in wave number. When we compare these results with our vibrational spectra, in particular, the 4000 K line of figure 3.11a, we do see some similarities. Of course, as the density is decreased or the temperature is increased, we lose this resemblance. Another important factor to keep in mind when examining the vibrational spectra of phases other than crystalline is the effect of diffusion. Diffusion will elevate the spectral line at the lowest wave numbers, having the strongest effect at 0 cm^{-1}. At the lowest density, for both temperatures, in figure 3.11d the vibrational spectrum because featureless and approaches that of an atomic gas where it is controlled purely by diffusion.

$$I(\omega) = Re[\int_0^\infty m \langle v(t+\tau)v(t)\rangle \, e^{-i\omega t}] \tag{4.1}$$

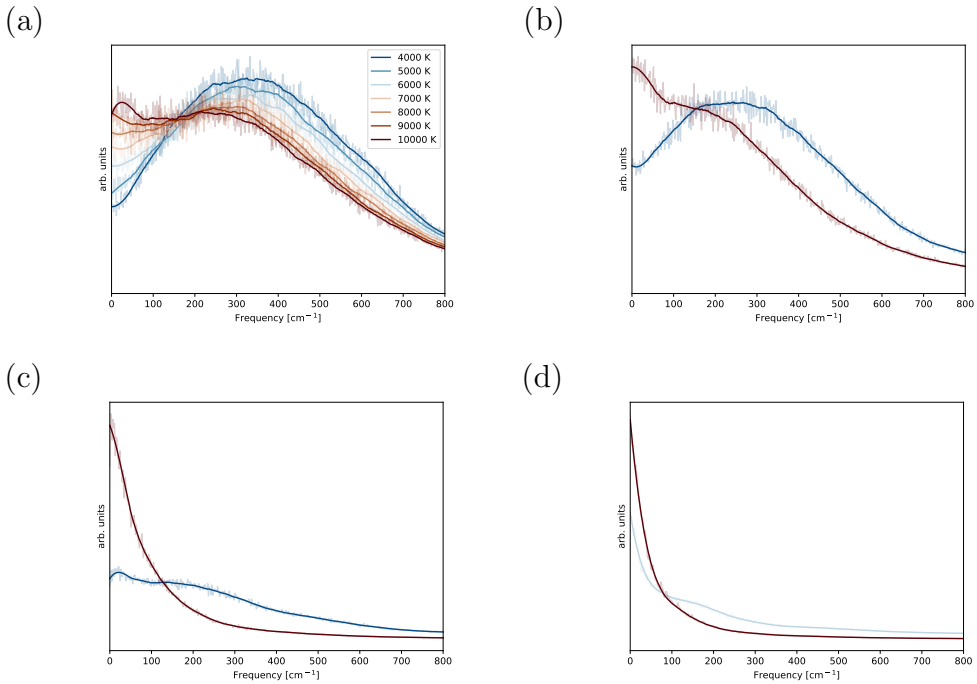


Figure 3.11: Total vibrational spectra at several densities of fluid MgO are obtained from the velocity-velocity self-correlation function. All isotherms are shown for the highest density (a). Only the values computed at the lowest and the highest isotherms are shown for the others. The noisy data, shown in the background, was filtered using the Savitzky-Golay filter Savitzky and Golay (1964). At high density and low temperature, the spectrum shows a broad peak around 300 - 500 cm^{-1}. The peak is smoothed out with decreasing density and increasing temperature. a, b, c, and d correspond respectively to 3.29, 2.63, 0.90, and 0.52 g/cm^3 densities.

3.5 Bader charge

One last analysis we do on MgO is to perform a Bader charge analysis. This is done in order to give insight into the electronic properties of the vapor phase. We perform highly accurate static calculations using the tetrahedron method with corrections from Blöchl et al. (1994), instead of the usual Fermi smearing. We choose systems based on the significant presence of vaporous species, here at 6000 K and 0.68 g/cm^3. Results of one of the studied configurations are shown in figure 3.12. We can see that the atomic charges are slightly off the nominal values of +2 and -2, we speculate that this is either due shortcomings in the method or temperature related effects. We find an interesting relationship between the coordination, as calculated in section 3.2, and the Bader charge. We find that as the coordination number is decreased, the charge becomes more neutral. This would indicate that vaporous particles in our system are neutral, or close to neutral, in charge. Meaning we are dealing with an atomic gas, instead of an ionic one.

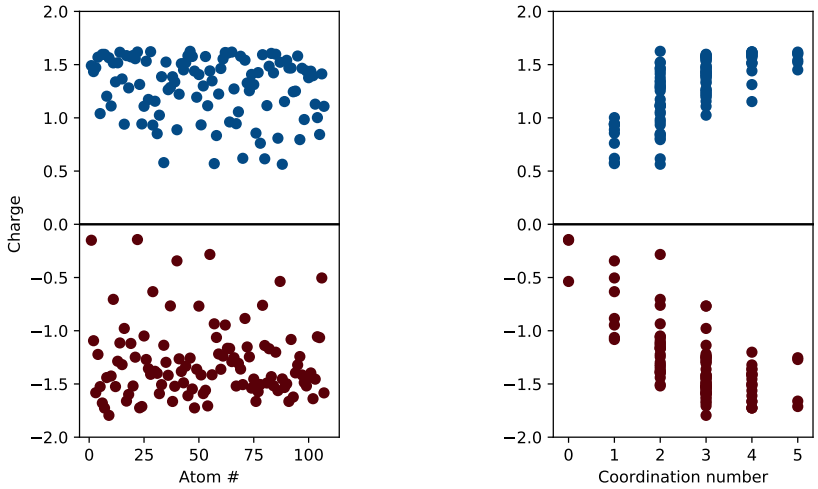


Figure 3.12: Atomic charges for all Mg (blue) and O (red) atoms for a representative snapshot inside the liquid-vapour dome (a). The order of the coordination polyhedron around each atom (b). Charge values close to the nominal (+2 and -2) correspond to highly coordinated atoms, which lie in the bulk liquid. Charge values close to zero correspond to low coordination numbers, as encountered on the liquid-vapour interface and in the vapour.

CHAPTER 4

Works associated with MgO

Contents

In this chapter, we will discuss yet unpublished work which is closely related to MgO in one way or another. We begin by putting MgO in the context of the whole MgO-SiO_2 material assembly. After, we perform a study on CaO in a similar manner as done for MgO. These materials are physically very similar and we compare their results. We end this chapter with a short study on the isotopic effects on vaporization species distribution.

4.1 Magnesium silicates

The MgO-SiO mineralogical system is arguably the most important one in geology. MgO and SiO_2,combined , account for more than 80% of the crust+mantle composition (Morgan and Anders (1980)) and about 75% of the moon (Morgan et al. (1978)). A number of studies have been done on materials of the MgO-SiO system under extreme conditions shortly after a giant impact. We want to combine some of these results and see if we are able to draw any relationships between material properties and the MgO/SiO ratio. In the following section we will be mostly using results from Green et al. (2018) on SiO_2, Xiao and Stixrude (2018) on $MgSiO_3$, Townsend et al. (2020) on Mg_2SiO_4, and our own work Bögels and Caracas (2022) on MgO. Not every type of characteristic is

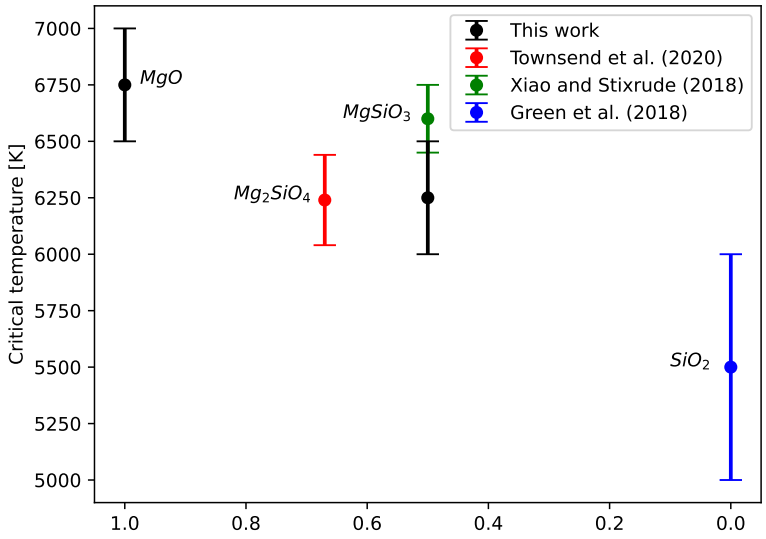


Figure 4.1: Critical points and associated error bars of MgO-SiO_2 materials. Critical points were determined by Green et al. (2018) (SiO_2), Townsend et al. (2020) (Mg_2SiO_4), Xiao and Stixrude (2018) ($MgSiO_3$), and Bögels and Caracas (2022) (MgO). Overall, we can see a trend of a slight decrease in critical temperature with decreasing MgO/SiO_2 ratio. We have added the recalculated value for $MgSiO_3$, as discussed in this chapter.

discussed in each publication. For this reason, our comparison will be limited to the critical point and speciation.

4.1.1 Critical point

We begin our comparison with one of the most essential parameters: the critical point. In order of chemistry, the critical temperatures (T_c) and associated critical densities (ρ_c) obtained are:

- MgO: 6500-7000 K, 0.45-0.60 g/cm^3 (Bögels and Caracas (2022))
- Mg_2SiO_4: 6240 $\pm$ 200 K, 0.52 $\pm$ 0.03 g/cm^3 (Townsend et al. (2020))
- $MgSiO_3$: 6600 $\pm$ 150 K, 0.48 $\pm$ 0.05 g/cm^3 (Xiao and Stixrude (2018))
- SiO_2: 5000-6000 K, 0.5-1 g/cm^3 (Green et al. (2018))

If we take these values and put them together on one plot like in figure 4.1, we notice that the critical temperature increases with the amount of MgO. This is somewhat expected as MgO is a refractory material and its critical temperature is found to be higher than that of pure SiO_2. The number of data points is quite limited, so talking about an 'outlier' seems somewhat disingenuous.

However, we would like to propose that the critical temperature for $MgSiO_3$ is too high. When comparing the methodologies of the aforementioned works, most of them are quite similar in regards to the simulation box, except for $MgSiO_3$. The works from Green et al., Townsend et al., and Bögels and Caracas all apply a filled simulation box with periodic boundary conditions, Xiao et al. practice a slightly different approach using a slab simulation box surrounded by a vacuum. Slab simulation boxes are of course widely applied in DFT-MD, most notably when studying surface characteristics or catalysis. However, we have some questions on its applicability when it comes to determining the critical point. For this reason, we perform some simulations on $MgSiO_3$ ourselves using the same methodology applied for MgO in order to compare our results with the previous one. Since we are only interested in the critical point we perform our simulations over quite a narrow range, as we know roughly where to expect the critical point. In order to obtain a reliable average pressure, we run our simulations for at least 5 ps, this does mean that other characteristics like speciation are less reliable from these simulations. The equations of state (EOS) obtained for five isotherms are shown in figure 4.2. From our new simulations, we estimate the critical temperature to be slightly lower than previously determined. This would mean that it fits much better in the overall trend in the MgO/SiO system we previously identified. By looking at the EOS we can note there must be some problem with the fit as they overlap at 1.4 g/cm^3 in the figure. The fit itself is also worthy of discussion, and there is little consistency between the four studies in this regard. Both MgO and Mg_2SiO_4 are fit using a cubic function, but with some differences with respect to the intercept and fitting per isotherm. $MgSiO_3$ is fit using a Wegner expansion (Vega et al. (1992)) while SiO_2 applies a high order polynomial to relate the pressure to temperature and density. The Wegner expansion is a method to describe the coexistence curve as a function of a the effective critical exponent, this latter parameter is a measure of the shape of the coexistence curve. Using a Wegner expansion allows the user to determine the system's critical temperature with liquid-vapour coexistence data. All this being said, we do believe there is some value in trying to establish a relationship between MgO/SiO ratio and the critical temperature and we believe that our newly calculated temperature for $MgSiO_3$ fits this relationship better. We do not expect the relationship between the critical temperature and the chemistry of the system to be linear. Looking at the binary phase diagram at low-pressures of this system we find non-linear liquidus phase curves. If the liquidus temperatures translate into the critical points in any way we have reason to assume a linear relationship here either.

4.1.2 Coordination and species fraction

Starting out with the species fraction we can compare only the MgO species, as no species fractionation data is given for SiO_2. Because of how the $MgSiO_3$ simulations are done we obtain vapour species fraction only with respect to temperature, not density. For this reason, we will only compare the systems at different temperatures and average the vapour fraction for MgO and Mg_2SiO_4 over roughly 0.5 - 1.0 g/cm^3. Looking at the temperature range of 5000 to 5500 K, we see that for both magnesium-silicates O_2 and SiO dominate the vapour species fraction with Mg, SiO_2, O following thereafter, MgO and Si have the lowest contribution. In contrast, in MgO, we see on average that Mg is by far the dominant vapour species followed by MgO, O_2, and O in order of occurrence. At higher temperatures, 6000 to 6500 K, the differences between the vapour fractions of the magnesium-silicates decrease and the distribution becomes slightly more equal. Still, O_2 is the dominant vapour species, and the species after that are more equal in fraction with SiO, O, and Mg getting closer to each other. Si and MgO are still consistently under-represented in

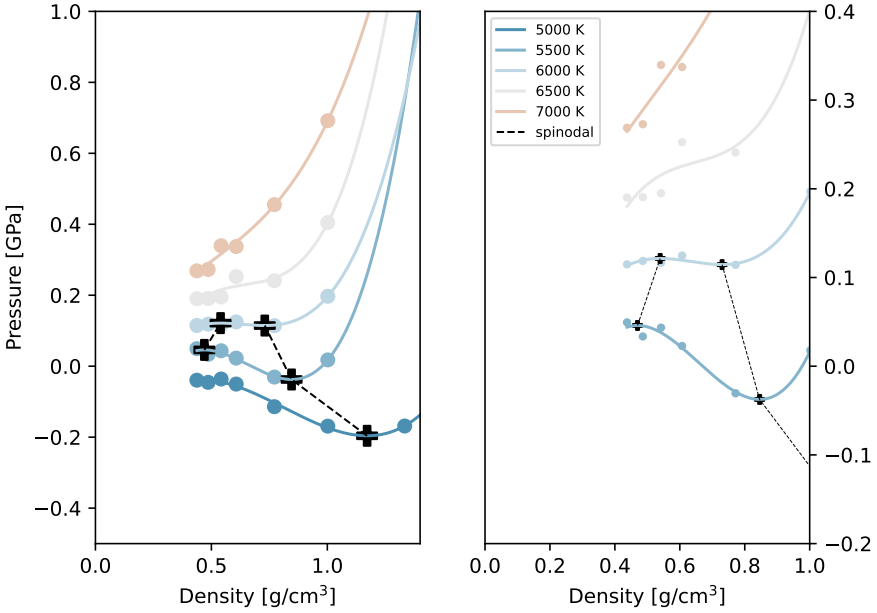


Figure 4.2: Variation of pressure as a function of density for various isotherms with a $MgSiO_3$ system. The solid lines are equations of state (EOS) defined using a cubic function. The minima and maxima are spinodal points which correspond to the outer limits of the stable liquid-vapour region. The critical point is found between the isotherm with the last minima and maxima and the first without. In the case of $MgSiO_3$ we estimate here that the critical temperature lies between 6000 and 6500 K, with a critical density of between 0.5 to 0.7 g/cm^3

the vapour. In MgO, the story stays more or less the same as before, with Mg being the most present followed by MgO, O_2 and O vapour fraction are close to equivalent at the higher temperatures for MgO. One important thing to keep in mind when comparing these results is that both magnesium-silicates vaporize incongruently while MgO was found to vaporize congruently. Knowing the difference in critical temperature (and boiling point) between MgO and SiO_2, it makes sense that in the magnesium-silicates we see an elevated presence of Si in the vapour when compared to Mg.

Comparing our result with work done by Fegley et al. (2016), we do see some similarities in the vapour distribution. In their work, they present a vapour composition of 61% Mg, 24% O_2, 13% O, and 2% MgO at 2000 K. Now comparing that to our results at 6000 K and 0.90% we get a relative composition of 62.4% Mg, 9.5% O_2, 13.4% O, and 14.7% MgO. Of course, these values depend on what temperature and density you choose but on the whole, we see a good agreement with the amount of Mg and O to a lesser extent. Our greatest differences are in the prevalence of O_2 and MgO. While the decrease in O_2 can be quite easily explained by the difference in temperature, higher temperatures will favour isolated O, and the difference in MgO is trickier. A possible explanation is that while Mg content in the vapour does not seems to increase appreciably with temperature,

O content does. This increase in O in the vapour could facilitate the formation of MgO species. Our values also seem to imply that MgO vaporizes incongruently, as ($\%O + 2 * \%O_2 = / = \%Mg$). However, as stated earlier, we do see it vaporizes congruently so we propose that a large amount of O in the vapour species can be found in more complex vapour species.

The coordination number will be compared for SiO_2, $MgSiO_3$, and MgO. At 4000 K both silicate materials have Si-O coordination close to 4. The Mg-O coordination in $MgSiO_3$ is 5, which is the same in MgO at higher densities. However, at lower densities, a coordination of four becomes more prevalent. The Si-O coordination seems much less affected by temperature than the Mg-O one, from 4000 to 6500 K the coordination decreases by only 1 in $MgSiO_3$ and SiO_2. The Mg-O coordination on the other hand decreases by more than 1.5 in $MgSiO_3$ and in MgO we see a change of roughly 2. Whether this relationship is related to any particular chemistry is doubtful. More likely is that both coordination numbers trend towards gas-like behaviour, which would have a lower coordination number, and as Mg-O starts at higher coordination, it has to decrease more sharply with increasing temperature.

4.2 CaO

The other refractory material studied in this work is CaO. CaO has a lot of similarities with MgO, as described in chapter 1. For this reason, we want to see what are the differences, if any, between the two materials and find an explanation for them. This body of work is not as extensive as done for MgO. We will discuss the critical point and structural properties of CaO.

4.2.1 Critical point

Just like with MgO and $MgSiO_3$ before, we construct EOS on various isotherms. As we have some expectation for the critical temperature of CaO, we limit our isotherms to 5500, 6000, 6500, and 7000 K. We perform simulations at densities ranging from 2.46 g/cm^3 to 0.37 g/cm^3. The resulting EOS are displayed in figure 4.3. The last isotherm to contain a minimum and maximum is 6000 K, so we estimate the critical temperature to be between 6000 and 6500 K. From the last minimum and minimum we can also estimate the critical density range to be between 0.5 and 0.7 g/cm^3. This puts CaO at a slightly lower critical temperature than the one we found for MgO. The lower critical temperature for CaO is consistent with the lower boiling point of CaO. This is also consistent with respect to the melting temperatures of both materials, although the melting temperature of CaO leaves some uncertainty and the upper bound of this uncertainty is above the melting temperature of MgO.

4.2.2 Structural properties

Pair distribution function

Just like we did before with MgO we will take a look at the PDF to gain some insight into the structural behaviour of our CaO system. Figure 4.4 shows the PDF for Ca-O and O-O calculated for the two isotherms of 6000 and 7000 K. Instead of describing the PDF on their own we will immediately compare them to the PDF of MgO, as they show a lot of similarities. We will start by comparing the cation-oxygen PDF of both systems. The exclusion radius for CaO is ever so slightly larger than for MgO, a difference of about 0.1-0.2 Å. This is likely caused by the difference

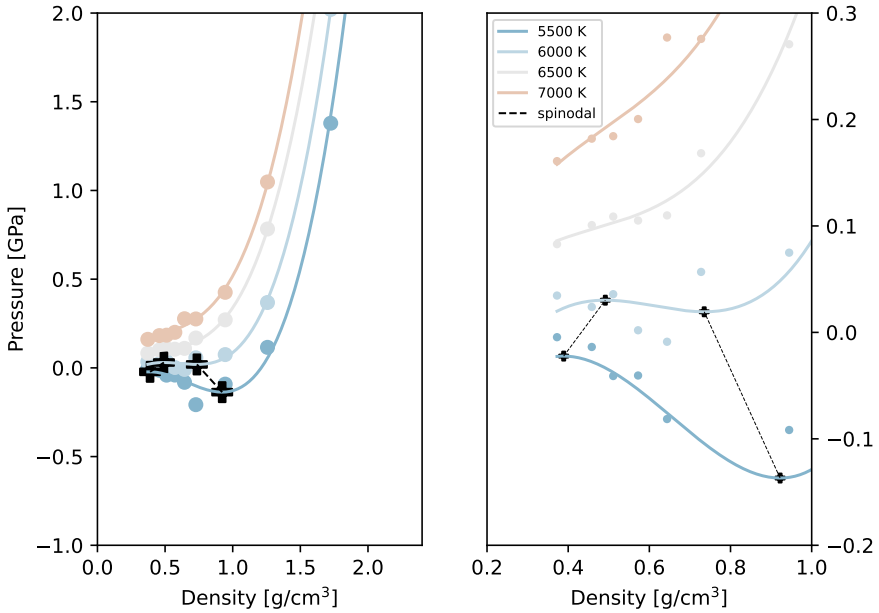


Figure 4.3: Variation of pressure as a function of density for various isotherms with a CaO system. The solid lines are equations of state (EOS) defined using a cubic function. The minima and maxima are spinodal points which correspond to the outer limits of the stable liquid-vapour region. The critical point is found between the isotherm with the last minimum and maximum and the first without. In the case of CaO, we estimate here that the critical temperature lies between 6000 and 6500 K.

in atomic radius between Mg and Ca. Looking at the first minima of the PDF, the limit of the first coordination sphere, we can see that for Mg-O this distance is marginally shorter than for Ca-O. We see the Mg-O bond length ranging from 3.1 to 3.4 Åwhile the Ca-O bond length ranges from roughly 3.3 to 3.6 Å. On the whole, both PDF are very similar, which suggests that the liquid structure of both systems is nearly identical.

Looking at the O-O PDF for both CaO and MgO, the differences become nearly impossible to see (if you look past the rough CaO graph due to sample size). For the first peak related to molecular oxygen we expect the two to have the same behaviour. Molecular oxygen stays the same chemical in both systems. For the second peak, related to O-O distance within the liquid, which is a three-body problem related to Ca-O distance and O-Ca-O bond angle, the maximum is at a larger distance in the CaO system than in the MgO one. We can resolve this discrepancy by thinking of the Cation-Oxygen PDF, where the Ca-O distance was found to be slightly larger than the Mg-O distance. This larger Ca-O distance will have double the effect on the O-O distance in the fluid, since it is an O-Ca-O problem.

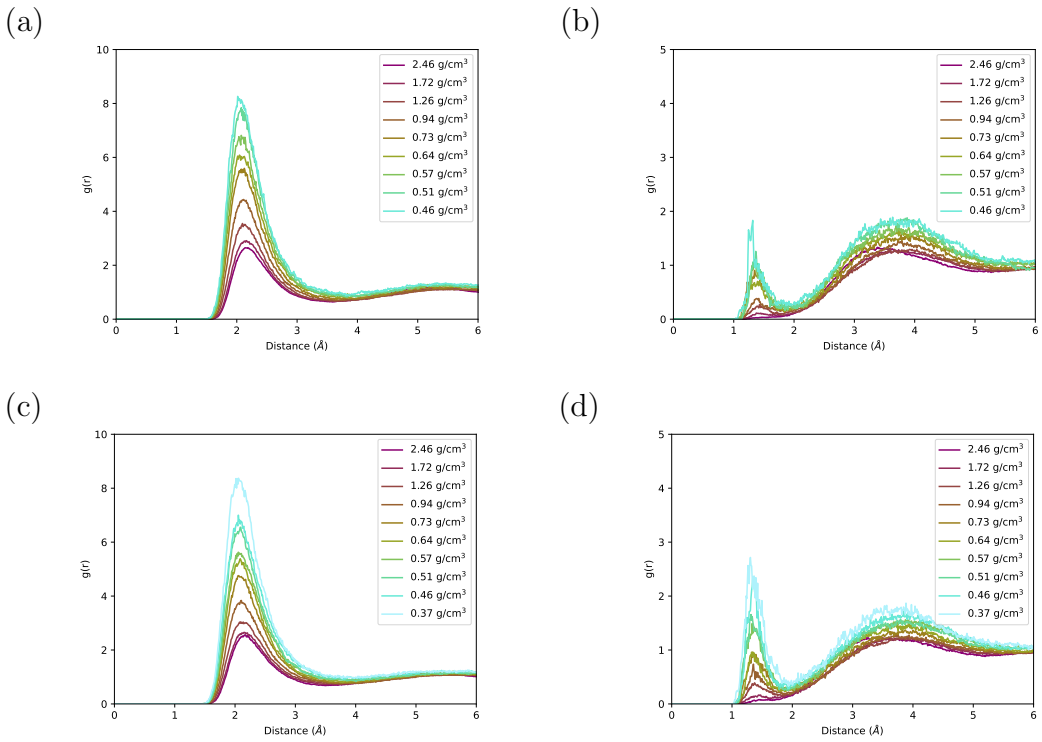


Figure 4.4: The pair distribution function for Ca-O (a, c) and O-O (b, d) at two isotherms, 6000 (a,b) and 7000 (c,d) K, and densities ranging from 2.46 to 0.37 g/cm^3. The first maxima yield a good approximation of the average bond distances. The first minima yield the radius of the first coordination sphere. As mentioned in section 2.4 we use this value to establish interatomic bonding in the post-processing

Speciation

The vast majority of our simulations in the CaO system are too short (maximum of 5 ps) to perform any meaningful speciation analysis. For this reason, we take a single isochore and run three additional simulations for each temperature, for 5 ps each, meaning we will have a total of 20 ps per system. The results of these simulations are summarized in figure 4.5. We calculate the speciation behaviour at 0.94 g/cm^3 for CaO and we will compare it with the results for MgO at 0.90 g/cm^3. Here the differences are more evident than with the PDF before. While the change in coordination for MgO from 5000 to 7000 K is quite pronounced, this is not true for CaO. In CaO the coordination stays almost constant from 5500 to 7000K, with CaO_2 and CaO_3 being the prevailing dominant species. This large difference between CaO and MgO is unexpected after seeing how similar the PDF are. We propose this difference is due to a sampling issue, not anything related to the physics/chemistry of the system. A closer analysis of the speciation of MgO using ML later on in this manuscript should answer this question.

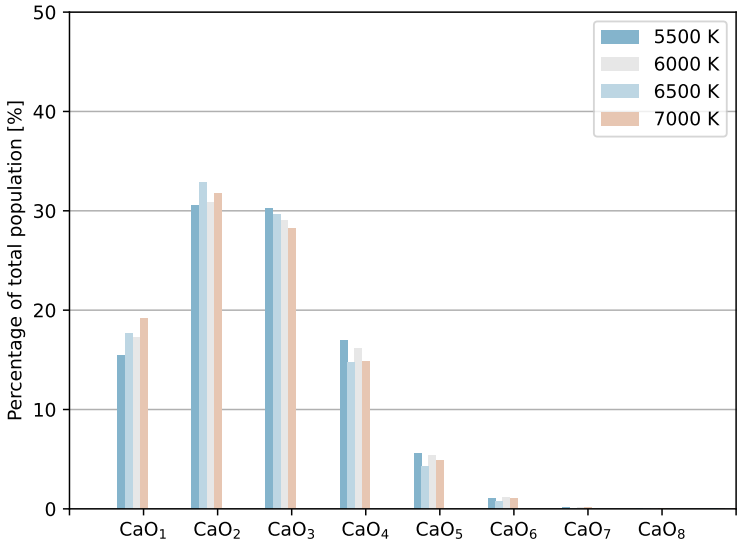


Figure 4.5: The speciation of CaO_x at a density of 0.94 g/cm^3 as a function of temperature. The coordination polyhedra around each atom are obtained using the analysis of the pair distribution function.

Like with MgO we perform the polymer analysis method to study species isolated from the bulk liquid, our incipient vapour. Table 4.1 shows the results for the polymer analysis at 5500 and 6000 K for 0.94 g/cm^3. As briefly described in section 2.4 we believe the configuration of the system to have a non-negligible impact on the polymer analysis method. For this reason, we take four unique configurations, run a simulation with each configuration, and perform the analysis per simulation. In the table we can see that the absolute difference can be quite minor, with some extreme differences in the O species percentage at system 2 at 6000 K. However, these minor

differences cause big relative differences for the minor species. For a more quantitative comparison we have also included the standard deviation for every species in the table.

Comparing the average results with the results obtained for MgO at 6000 K and 0.90 g/cm^3, we see a drastic difference in the percentage of atomic oxygen. If we refer again to the work of Fegley et al. (2016), we find a vapour content of 55% Ca, 35% O, and 10% O_2 at 2000 K. Looking at our results at 6000 K and 0.90 g/cm^3, we find a relative content of 57.2% Ca, 40.5% O, and 2.3% O_2. These results are very similar and the difference in O percentage can be explained by the temperature difference.

	5500 K						6000 K					
	sys1	sys2	sys3	sys4	avg	std. dev.	sys1	sys2	sys3	sys4	avg.	std. dev.
O	18.70	19.95	20.62	20.34	19.90	0.74	18.47	24.74	18.98	17.85	20.01	2.76
O_2	0.14	0.11	0.63	0.29	0.29	0.21	1.08	0.48	0.90	1.99	1.11	0.55
Ca	27.78	25.71	27.89	30.47	27.96	1.69	27.22	27.88	29.07	28.76	28.23	0.73
CaO	16.92	18.29	16.90	19.07	17.80	0.93	16.53	17.31	16.52	17.85	17.05	0.56
CaO_2	3.71	3.27	4.13	2.31	3.35	0.67	3.96	3.22	3.24	3.83	3.56	0.33
CaO_3	0.27	0.44	0.37	0.49	0.39	0.08	0.21	0.22	0.24	0.43	0.28	0.09
CaO_4	0.00	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.06	0.02	0.02	0.02
Ca_2O	4.48	3.59	4.76	4.05	4.22	0.44	5.62	4.27	4.68	3.80	4.59	0.67
Ca_3O	0.15	0.24	0.21	0.42	0.25	0.10	0.13	0.59	0.27	0.21	0.30	0.17
Ca_4O	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
Ca_2O_2	4.56	4.18	4.14	3.39	4.06	0.42	5.01	2.73	3.93	3.89	3.89	0.81
Ca_2O_3	2.09	1.44	1.09	1.19	1.45	0.39	1.95	1.21	1.13	1.79	1.52	0.36
Ca_3O_2	0.87	1.18	1.63	0.89	1.14	0.31	2.32	1.31	1.80	1.55	1.74	0.38

Table 4.1: Speciation analysis (polymer method) of CaO at 0.94 g/cm^3, at 5500 and 6000 K. Four unique configurations were simulated for each temperature, and an average is calculated.

4.3 Oxygen isotope

Oxygen isotopes are an important part of the story for the lunar formation, as discussed in section 1.3.1 we find a large degree of similarity between the Earth and the Moon's oxygen isotopic composition. For this reason, we attempt to study the isotopic fractionation and vapour species formation that may occur during the vaporization of MgO. Previous simulations function as a baseline with 100% ^{16}O. We run two more types of systems, one being 100% ^{18}O and another being an even 50/50 split between ^{16}O and ^{18}O. For these two new system types, we run three unique configurations per system type, each running for 6 ps for a total of 18 ps. In order to mimic ^{18}O in DFT-MD, we use the standard O potential we used before and increase its mass.

	Mg(^{16}O,^{18}O)		**Mg^{16}O**		**Mg^{18}O**
	6000 K	8000 K	6000 K	8000 K	8000 K
^{16}O	7.37	5.52	8.62	11.48	N/A
^{18}O	8.12	6.88	N/A	N/A	12.23
^{16}O-^{16}O	1.29	1.12	8.58	3.89	N/A
^{18}O-^{18}O	0.72	1.02	N/A	N/A	3.82
^{16}O-^{18}O	2.35	1.88	N/A	N/A	N/A
^{18}O-^{18}O-^{16}O	N/A	0.02	N/A	N/A	N/A
^{18}O-^{16}O-^{16}O	0.03	N/A	N/A	N/A	N/A

Table 4.2: Cluster analysis (polymer method) of oxygen in a MgO system. Three systems are studied, all ^{16}O, all ^{18}O, and half/half. All systems are studied at 8000 K, and two of them at 6000 K. We look at molecular oxygen and any ozone that forms. Values are percentages of total lifetime contribution.

We compare all three system types at a cell size of 24x24x24 Å, which is 0.52 g/cm^3 for the Mg^{16}O system, and at 8000 K. Additionally, we look at the Mg^{16}O and Mg(^{16}O,^{18}O) at the same density but at 6000 K. We look at the speciation by analyzing the coordination of oxygen up to and including the first coordination sphere and by doing the polymer method to target vapour species in particular. The results from the polymer speciation are displayed in table 4.2. Mg(^{16}O,^{18}O) is the most interesting system; the other two systems function more as extremes than giving much useful information. We can see that ^{18}O, as a single ion, tends to vaporize more readily than its lighter isotope. This behavior is opposite of what we would expect from what we know of equilibrium isotope fractionation. This indicates that method should be scrutinized closely and more detailed simulations might be necessary. When looking at the molecular oxygen of the three different varieties, no clear pattern emerges. This holds true when comparing the oxygen molecules of the two extreme cases as well. At 6000 K we have only ozone of ^{18}O-^{16}O-^{16}O variety, while at 8000 K it is always ^{18}O-^{18}O-^{16}O.

The average lifetimes for molecular oxygen, in all cases, stays constant around 80 fs. The same holds true for ozone, where the lifetime is around 20 fs. There does not seem to be any noticeable influence from the isoptic difference to species' lifetimes.

	Mg(^{16}O,^{18}O)		**Mg^{16}O**		**Mg^{18}O**
	6000 K	8000 K	6000 K	8000 K	8000 K
^{16}O-^{16}O	118.16	84.64	105.35	79.23	N/A
#	173.00	319.33	1364	4196	N/A
^{18}O-^{18}O	93.04	88.14	N/A	N/A	86.78
#	113.33	260.33	N/A	N/A	1084.33
^{18}O-^{16}O	102.67	81.07	N/A	N/A	N/A
#	148.67	262.00	N/A	N/A	N/A
^{16}O-^{16}O-^{16}O	21.22	22.76	21.97	19.84	N/A
#	4.00	20.00	34	297	N/A
^{18}O-^{18}O-^{18}O	32.90	26.48	N/A	N/A	22.53
#	2.33	6.00	N/A	N/A	87.67
^{18}O-^{18}O-^{16}O	15.75	16.52	N/A	N/A	N/A
#	2.50	8.00	N/A	N/A	N/A

Table 4.3: Average lifetime in fs and count (#) for species obtained using the cluster analysis (polymer method) of oxygen in a MgO system. Three systems are studied, all ^{16}O, all ^{18}O, and half/half. All systems are studied at 8000 K, and two of them at 6000 K. We look at molecular oxygen and any ozone that forms. Both count and lifetime results are averaged from the three configurations to obtain the final results, which is why we have some fractional counts

CHAPTER 5

MgO and machine learning

Contents

In this last chapter, we will discuss our work done using ML potentials in MD. We will briefly walk through the process of applying the ML algorithm and the validation process since especially the latter is an important step. After convincing the reader of the quality of our potential, we will show some results obtained from simulations ran with the ML potential. We will discuss how ML-MD can improve our understanding of the vaporization process and how it enables us to get well-converged viscosity values with relative ease.

5.1 Training

With ML potentials the only thing that really matters is the final potential you make, the potential that you decide to use in your MD. Regardless, we believe it to be worthwhile to briefly discuss our potential creation process. The challenging part of ML-MD lies not in the simulation itself, but in the making of the potential. We want the reader to understand the choices made for the potential. We also want to improve their understanding of the training process.

In order to obtain a good potential it is extremely important that the forces and energies in your training data are highly accurate. Convergence tests we performed indicate, that this cannot be achieved without applying a dense grid of k-points. Our MD simulations from the previous

chapters were all performed using only the gamma point, a 1x1x k-grid. In order to find a balance between computational time and quality, we use a k-spacing of 0.2 Å, VASP then automatically determines the nearest whole integer k-grid it can apply. In practice, this means the k-grid ranges from 2x2x2 at the lowest density and 3x3x3 at the highest. The re-made training data was used to produce a potential that was able to run stable simulations and had an adequate level of quality, i.e. acceptable model validation results. In an attempt to improve the potential, we added training data at an extremely low density, 0.27 g/cm^3. The thought here was that at this low density we would be able to sample more vapour species and improve the potential's affinity for vaporization. Another way we attempted to improve the model was by including training data of different system size. This is a common practice for all types of ML potentials. Besides adding more training data, we also tweaked the settings of the GAP algorithm. We adjusted the values for the cutoff and the default sigmas of the force and the virial. The default sigmas should be set to a value that is similar to the value of one standard deviation of your training data. We found that trying to decrease the default sigma past a certain point would produce nonphysical potentials.

After all the various tests and additions of training data, we found that we were not able to improve the potential to a significant extent when compared to the original. Figure 5.1 shows the 2-body potentials created which, together with the GAP potential, describe the system. The 'v' denotes the version, where v1 is the original potential, v2 is the original training data with the low-density training data included, and v3 is the same training data as v2 with smaller system size training data included. We can see that, for the majority of the distance, the three potentials overlap. The biggest differences are seen in the repulsive part, which should have little effect on the simulation. The repulsive part of the Mg-Mg pair potential shows the biggest difference and the most unphysical behaviour. If the pair potential was realistic it would increase in energy constantly as it gets closer to zero distance, as we see with the other pair potentials. We are not certain why this behaviour occurs, our best guess would be a sampling issue. This error should not have an impact on the simulations run with the potentials, as it is unlikely for an Mg atom to climb up the energy barrier and fall in the local minimum at 0.5 Å. We can check whether this occurs by looking at the PDF and looking for an abnormality in the Mg-Mg PDF. This pair potential only becomes a problem when we decrease the volume of a simulation carelessly. The simulation software we perform out ML-MD with performs periodic remapping on atoms that fall outside of the bounds of the cell. With this method, an Mg atom can be placed close enough to another one to make it fall in the local potential well. We see this in the PDF as shown in figure 5.2. We make sure to always check the PDF of any simulation to verify no Mg fell into the unphysical local potential well.

Lastly, an important part of any ML process is iteration. The iterative process in the case of an ML potential looks like the following; a potential is made and simulations are run with it, some (or all) of the simulations will fail because it encounters an environment the potential cannot describe. You take the points of failure and make additional training data with it and remake the potential. After we made our potentials and ran simulations with them we never encountered any failure, meaning we could not effectively apply the iterative process. Even simulations ran for more than a nanosecond showed no signs of instability. If we did want to apply an iterative process we could have run some simulations parallel to each other and taken out the snapshots where the error is the largest.

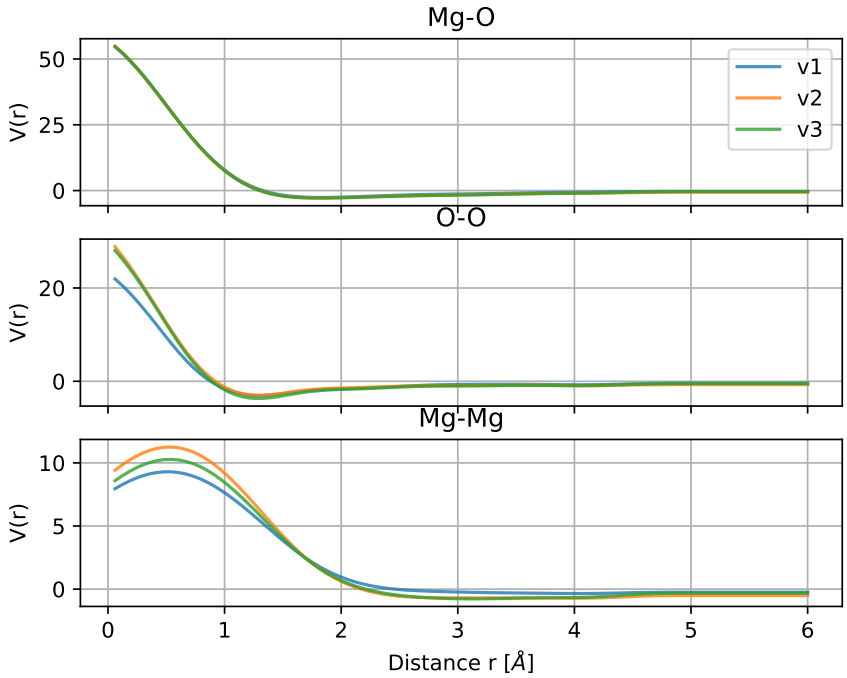


Figure 5.1: 2-Body potentials as calculated with the GAP method. These potentials, together with the GAP, are used to run the ML-MD. The 'v' indicates the version, where v1 consists of the original training data, v2 contains an extra set of low-density training data, and v3 contains an extra training data set of smaller system size on top of the v2 training data. On the whole, the pair potentials are very similar, the biggest differences can be found in the repulsive part of the potential. The repulsive part, 0.5 Ådistance, of the Mg-Mg potential, shows unphysical behaviour of local potential well at 0 Å. This could be caused by poor sampling. It should have no effect on the simulations as long as we are careful not to manually place Mg atoms too close to each other by accident.

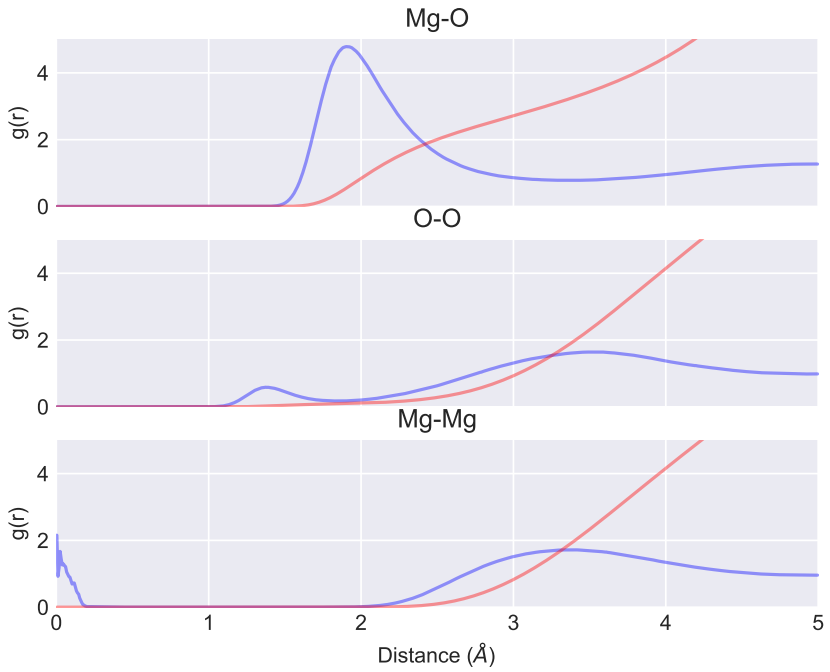


Figure 5.2: Pair distribution function as calculated for a 20 ps simulation of 1.24 g/cm^3 and 6000 K in blue, with the integral in red. In this example the cell volume was carelessly decreased, this allowed the periodical remapping done by the simulation software to place Mg atoms in close proximity to other Mg atoms. This put the Mg atoms in the local potential well, as described in the previous figure, and kept Mg atoms unphysically close to each other.

Density [g/cm^3]	Std. dev.	RMSE
3.29	0.53	8.16
2.63	0.36	0.43
2.14	0.21	0.25
1.76	0.19	0.64
1.23	0.13	0.98
0.9	0.09	0.75
0.78	0.11	0.64
0.67	0.07	0.53
0.59	0.09	0.48
0.27	0.05	0.07
1.9*	0.32	0.38
2.4*	0.39	0.48
2.63*	0.73	0.76
2.14*	0.28	0.29
1.76*	0.29	0.66

Table 5.1: Standard deviation and Root Mean Square Error (RMSE) calculated of the average normale pressures, shown in figure 5.3. *smaller system size

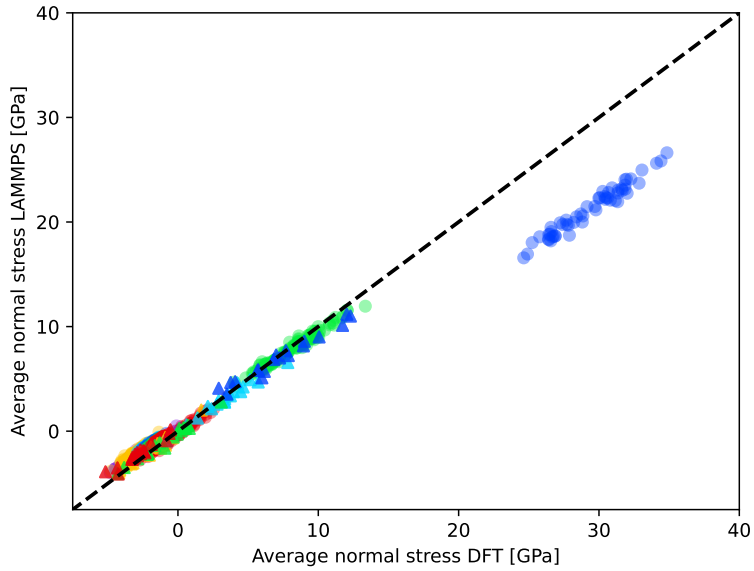
5.2 Validation & testing

After constructing the potential, we applied various methods of validation, the first being to check the average normal pressure. Using the ML potential we perform simulations using LAMMPS (Large-scale Atomic/Molecular Massively Parallel Simulator) with the same configurations that were sampled as the training data. This training data was originally produced using static first-principles calculations in VASP. As the energies converge first, then forces and lastly stresses[1] the normal stress should prove to be a good measure for the quality of the potential. Figure 5.3 shows the resulting values for the average normal stress, comparing the values from the original DFT-MD (VASP) training data and the values obtained with the potential in MD (LAMMPS). We calculate the normal stress by averaging the diagonal components of the stress tensor. We can see that most of the systems plot nicely on the $x = y$ line, except for the highest density one at 3.29 g/cm^3. We are not too worried about this system as most of our curiosity is focused on the liquid-vapour co-existence region at much lower densities. From experience, we know we can fix this offset by including training data of a denser system. We saw a similar kind of offset before at 2.14 g/cm^3 and after including training data at 3.29 g/cm^3 this offset disappeared. For a more quantitative comparison, we have included the standard deviation and the Root Mean Square Error (RMSE) in table 5.1. The standard deviation all fall below 0.8, with an overall average of 0.256. The RMSE are not as favourable, disregarding the 3.29 g/cm^3 data they all fall below 1.0 GPa, with an average of 0.525 GPa. The low standard deviation indicates that the error a system has compared to the true value is quite constant. Most data sets in figure 5.3 fall nicely on a line parallel to $x = y$. Although the RMSE is not that big in absolute terms, relatively speaking it is quite high, especially considering most systems fall between -3.0 and 0.0 GPa.

We also look at the convergence of the forces. To do this, we pick a snapshot at random from

[1] From discussions with Prof. Csányi.

(a)



(b)

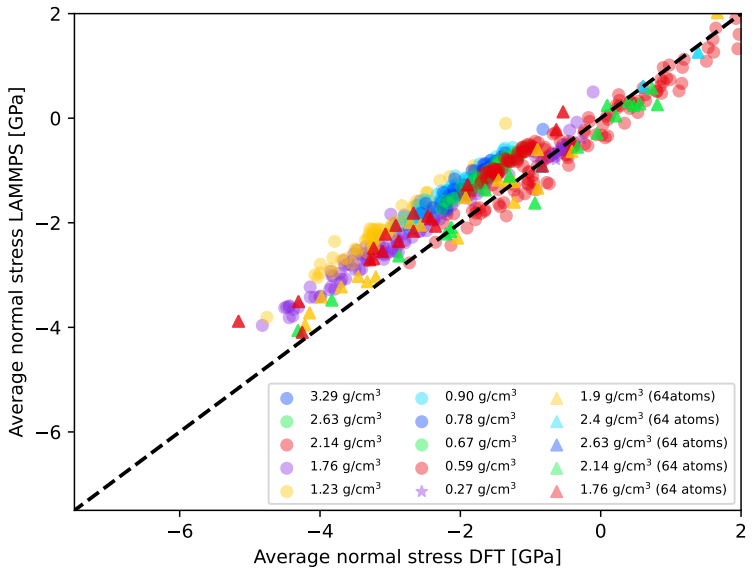
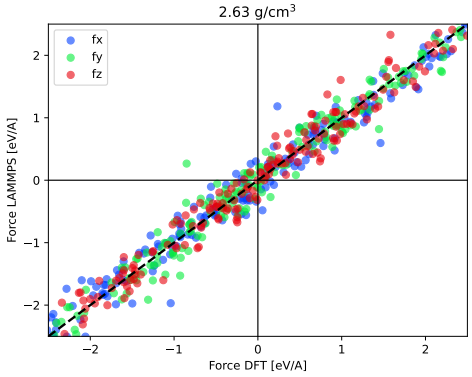


Figure 5.3: Average normal stress as calculated by the original DFT-MD training data and by the ML-MD simulations of the same configuration. Figure a shows all of the data and b shows a zoomed in portion containing the majority of the data. All of the systems fall within a standard deviation of 0.8, the results of the 3.29 g/cm^3 are offset, however. All of the Root Mean Square Errors (RMSE), except for 3.29 g/cm^3, are below 1.0 GPa.

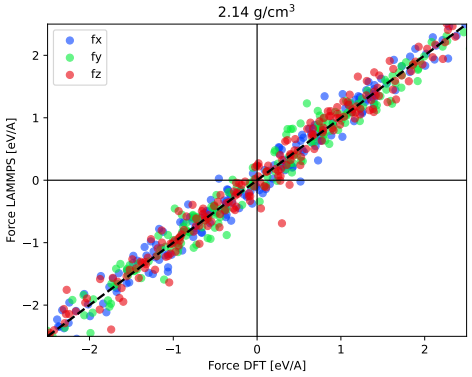
each type of system and compare the x,y, and z force components. This evaluation can be seen in figure 5.5. On the whole, we see that the error in the forces is most significant in the lowest density system, at 0.27 g/cm^3. The smallest error can be found in systems with a small system size. The larger error at 0.27 g/cm^3 can possibly be explained by the lower amount of training data compared to the denser systems. The smaller error found in the smaller system size configurations is likely, at least partially, explained by a denser k-grid. Although the k-spacing is always kept at 0.2 Å, the way VASP determines the k-grid by the nearest integer enabled the smaller system configurations to have 4x4x4 grids (except for the 1.76 g/cm^3 one). While all other 216 system size configurations have 3x3x3 or 2x2x2 grids.

The last type of validation we are showing in this section is the resulting PDF from a longer simulation, figure 5.6. As we are particularly interested in the vaporization and the nature of the liquid/vapour structure, we feel the PDF gives us the most rigid evaluation of the accuracy of the potential.

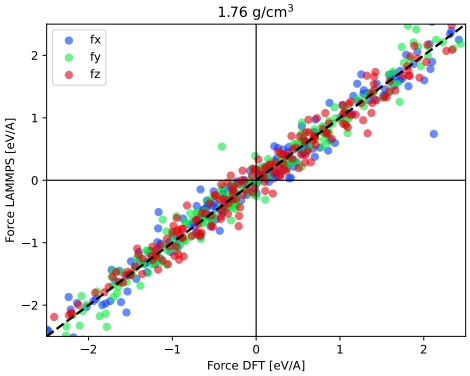
(a)


2.63 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

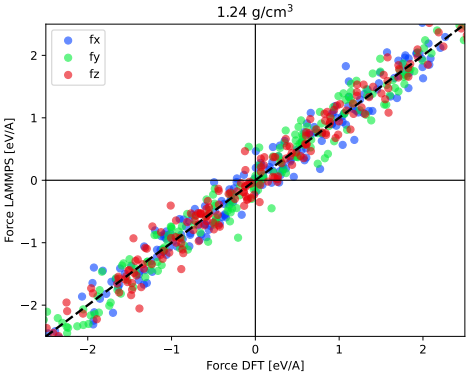
(b)


2.14 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

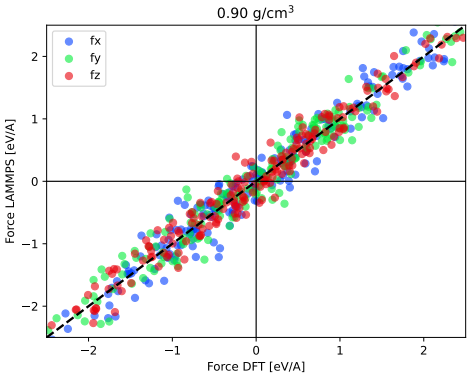
(c)


1.76 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

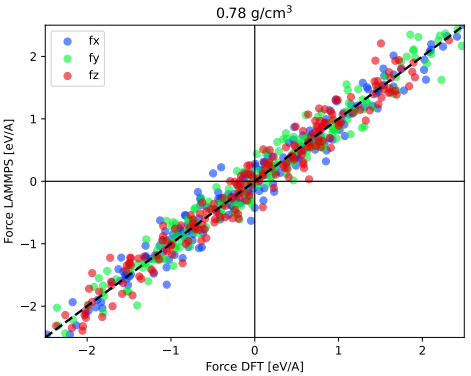
(d)


1.24 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

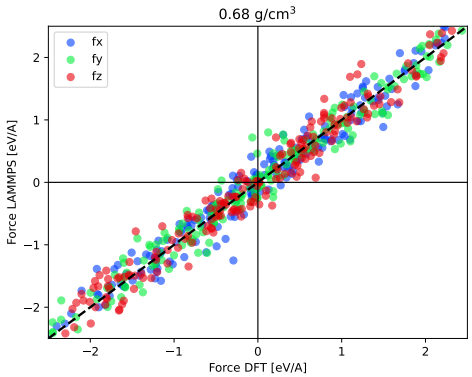
(e)


0.90 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

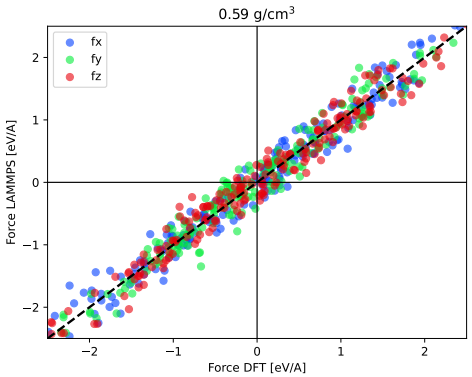
(f)


0.78 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

(g)


0.68 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

(h)


0.59 g/cm^3
fx
fy
fz
Force LAMMPS [eV/A]
Force DFT [eV/A]

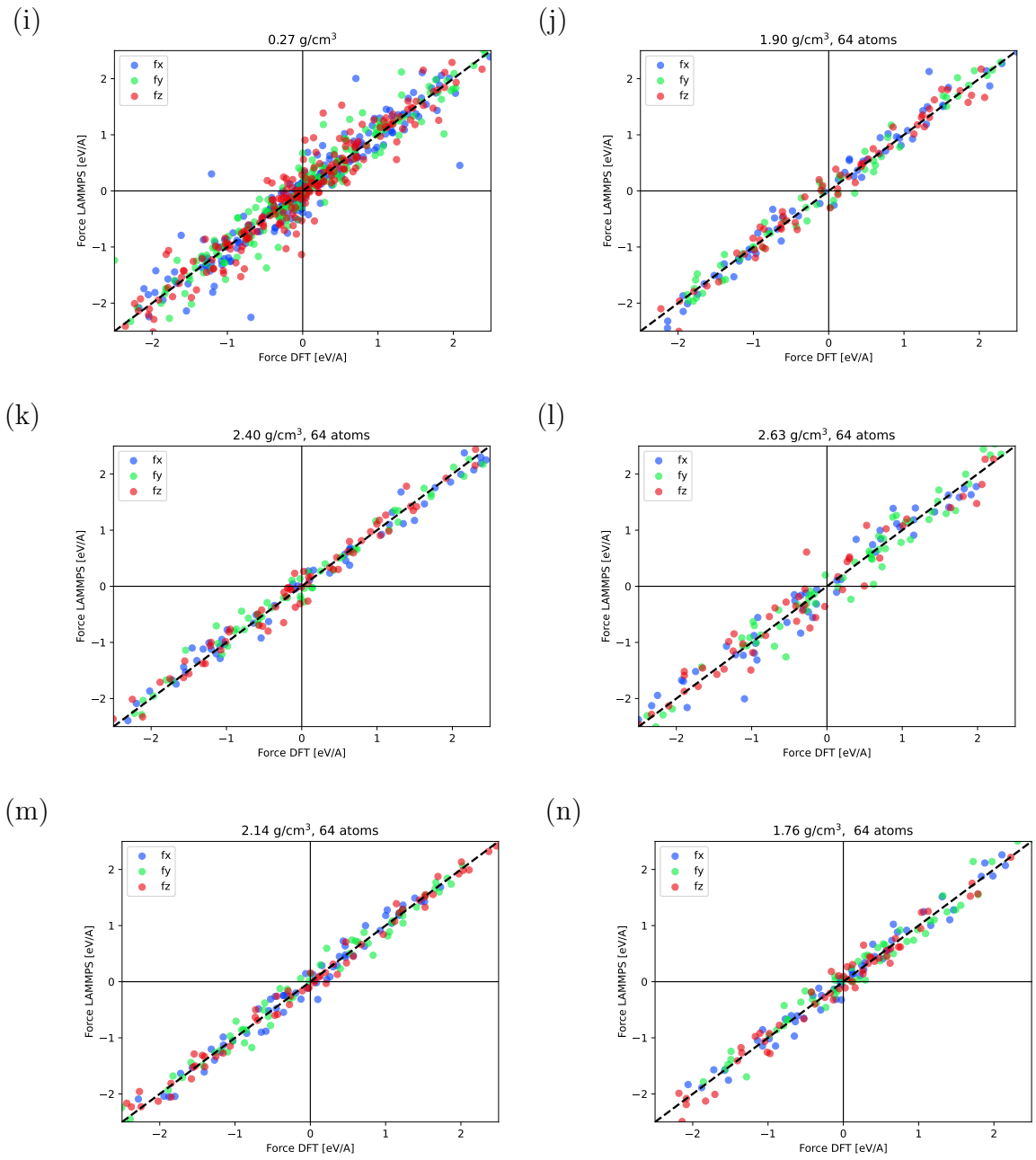
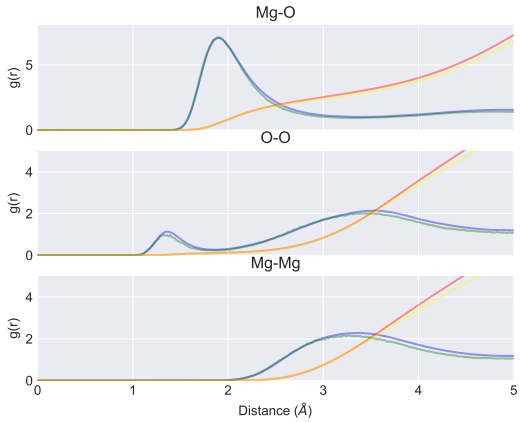
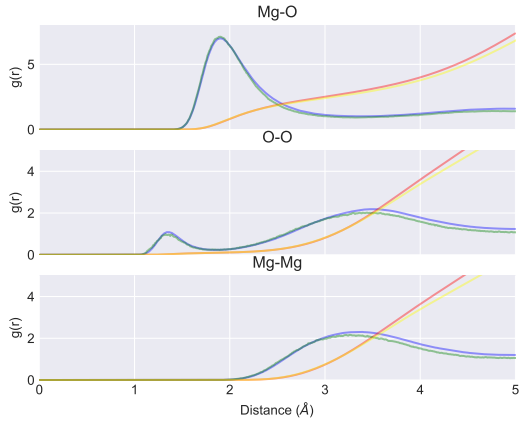


Figure 5.5: Difference in force components x, y, and z for each atom in a random snapshot for each system configuration. Force errors appear the largest in the 0.27 g/cm^3 system and the smallest in the small system sizes.

(a)



(b)



(c)

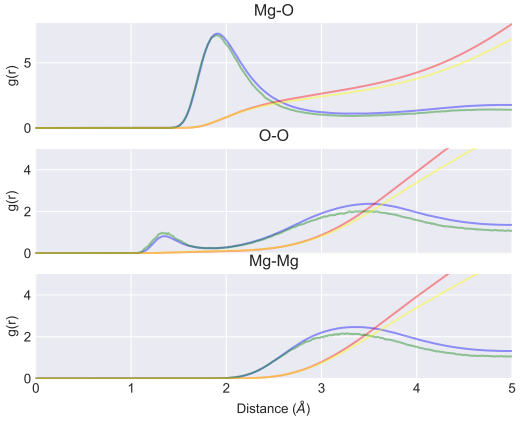


Figure 5.6: Pair distribution functions from simulations ran at 0.59 g/cm^3 and 6000 K. Red and blue lines are calculated with ML-MD, being the PDF and the integral respectively. The green and yellow are the PDF and integral from ML-MD and function as the reference. Figures a, b, and c correspond to simulations run with potentials v1, v2, and v3 respectively.

5.3 Results

In the following section, we will discuss the results obtained from simulations ran with the ML potential. Using the ML potential we are able to scale up our simulations from 216 atoms to 27000 atoms and run for much longer timescales without significant computational cost. In particular, we will take another look at the speciation and discuss the structure of the liquid/vapour system.

5.3.1 Structure

Before we get into the actual results of the speciation analysis, let us talk about how the simulation behaves: What does it look like? Before, when we ran simulations with a system size of 216 atoms, it was always difficult to speak of any real 'vapour' forming. Most of the time you would form a void slab or cylinder through which a handful of isolated molecules and atoms would flow if you were lucky.

Now, with our simulation size of 27000 atoms, we start to look a lot more like a real physical system, a real liquid with vaporous pores forming inside of it. Our maximum simulation cell size measures 120 Å which is of course still a far cry from the nm simulations we see with classical MD, but considering where we came, from it is a major improvement. In order to visualize the LAMMPS simulations results we make use of OVITO. Published by Stukowski (2010), OVITO contains some helpful visualization tools and combining them enables us to perform some valuable visual analyses, next to our numerical analyses. OVITO allows for cluster analysis, based on given bond lengths. Combining this cluster analysis together with a surface mesh allows us to construct visuals that seem liquid-like in their appearance and behaviour. An example of what this mesh does is shown in figure 5.7. This increases the visual clarity and allows us to focus on the more interesting details; the vapour species.

The figures in the last example are made from a simulation of a low-density system, at 0.59 g/cm^3. Naturally, at denser systems, we still expect to find some bubbles containing vapour species. In order to visualize them, we need to make slices of the model and in that way, we can visualize bubbles inside of a bulk liquid, as is shown in figure 5.8. Although this is a labour-intensive process it can still be applied to study individual bubbles when needed. In the bubble in figure 5.8b we see four vapour species; an isolated oxygen and magnesium atom, one MgO particle, and a MgO_2 particle.

5.3.2 Vaporization

We will now examine how our ML-MD performs when it comes to the vaporization process, which is one of our primary interests. We hope to achieve a greater level of accuracy when it comes to predicting the distribution of the vapour species. As we saw in section 4.2.2, our DFT-MD simulations are quite alright at predicting a consistent distribution for the dominant species. The relative differences, however, are quite substantial and this is especially evident with the minor species.

In order to analyse the consistency of the speciation distribution within ML-MD simulations we approach the problem in a similar manner as we did before with CaO. We set up four unique systems, with the same density and temperature but different starting configurations, and perform the analyses on each of these systems. Due to the nature of ML-MD, the amount of data obtained is substantially higher than what we had to post-process with DFT-MD before. For this reason, we make use of a larger timestep of 25 fs, as opposed to 1 fs. Even with a timestep of 25 fs, performing

(a)

(b)

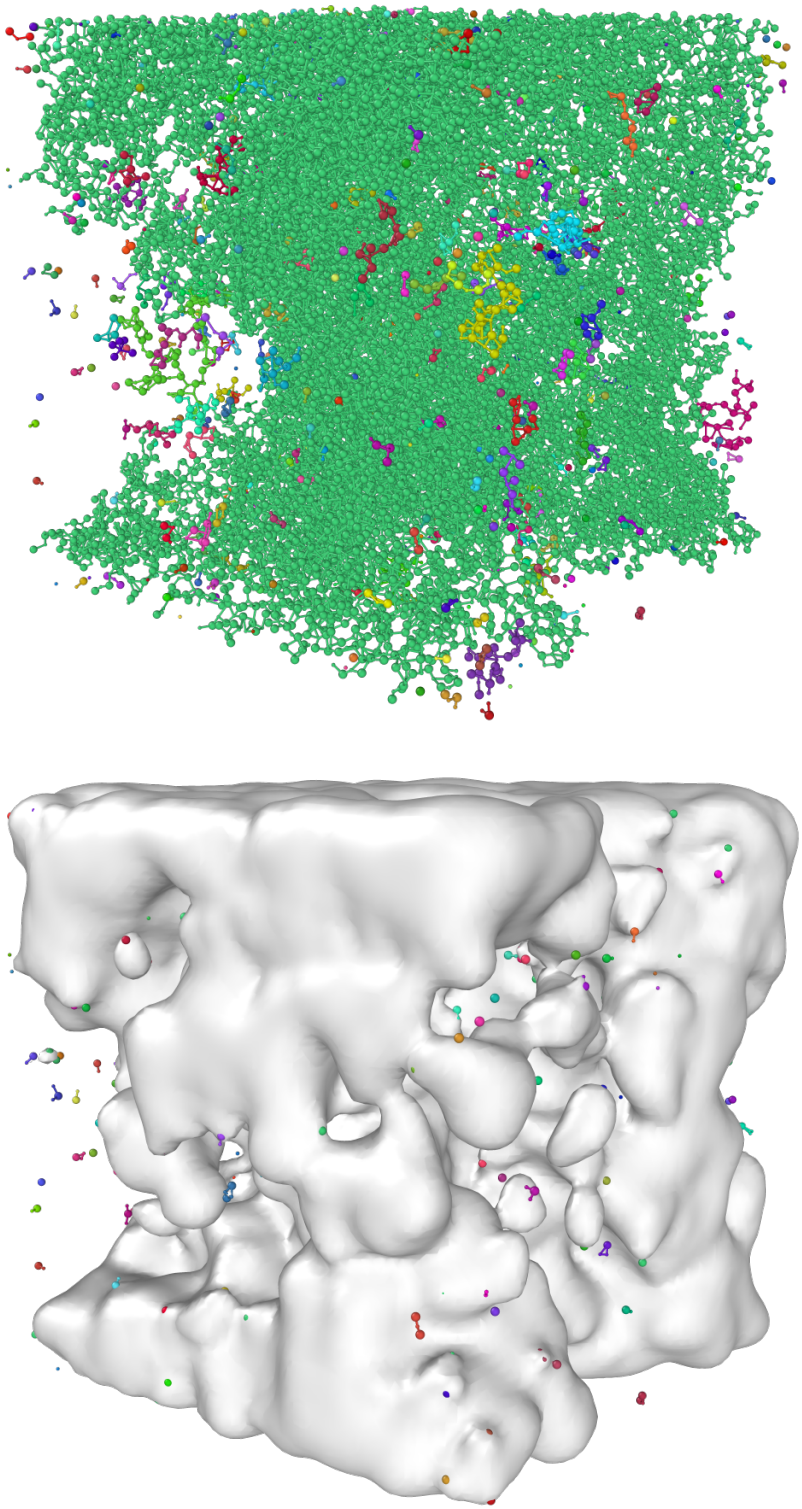
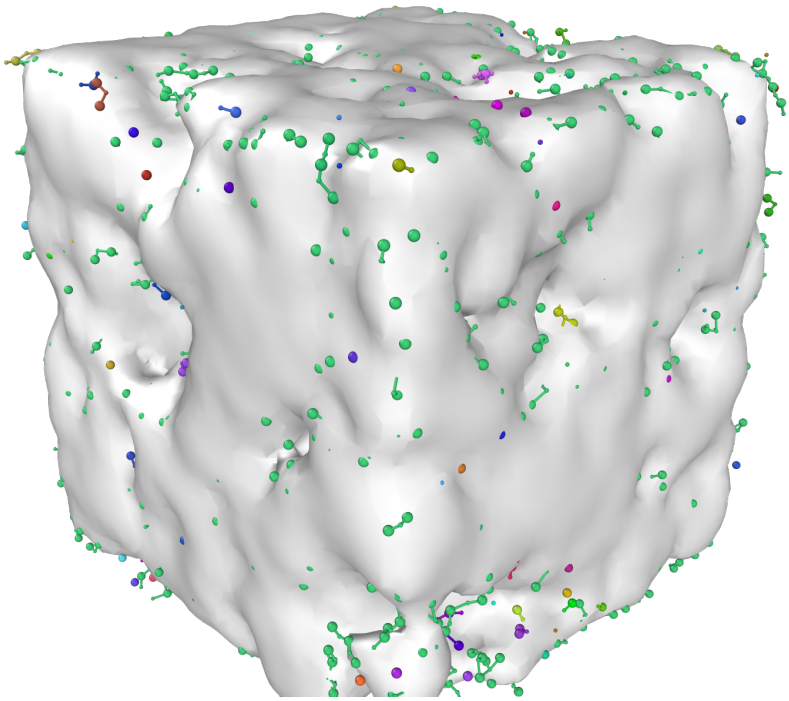


Figure 5.7: Visualization of the simulation in OVITO using a cluster analysis method combined with (b) and without (a) a surface mesh. This improves visual clarity greatly and gives the impression of dealing with a real physical liquid/vapor system.

(a)



(b)

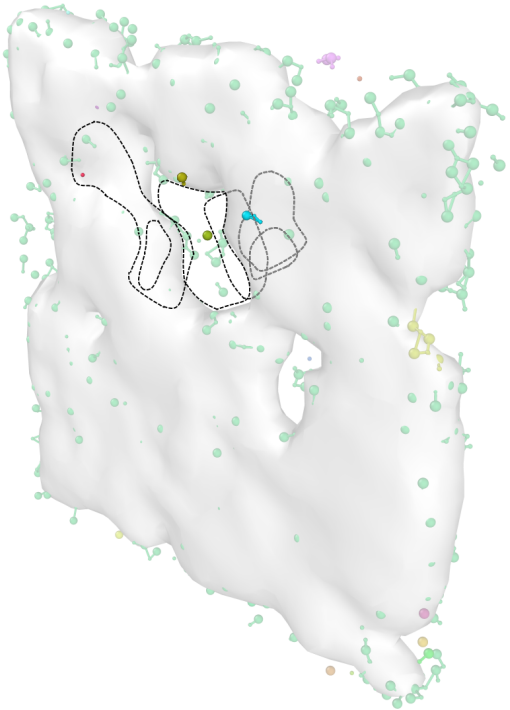


Figure 5.8: Bubbles inside of a bulk liquid, inside of the unit cell, can be inspected using slices. Here we take slices of a single frame (a) and draw the walls of the bubble. Stacking the walls we can get a sense of the morphology of the bubble (b). We also can see what vapour species are contained within, in this case, an isolated oxygen and magnesium atom, one MgO particle, and a MgO_2 particle.

	6000 K, 0.90 g/cm^3					
	sys1	sys2	sys3	sys4	std. dev	CaO std. dev.
O	8.69	9.04	8.71	8.76	0.14	2.76
O_2	6.49	6.97	6.23	6.40	0.27	0.55
Mg	65.13	64.57	65.31	65.02	0.27	0.73
MgO	2.60	2.40	2.66	2.68	0.11	0.56
MgO_2	1.02	1.15	1.07	1.04	0.05	0.33
MgO_3	0.64	0.76	0.66	0.64	0.05	0.09
MgO_4	0.67	0.80	0.66	0.61	0.07	0.02
Mg_2O	2.61	2.38	2.71	2.74	0.14	0.67
Mg_3O	2.74	2.61	2.67	2.61	0.05	0.17
Mg_4O	1.35	1.25	1.24	1.21	0.05	0.00
Mg_2O_2	0.39	0.36	0.40	0.42	0.02	0.81
Mg_2O_3	0.18	0.18	0.18	0.16	0.01	0.36
Mg_3O_2	0.40	0.34	0.38	0.43	0.03	0.38

Table 5.2: Vapor species with a maximum size of five atoms distribution obtained using ML-MD simulations. Three unique systems were constructed and run under the same density and temperature conditions. We show the standard deviation obtained on the right. For a reference we provide the standard deviation obtained for CaO with DFT-MD, from chapter 4.2.2

the speciation post-processing on 200 frames, i.e. 5 ps, takes about 2-3 days because of the poor level of optimization. In table 5.2 you can see the results of these simulations. On the whole, we see that the standard deviation decreases considerably for almost all of the vapour species, with respect to the standard deviation calculated over the four systems of CaO at 6000 K. This confirms that system size plays a large role in in the precision of the obtained vapour species distribution numbers. Due to a mistake in the thermalization process and the subsequent simulations, the actual values obtained are inaccurate. Part of the thermalization was included in the final simulation that was submitted for post-processing, while this part should have been cut out. This means that part if the simulation was poorly mixed. However, because this thermalization is only a smaller part of the whole simulation ±10% this does not significantly influence the level of precision we show with this method. In order to check whether we can reach even higher levels of convergence we perform longer time-scale simulations, the results of which are shown in table 5.3. We use three unique systems, running each system for 0.2ns with a timestep of 500fs. Meaning that every 0.1 ns contains the same number of frames as before, 200. Again we see a sharp decrease in the standard deviation obtained when compared to the previous short-time-scale results. Combining large-scale simulations over long time scales greatly improves the level of precision we obtain for the vapour species distribution.

With the knowledge of the aforementioned tests we perform 200 ps simulations for every system and analyze the coordination and polymer speciation. Results of the coordination are shown in figure 5.9. We see some minor differences at 5000 K, where DFT-MD predicts a somewhat lower overall degree of coordiation compared to ML-MD. At 6000 K we see nearly identical results comparing ML-MD and DFT-MD. The fact that the ML-MD corresponds better to DFT-MD at higher temperatures is likely related to the fact that at higher temperatures the liquid tends to become more homogenized. The structure of the liquid at 6000 K is simpler, the distribution is flatter,

	6000 K, 0.59 g/cm^3							
	sys1		sys2		sys3			
	0.0-0.1ns	0.1-0.2ns	0.0-0.1ns	0.1-0.2ns	0.0-0.1ns	0.1-0.2ns	avg.	std.dev.
O	13.19	12.78	13.07	12.59	13.07	12.59	12.88	0.06
O_2	2.12	2.01	1.96	1.98	1.96	1.98	2.00	0.08
Mg	28.64	28.65	28.23	28.09	28.23	28.09	28.32	0.19
MgO	13.72	13.55	13.63	13.28	13.63	13.28	13.51	0.04
MgO_2	6.99	6.85	6.88	6.85	6.88	6.85	6.88	0.06
MgO_3	0.94	1.05	0.93	0.95	0.93	0.95	0.96	0.01
MgO_4	0.11	0.11	0.09	0.10	0.09	0.10	0.10	0.01
Mg_2O	5.75	5.78	5.79	6.02	5.79	6.02	5.86	0.02
Mg_3O	0.32	0.28	0.33	0.33	0.33	0.33	0.32	0.01
Mg_4O	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.00
Mg_2O_2	6.49	6.70	6.69	6.91	6.69	6.91	6.73	0.09
Mg_2O_3	2.76	2.74	2.70	2.74	2.70	2.74	2.73	0.03
Mg_3O_2	2.29	2.32	2.30	2.24	2.30	2.24	2.28	0.00

Table 5.3: Vapor species with a maximum size of five atoms distribution obtained using ML-MD simulations. Three unique systems were run for 0.2 ns each with a timestep of 500 fs. The 0.2 ns were split up into two parts in order to improve processing speed. The standard deviation is calculated over the values of the 0.0-0.1ns results.

than at 5000 K. We do not rule out that this difference is caused by a poorly constructed training data set. The fact that the results appear quite similar means that structure of the fluid stays the same no matter what size the system is, which is to be expected. It also tells us that for sampling the coordination a small system size would be sufficient. Looking at the vapor species by using the polymer speciation analysis we obtain the percentages shown in tables 5.4 and 5.5. Figure 5.10 shows the percentage vapor distribution per isotherm. Comparing our ML-MD results to our DFT-MD results, several differences can be identified. The most prominent being a significant increase in the amount of isolated oxygen in the vapor species distribution. In general we note an increase in the amount of vapour in the system, this increase is much more noticeable at higher densities. One possible mechanism that could facilitate vapor species production is a change in how superficial tension acts at different system sizes. A larger system could decrease the height of the energy barrier required to overcome superficial tensions and break out of the fluid. This is difficult to prove using the simulations we have done so far, and would require more testing. At the lowest densities the vapor species distributions from ML-MD and DFT-MD become more similar, but there still are some noticeable differences.

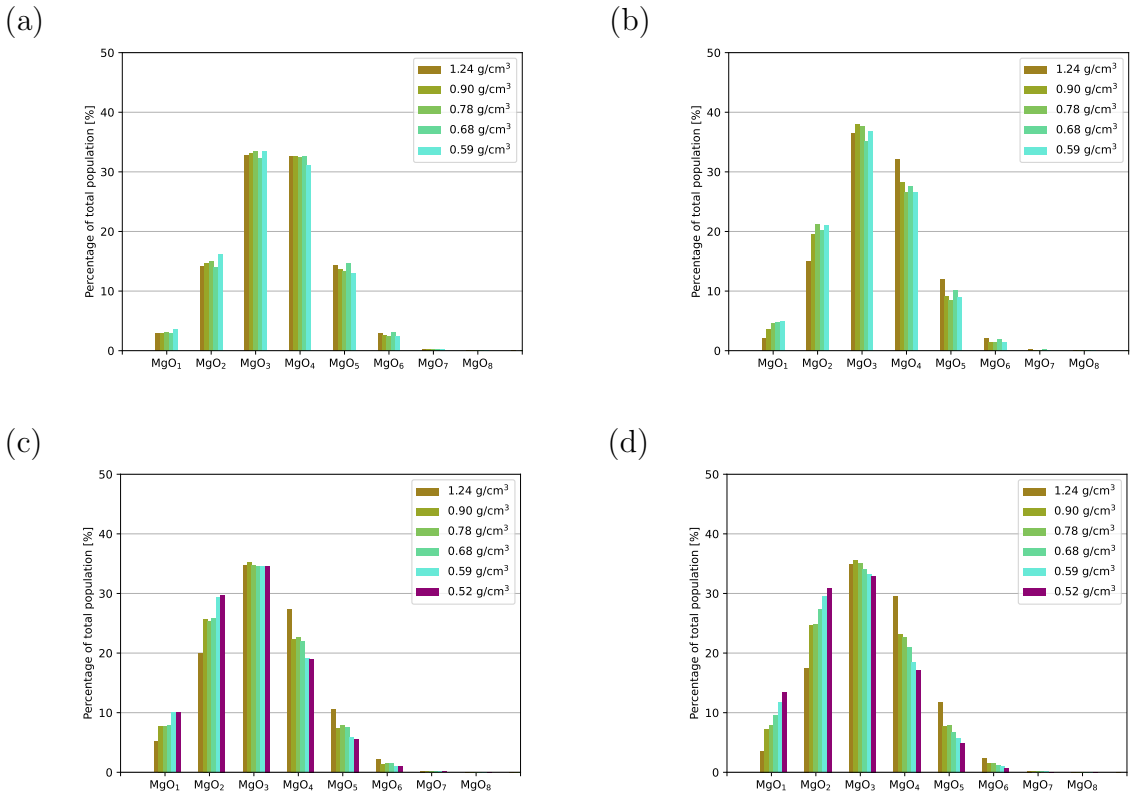


Figure 5.9: (a,c) Coordination number of the liquid in the ML-MD simulations. For comparison the coordination obtained with DFT-MD is given in the right column (b,d). The top row are simulations along the 5000 K isotherm, the bottom row are at 6000 K.

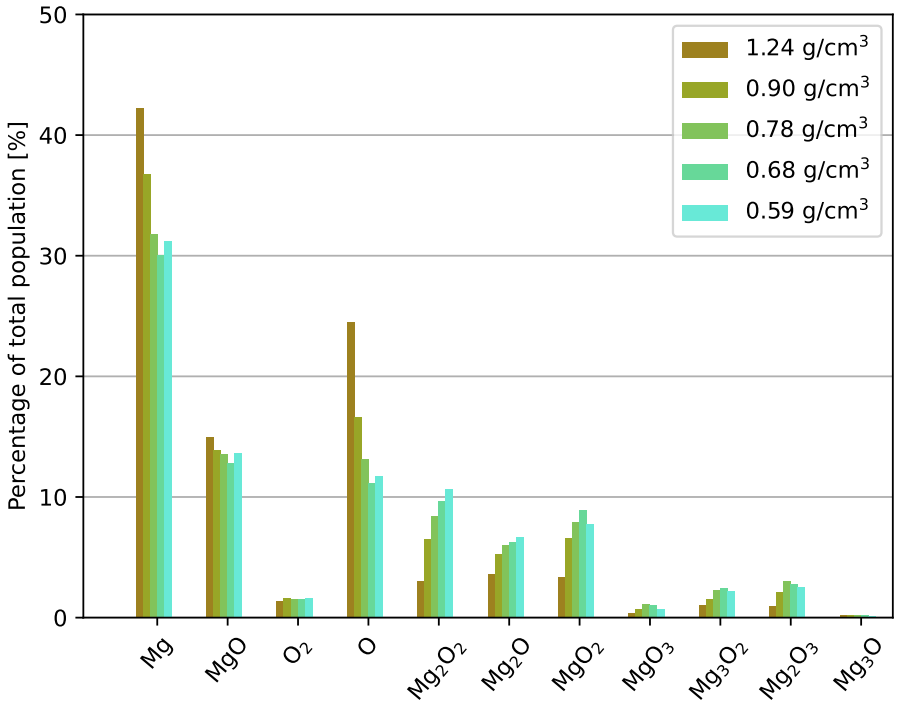
Species \ Density [g/cm^3]	Cluster percentage [%]				
	1.24	**0.90**	**0.78**	**0.68**	**0.59**
O	24.48	16.58	13.17	11.13	11.73
O_2	1.33	1.60	1.52	1.57	1.63
Mg	42.20	36.74	31.82	30.04	31.25
MgO	14.94	13.89	13.58	12.80	13.62
MgO_2	3.34	6.55	7.89	8.90	7.72
MgO_3	0.38	0.70	1.10	1.07	0.73
MgO_4	0.02	0.03	0.07	0.04	0.13
Mg_2O	3.57	5.29	6.00	6.29	6.65
Mg_3O	0.21	0.18	0.24	0.22	0.11
Mg_4O	0.00	0.00	0.00	0.01	-
Mg_2O_2	2.99	6.53	8.44	9.66	10.64
Mg_2O_3	0.97	2.13	3.05	2.77	2.54
Mg_3O_2	1.07	1.57	2.27	2.43	2.18

Table 5.4: Cluster lifetime percentage contribution to the total lifetime of all clusters in the system for every system on the 5000 K isotherm. Simulations using ML-MD of a 27000 atom system for 200 ps, sampled every 2500 fs.

Species \ Density [g/cm^3]	Cluster percentage [%]					
	1.24	**0.90**	**0.78**	**0.68**	**0.59**	**0.52**
O	24.32	19.49	16.19	13.54	13.65	10.85
O_2	1.48	2.11	2.39	2.11	2.12	1.97
Mg	48.53	38.43	33.19	30.19	28.37	26.03
MgO	14.88	15.51	14.36	13.97	13.52	13.15
MgO_2	2.12	4.69	6.07	6.56	6.91	7.58
MgO_3	0.19	0.60	0.92	0.92	1.05	1.10
MgO_4	0.03	0.05	0.10	0.09	0.11	0.11
Mg_2O	3.29	5.10	5.34	5.84	6.04	6.02
Mg_3O	0.31	0.37	0.35	0.40	0.32	0.31
Mg_4O	-	0.01	0.01	0.01	0.01	0.01
Mg_2O_2	1.81	3.48	5.31	6.11	6.24	7.29
Mg_2O_3	0.26	1.36	2.10	2.62	2.66	3.02
Mg_3O_2	0.52	1.41	1.83	2.12	2.30	2.51

Table 5.5: Cluster lifetime percentage contribution to the total lifetime of all clusters in the system for every system on the 6000 K isotherm. Simulations using ML-MD of a 27000 atom system for 200 ps, sampled every 2500 fs.

(a)



(b)

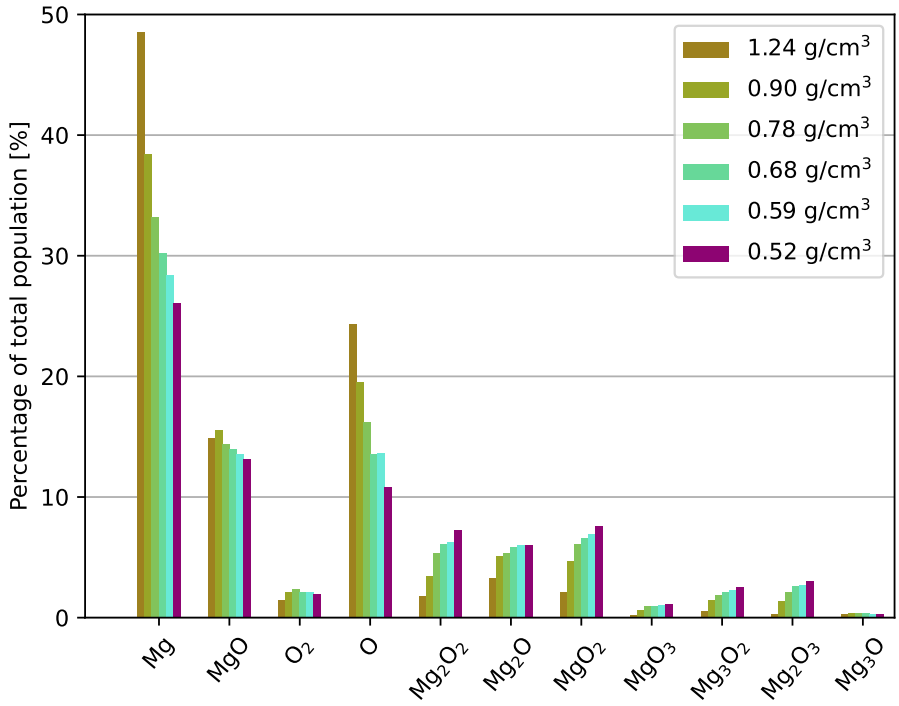


Figure 5.10: Polymer speciation analysis of the 5000 K (top) and 6000 K (bottom) isotherms. Vapor species are put in the same order as the one used in 3.3, for ease of comparisson.

5.3.3 Viscosity

As we stated before, ML enables us to run much longer simulations in a fraction of the time it would take in DFT-MD. One thermodynamics property that benefits greatly from longer simulation time is viscosity. The viscosity is calculated from the auto-correlation function of the off-diagonal values of the stress tensor, in equation form eq. 3.1. Where σ_{ij} is an off-diagonal stress tensor value, τ is the lag in the auto-correlation, V is the volume, T is the temperature and k_B is the Boltzmann constant. As the equation shows we are required to calculate the ACF first and then take its integral. In an ideal world, the correlation goes down to zero and stays there for an infinite τ. However, in real systems there is a tail, noise if you will, that disrupts the signal of the integral, meaning that the integral has to be cut off at a certain *tau*.

$$\eta = \frac{V}{3k_BT} \sum_{i,j} \int_0^\infty \langle \sigma_{ij}(t+\tau)\sigma_{ij}(t) \rangle \, d\tau \tag{3.1}$$

In order to illustrate this, together with comparing DFT-MD and ML-MD results, we show some results of the integral values of the auto-correlation with respect to *tau* in figure 5.11. The coloured lines are the XY, XZ, and YZ components, while the black is the average. We can nicely see how the individual components, and subsequently the average, becomes more converged with respect to one another as the simulation length is increased.

	Viscosity [Pa s]		
Density [g/cm^3]	4000 K	5000 K	6000 K
2.63	6.7×10^{-4}	5.4×10^{-4}	4.7×10^{-4}
2.14	3.5×10^{-4}	2.8×10^{-4}	2.4×10^{-4}
1.76		1.6×10^{-4}	1.4×10^{-4}

Table 5.6: Viscosity values obtained for all-liquid systems at 4000 K, 5000 K and 6000 K. Values were obtained using the shear-stress ACF, taking the average integral value over the 100-250 fs interval and applying eq. 3.1.

Besides simulation length, we can again look at how system size impacts the integral values of the shear stress ACF. In figure 5.12 we show two results; a 216-atom system ran for 1 ns and the other a 27000-atom system ran for 0.2 ns, both are ML-MD at the same density and temperature. The 0.2 ns took about 2.5 times as long to complete as the 1 ns one. From these results, it is evident that the derived integral is much better converged over long time scales than larger systems. We do note that the two simulations provide a similar value for the average integral at lower τ values, which is the most important part. To calculate the viscosity we will use the average of the average integral over a τ range of 100-250 fs.

After performing the shear-stress ACF on 1 ns simulations of all applicable systems we obtain values for the viscosity, summarized in table 5.6. Now, if we compare these values with a recent work by Karki et al. 2019 (see figure 5.13) we can see that we are quite close, our values are about 4×10^{-4} Pa s lower. In their work they also perform NVT DFT-MD simulations, with 100-120 atoms and temperatures ranging between 6000 and 2000 K. They perform their simulations around zero pressure conditions and use a cut-off energy of 450 eV (slightly lower to our 550 eV). We seem to consistently underestimate the viscosity compared to their obtained values. The simulations in the work from Karki et al. were performed at zero pressure whilst our simulations are at pressures ranging between 17.2 GPa to 0.7 GPa. As pressure positively influences viscosity we would expect our values to be higher than the ones from Karki et al. Even the value closest to zero pressure we have (0.7 GPa at 5000 K) underestimates the viscosity by about 6×10^{-4} Pa s compared to Karki et al's results.

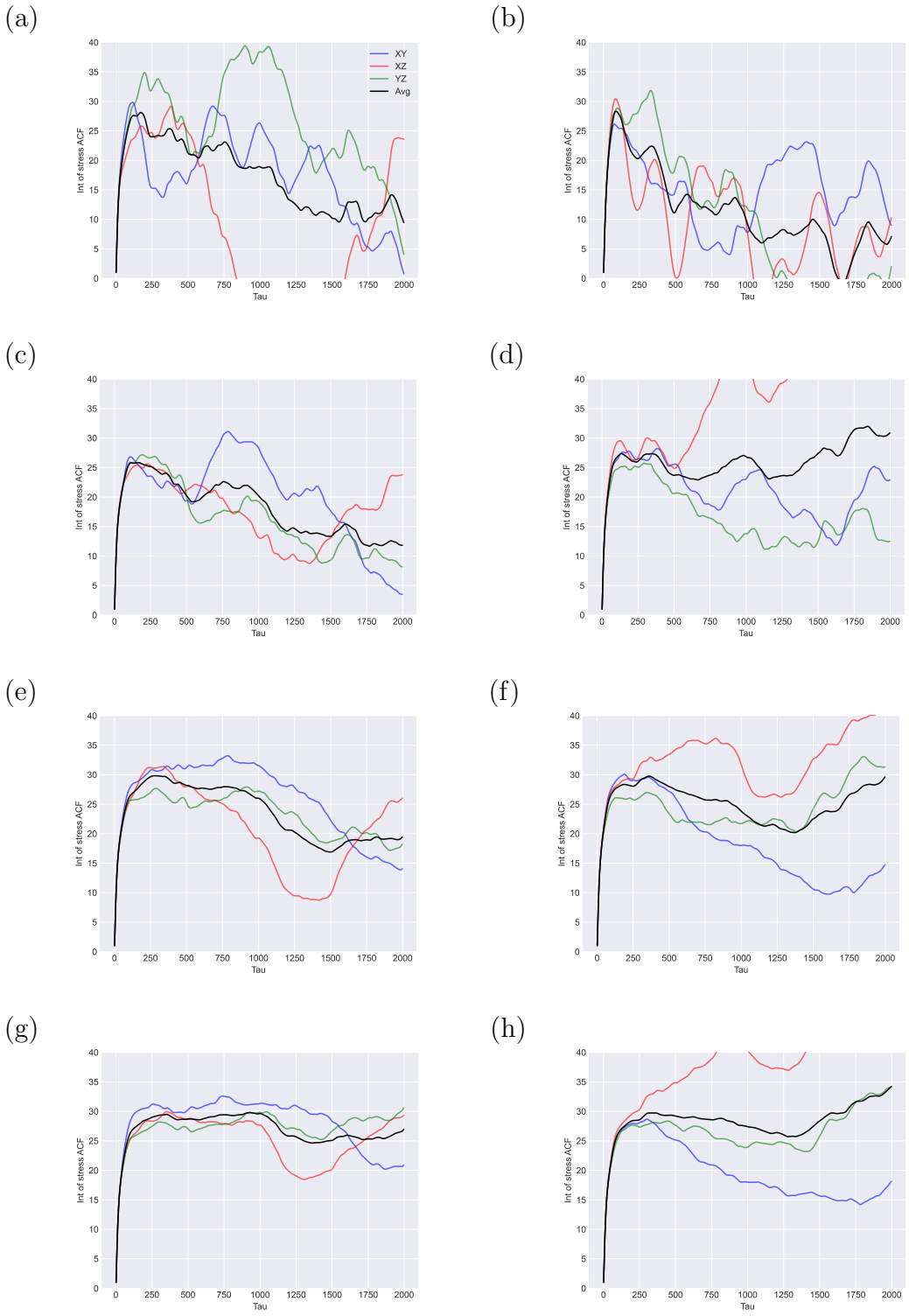


Figure 5.11: Integral of the shear stress ACF determined at four time-lengths. a, b: 5 ps, c, d: 20 ps, e, f: 60 ps, g, h: 100 ps. The left column comes from an DFT-MD simulation while the right column is produced with ML-MD. Both simulations are of a system at 4000 K and 2.14 g/cm^3

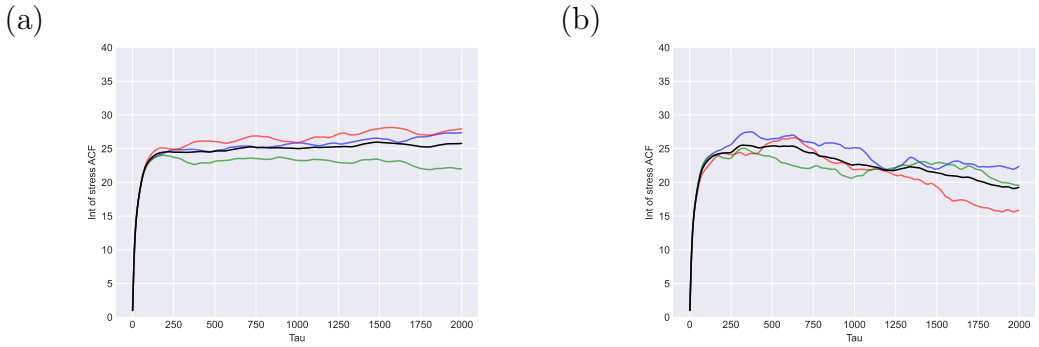


Figure 5.12: Integral of the shear stress ACF determined at two different time and size scales. a: 216-atom system for 1 ns, b: 27000-atom system for 0.2ns. Both systems are at 5000 K and 2.63 g/cm^3. We see a better level of convergence in the integrals of the shear stress ACF at longer simulation times. The simulation that produced the right figure took 2.5 times longer than the simulation of the left figure.

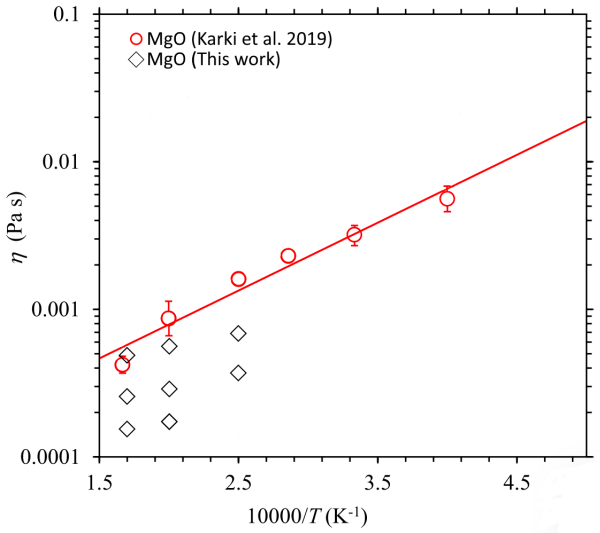


Figure 5.13: Viscosity values as a function of temperature as calculated by Karki et al. (2019) in red with the values obtained here in black added for comparison. Figure adapted from Karki et al. (2019)

CHAPTER 6

Conclusion

Contents

6.1 Refractory oxides

In this body of work we study MgO and CaO, we determine their critical point and speciation behavior. We determine the critical point of MgO be between 6500 K and 7000 K and 0.45-0.6 gm/cm^3. For CaO we estimate the critical point to be situated between 6000 K and 6500 K and a density range of 0.5-0.7 g/cm^3. In terms of coordination we see a much bigger change in coordination for MgO at a temperature range of 5000 K to 7000 K than what we find for CaO, whose coordination distribution remains nearly constant through all temperatures analyzed. Both MgO and CaO vaporize congruently and we find that CaO greatly favors isolated oxygen in the vapor species in comparison to the MgO vapor species fraction. By putting MgO in the context of all MgO-SiO_2 materials that have been studied before, we get a sense how the MgO/SiO_2 ratio affects the critical point, coordination, and vapour species fractions. The critical temperature increases as MgO/SiO_2 increases, this relationship does not appear to be linear. In terms of vapor species fraction SiO and O_2 are much more dominant in the magnesium-silicates than in MgO, where Mg is by far the dominant vapor species.

6.2 ML potentials

ML potentials is a relatively young field in MD and is still very much evolving. Within this manuscript, we try to apply ML-MD to field for which it has not been used before, the liquid/vapor system. We show how applying ML-MD can greatly increase the level of precision of the vapor species distribution. A great level of convergence in this case does not necessarily mean a high level of accuracy. We know the ML potential we use is not perfect, so we can not expect the results to be of exact chemical accuracy, but they should not be far off either. We also still have the question as to why our ML-MD simulations predict a significantly larger amount of vapor than our DFT-MD simulations, which is especially noticeable at higher densities. We provide a possible explanation but definitive proof would require more work.

Defining a chemical bond in MD is an interesting subject. In static DFT one could analyze a specific bond when doing surface chemistry by looking at charge density differences or by analyzing the Born effective charges (or Bader charges). In MD we do not have this ability as the number of bonds is orders of magnitude larger and the bonds change every couple of time-steps. For this reason we have to rely on a statistical approach to define bonding, for the time being this is the best way we can do it. Whilst the study of vaporization can make great use of both the time and size scale increase enabled by ML-MD, we see that the size has a much smaller positive effect on the stress ACF than the time length of the simulation.

By performing a rough calculation we can estimate the net gain (or loss) in computational time by comparing DFT-MD and ML-MD. We will take the v1 potential for the ML-MD as this version already proved to be quite robust and the subsequent iterations did not add much to its accuracy. The training data used to construct the v1 potential was made using only 350 CPU hours, the GAP algorithm itself also adds a couple more hours so we can say 400 in total. This number excludes the time required to obtain the initial configurations that are used for the high-accuracy training data. Now if we compare this to only the 5000 K isotherm calculated using DFT-MD we use 560 CPU hours. Keep in mind that these CPU hour values are not actual wall time, which would depend on the amount of nodes/CPUs used. For the sake of ease, we will assume the HPC job settings used are the same. We can run 200 ps or 27000 atoms with the ML in about 10 CPU hours and a 1 ns simulation of 216 atoms takes only 3 hours. We can use this ML potential for more than one isotherm, so the comparison of computational time becomes even more skewed in favour of ML-MD when including the 4000 and 6000 K isotherm. We can quickly see how ML-MD is computationally much more efficient than DFT-MD, but it does come at a slight accuracy cost when it comes to the stress and forces.

One nagging thought that remains is how these results would compare to ones obtained with a classical potential. An ML-MD is very similar to classical MD in many ways, so it would be interesting to see how, in particular, the vaporization results would look with a classical potential. From Bartók et al. 2018 we know that the GAP should outperform classical interatomic potentials on average, but the GAP can be beaten by classical potentials designed for specific uses. We should also look at other ML potential methods like DeePMD and ACEsuit and do proper comparisons between these methods to identify their respective strengths and weaknesses. As the field is still relatively young, there has not been a lot of effort yet comparing ML potential methods with each other. This can make it difficult for newcomers in the field to make an educated choice of method.

6.3 Implications

Most modern day lunar formation models concerning giant impacts are simulated using modified hydrodynamic codes like HERCULES and GADGET-2. These codes rely on one or multiple equations of state in order to accurately calibrate the materials involved in the simulation. Benz et al. (1989) state that the effect of the equation of state on the amount of mass that gets thrown into orbit, to form the proto-lunar disk, is not significant. However, the equation of state does have a significant influence on how this mass is distributed in the disk. Most modern SPH simulations are ran using some form of the ANalaytical Equation Of State (ANEOS).

ANEOS was developed originally in the 1970s by Thompson and Lauson (Thompson (1970) and Thompson and Lauson (1974)) at Sandia National Laboratories. ANEOS is a set of equation of state routines to be applied in hydrodynamics simulations. Over the past decades several iterative improvements have been made to ANEOS to make it more accurate. One of the more noteworthy improvements of the ANEOS includes the one from Thompson (1990), onto which most modern ANEOS appear to be built. Some of the most seminal publications concerning the modern giant impact hypothesis, performed their SPH simulations using an adaptation of ANEOS by Melosh (2007). This modified ANEOS, commonly refered to as M-ANEOS, is used to obtain an EOS of forsterite (Mg_2SiO_4). The majority of the giant impact simulations give the proto-Earth a chemical composition of roughly 1/3 iron and 2/3 forsterite.

There are some other equations of state applied in giant impact simulations, like the Tillotson (1962) equation of state in some of the more earlier works in the field and more recently a hard-sphere equation of state as described by Hosono et al. (2019). Hosono et al. suggest using a hard-sphere equation of state to simulate the liquid in the system because the ANEOS (and M-ANEOS as a result) are essentially constructed for solids. They note that in solids the internal energy is the main contributor to the Helmholtz free energy, while realistically in liquid systems the configurational entropy is the dominant contributor to the Helmholtz free energy. In systems where this holds true, the degree of heating upon shock compression is much higher.

Recent developments in ANEOS by Stewart et al. (2020) improve the vapour curve and temperatures in the liquid field. This was done by adding a user-defined parameter, the specific heat capacity. This new ANEOS seems to be an improvement over the older one, although it is missing some solid phases at elevated pressures. More recently in Stewart et al. (2022) this new ANEOS was applied to a pyrolitic material ($Ca_{0.87}Fe_{2.03}Mg_{20.22}Al_{1.98}Si_{16.27}O_{58.63}$) in order to better reflect the mantle composition during giant impact simulations.

Sadly, it is not as straightforward to implement the equation of state data presented in this manuscript to an SPH simulation. The parameters used to construct the ANEOS, for example, are not the same parameters we are able to obtain with our DFT-MD simulations. However, one way we can use our results is for comparison, like experimental data in a sense. Although our DFT-MD results might not be as accurate as some experimental results closer to ambient conditions, they can still act as valuable points of reference. From our simulations we able to describe the phase of the system for a wide range of P-T-ρ conditions. We give information about both dynamical and structural properties. In the case of the vapour phase we describe relative amounts of a variety of complex vapour species. With this knowledge we should be able to make adjustments to an ANEOS to make it as realistic as possible. Take for instance Lock and Stewart (2017), in this SPH simulation the critical point of forsterite is put at 8810 K and 1.68 g/cm^3. We know now from DFT-MD simulations by Townsend et al. (2020) that this is a gross overestimate from a realistic critical point, both in temperature and in density.

The exploration of the liquid-vapour phase space using DFT-MD is a relatively new endeavor and it will take some time before the equations of state used in SPH simulations will able to make full use of the results obtained using DFT-MD simulations. Perhaps due to this very reason publications using DFT-MD to study important geological materials in the liquid-vapour space are frustratingly vague as to the implications of the results. In the vast majority of the works these implications do not go further than a few words about how the low-density, high-temperature material properties are of interest for studies related to giant impacts. This manuscript is also somewhat lacking in this regard, so we have attempted to describe the history and role of equations of state applied to SPH simulations. We also provide a vision of how the results of DFT-MD simulations could be applied to improve the accuracy of equations of state and the SPH simulations to which they are applied. One promising item of note are collaborations between members of the SPH simulations community and the DFT-MD community, examples of this are Stewart et al. (2020) and Caracas and Stewart (2023).

6.4 Future work

One major facet that should be worked on, if we want to continue to work with ML-MD, is the post-processing. Currently, using UMD the post-processing is done on a single CPU and often times the scripts build matrices of sizes that are unnecessary. This leads to very slow processing times and high amounts of memory wastage. The way forward, as we already discussed internally, is to parallelize the post-processing and simplify connectivity matrices with binary ones.

In order to understand how our work on refractory oxides can have an implication on the formation of the Moon, it would require our results to be included in lunar formation models. On their own, we present novel material science results on important geo and industrial materials. If we want to know how these new findings can influence lunar formation models, more complex SPH models should be constructed.

ML-MD is the future of MD and it will be interesting to see what new ML schemes are developed the coming years. The scientific community should take great care when working with this powerful tool. It is relatively easy to produce results, but assessing and proving the quality of these results can be difficult but it is of vital importance.

Acknowledgements

First and foremost I have to thank Razvan for giving me this opportunity and guiding me throughout the project. I thoroughly enjoyed this adventure and would do it all over again without hesitation. I am grateful to Johanna for supporting me during my work and pushing me to finish the manuscript on time. I know I was not the most pleasant person to be around the last months so I admire your patience and compassion. I wish my old colleagues of the ERC IMPACT project and ENS Lyon all the best, I enjoyed all the lunches and coffees we shared and learned an immense amount from all of you. I would also like to acknowledge all the help from the support staff at PSMN, CBP, LGL-TPE, TGCC, and IDRIS, who all helped in moving the project along at times of technical or administrative hiccups. I feel fortunate to have a family who always supported me and nurtured my curiosity. Friends back home and in France helped me keep a healthy work-life balance which kept me sane. Although our cat Haricot does not seem the brightest, he still was a great support during the writing process by providing timely distractions. I know I have a lot to be thankful for and I feel a deep sense of gratitude for everything that happened to me.

Thank you. Bedankt. Merci.

Bibliography

Alder, B. J. and T. E. Wainwright (1957). "Phase transition for a hard sphere system". In: *The Journal of Chemical Physics* 27(5), pp. 1208–1209. ISSN: 00219606. DOI: 10.1063/1.1743957.

Alfè, Dario (2005). "Melting curve of MgO from first-principles simulations". In: *Physical Review Letters* 94(23), pp. 4–7. ISSN: 00319007. DOI: 10.1103/PhysRevLett.94.235701.

Bader, R. F.W. et al. (1979). "Quantum Topology of Molecular Charge Distributions". In: *Journal of the American Chemical Society* 101(6), pp. 1389–1395. ISSN: 15205126. DOI: 10.1021/ja00500a006.

Barr, Amy C. (2016). "On the origin of Earth's Moon". In: *Journal of Geophysical Research: Planets* 121(9), pp. 1573–1601. ISSN: 21699097. DOI: 10.1002/2016JE005098. URL: http://doi.wiley.com/10.1002/2016JE005098.

Bartók, Albert P. and Gábor Csányi (2015). "Gaussian approximation potentials: A brief tutorial introduction". In: *International Journal of Quantum Chemistry* 115(16), pp. 1051–1057. ISSN: 1097461X. DOI: 10.1002/qua.24927. arXiv: 1502.01366.

Bartók, Albert P. et al. (2010). "Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons". In: *Physical Review Letters* 104(13), pp. 1–4. ISSN: 00319007. DOI: 10.1103/PhysRevLett.104.136403. arXiv: 0910.1019.

Bartók, Albert P. et al. (2018). "Machine Learning a General-Purpose Interatomic Potential for Silicon". In: *Physical Review X* 8(4), p. 41048. ISSN: 21603308. DOI: 10.1103/PhysRevX.8.041048. arXiv: 1805.01568. URL: https://doi.org/10.1103/PhysRevX.8.041048.

Behler, Jörg and Michele Parrinello (2007). "Generalized neural-network representation of high-dimensional potential-energy surfaces". In: *Physical Review Letters* 98(14), pp. 1–4. ISSN: 00319007. DOI: 10.1103/PhysRevLett.98.146401.

Benz, W. et al. (1989). "The origin of the Moon and the single-impact hypothesis III". In: *Icarus* 81(1), pp. 113–131. ISSN: 00191035. DOI: 10.1016/0019-1035(89)90129-2. URL: https://linkinghub.elsevier.com/retrieve/pii/0019103589901292.

Blank, Thomas B. et al. (1995). "Neural network models of potential energy surfaces". In: *The Journal of Chemical Physics* 103(10), pp. 4129–4137. ISSN: 00219606. DOI: 10.1063/1.469597.

Blöchl, P.E. (1994). "Projector augmented-wave method". In: *Physical Review B* 50(24), pp. 17953–17979.

Blöchl, Peter E. et al. (1994). "Improved tetrahedron method for Brillouin-zone integrations". In: *Physical Review B* 49(23), pp. 16223–16233. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.49.16223. URL: https://link.aps.org/doi/10.1103/PhysRevB.49.16223.

Bögels, T. F.J. and R. Caracas (2022). "Critical point and supercritical regime of MgO". In: *Physical Review B* 105(6). ISSN: 24699969. DOI: 10.1103/PhysRevB.105.064105.

Born, M and R J Oppenheimer (1927). "Zur Quantentheorie der Molekeln". In: *Annalen der Physik* 389(20), pp. 457–484. URL: http://www.ulb.ac.be/cpm/people/scientists/bsutclif/bornop.pdf.

Calandrini, Eugenio et al. (2021). "Limits of the quasiharmonic approximation in MgO: Volume dependence of optical modes investigated by infrared reflectivity and ab initio calculations". In: *Physical Review B* 103(5), p. 54302. ISSN: 24699969. DOI: 10.1103/PhysRevB.103.054302. URL: https://doi.org/10.1103/PhysRevB.103.054302.

Cameron, A.G.W. and W.R. Ward (1976). "The origin of the Moon". In: *Lunar and Planetary Science Conference* 7.

Canup, Robin M. (2004a). "Dynamics of Lunar Formation". In: *Annual Review of Astronomy and Astrophysics* 42(1), pp. 441–475. ISSN: 0066-4146. DOI: 10.1146/annurev.astro.41.082201.113457.

Canup, Robin M. (2004b). "Simulations of a late lunar-forming impact". In: *Icarus* 168(2), pp. 433–456. ISSN: 00191035. DOI: 10.1016/j.icarus.2003.09.028. URL: https://linkinghub.elsevier.com/retrieve/pii/S0019103503002999.

Canup, Robin M et al. (2020). "Origin of the Moon". In: *New Views of the Moon II*.

Caracas, Razvan and Sarah T. Stewart (2023). "No magma ocean surface after giant impacts between rocky planets". In: *Earth and Planetary Science Letters* 608, p. 118014. ISSN: 0012821X. DOI: 10.1016/j.epsl.2023.118014. URL: https://doi.org/10.1016/j.epsl.2023.118014.

Caracas, Razvan et al. (2021). "Analyzing melts and fluids from Ab initio molecular dynamics simulations with the UMD package". In: *Journal of Visualized Experiments* 2021(175), pp. 1–20. ISSN: 1940087X. DOI: 10.3791/61534. arXiv: 2109.02916.

Clayton, Robert N. and Toshiko K. Mayeda (1996). "Oxygen isotope studies of achondrites". In: *Geochimica et Cosmochimica Acta* 60(11), pp. 1999–2017. ISSN: 00167037. DOI: 10.1016/0016-7037(96)00074-9. URL: https://linkinghub.elsevier.com/retrieve/pii/0016703796000749.

Cohen, Alan J and Roy G Gordon (1976). "Modified electron-gas study of the stability, elastic properties, and high-pressure behavior of MgO and CaO crystals". In: *Physical Review B* 14(10), pp. 4593–4605. ISSN: 0556-2805. DOI: 10.1103/PhysRevB.14.4593. URL: https://link.aps.org/doi/10.1103/PhysRevB.14.4593.

Coppari, F. et al. (2013). "Experimental evidence for a phase transition in magnesium oxide at exoplanet pressures". In: *Nature Geoscience* 6(11), pp. 926–929. ISSN: 17520894. DOI: 10.1038/ngeo1948.

Crameri, Fabio (2021). "Scientific colour maps (7.0.1)". In: *Zenodo*.

Ćuk, Matija and Sarah T. Stewart (2012). "Making the moon from a fast-spinning earth: A giant impact followed by resonant despinning". In: *Science* 338(6110), pp. 1047–1052. ISSN: 10959203. DOI: 10.1126/science.1225542.

Darwin, G.H. (1879). "On the bodily tides of viscous and semi-elastic spheroids, and on the ocean tides upon a yielding nucleus". In: *Philosophical Transactions of the Royal Society of London* 170, pp. 1–35. ISSN: 0261-0523. DOI: 10.1098/rstl.1879.0061. URL: https://royalsocietypublishing.org/doi/10.1098/rstl.1879.0061.

Deng, Hongping et al. (2019). "Primordial Earth mantle heterogeneity caused by the Moon-forming giant impact". In: *arXiv* 887(2), p. 211. ISSN: 23318422. DOI: 10.3847/1538-4357/ab50b9. arXiv: 1904.02407. URL: http://dx.doi.org/10.3847/1538-4357/ab50b9.

Faussurier, Gérald et al. (2009). "Evaluation of aluminum critical point using an ab initio variational approach". In: *Physical Review B - Condensed Matter and Materials Physics* 79(13), pp. 1–7. ISSN: 10980121. DOI: 10.1103/PhysRevB.79.134202.

Fegley, Bruce et al. (2016). "Solubility of Rock in Steam Atmospheres of Planets". In: *The Astrophysical Journal* 824(2), p. 103. ISSN: 0004-637X. DOI: 10.3847/0004-637x/824/2/103. URL: http://dx.doi.org/10.3847/0004-637X/824/2/103.

Fish, Ferol F. (1967). "Angular momenta of the planets". In: *Icarus* 7(1-3), pp. 251–256. ISSN: 10902643. DOI: 10.1016/0019-1035(67)90069-3.

Flyvbjerg, H. and H. G. Petersen (1989). "Error estimates on averages of correlated data". In: *The Journal of Chemical Physics* 91(1), pp. 461–466. ISSN: 00219606. DOI: 10.1063/1.457480.

Franchi, I. A. et al. (1999). "The oxygen-isotopic composition of Earth and Mars". In: *Meteoritics and Planetary Science* 34(4), pp. 657–661. ISSN: 10869379. DOI: 10.1111/j.1945-5100.1999.tb01371.x.

Gerstenkorn, H. (1955). "Über Gezeitenreibung beim Zweikörperproblem". In: *Zeitschrift für Astrophysik* 36, pp. 245–274.

Green, Eleanor C.R. et al. (2018). "Bulk properties and near-critical behaviour of SiO2 fluid". In: *Earth and Planetary Science Letters* 491, pp. 11–20. ISSN: 0012821X. DOI: 10.1016/j.epsl.2018.03.015. URL: https://doi.org/10.1016/j.epsl.2018.03.015.

Grimme, Stefan et al. (2010). "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu". In: *The Journal of Chemical Physics* 132(15), p. 154104. ISSN: 0021-9606. DOI: 10.1063/1.3382344. URL: http://aip.scitation.org/doi/10.1063/1.3382344.

Han, Jiaguang et al. (2008). "Terahertz dielectric properties of MgO nanocrystals". In: *Journal of Physical Chemistry C* 112(45), pp. 17512–17516. ISSN: 19327447. DOI: 10.1021/jp805880p.

Hartmann, W. K. and S. M. Larson (1967). "Angular momenta of planetary bodies". In: *Icarus* 7(1-3), pp. 257–260. ISSN: 10902643. DOI: 10.1016/0019-1035(67)90070-X.

Hartmann, William K. and Donald R. Davis (1975). "Satellite-sized planetesimals and lunar origin". In: *Icarus* 24(4), pp. 504–515. ISSN: 10902643. DOI: 10.1016/0019-1035(75)90070-6.

Haynes, W. M. (2017). *CRC Handbook of Chemistry and Physics*. Ed. by W. M. Haynes et al. 97th ed. CRC Press. ISBN: 9781498754293. URL: https://www.ptonline.com/articles/how-to-get-better-mfi-results.

Henkelman, Graeme et al. (2006). "A fast and robust algorithm for Bader decomposition of charge density". In: *Computational Materials Science* 36(3), pp. 354–360. ISSN: 09270256. DOI: 10.1016/j.commatsci.2005.04.010.

Herndon, J. Marvin (1992). "Nuclear fission reactors as energy sources for the giant outer planets". In: *Naturwissenschaften* 79(1), pp. 7–14. ISSN: 00281042. DOI: 10.1007/BF01132272.

Hohenberg, P. and K. Kohn (1964). "Inhomogeneous Electron Gas". In: *Physical Review* 136(3B). ISSN: 0378-620X. DOI: 10.1007/BF01198136. URL: http://link.springer.com/10.1007/BF01198136.

Hosono, Natsuki et al. (2019). "Terrestrial magma ocean origin of the Moon". In: *Nature Geoscience* 12(6), pp. 418–423. ISSN: 17520908. DOI: 10.1038/s41561-019-0354-2. URL: http://dx.doi.org/10.1038/s41561-019-0354-2.

Jackson, I.N.S. and R.C. Liebermann (1974). "Melting and elastic shear instability of alkali halides". In: *Journal of Physics and Chemistry of Solids* 35(9), pp. 1115–1119. ISSN: 00223697. DOI: 10.1016/S0022-3697(74)80129-0. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022369774801290.

Jacobson, Nathan et al. (2017). "Vaporization coefficients of SiO2 and MgO". In: *Journal of the European Ceramic Society* 37(5), pp. 2245–2252. ISSN: 09552219. DOI: 10.1016/j.jeurceramsoc.2016.12.042. URL: https://linkinghub.elsevier.com/retrieve/pii/S0955221916306926.

Jasperse, J. R. et al. (1966). "Temperature dependence of infrared dispersion in ionic crystals LiF and MgO". In: *Physical Review* 146(2), pp. 526–542. ISSN: 0031899X. DOI: 10.1103/PhysRev.146.526.

Jeanloz, Raymond et al. (1979). "B1-B2 transition in calcium oxide from shock-wave and diamond-cell experiments". In: *Science* 206(4420), pp. 829–830. ISSN: 00368075. DOI: 10.1126/science.206.4420.829.

Kanolt, C.W. (1913). "Melting Points of Some Refractory Oxides". In: *Journal of Chemical Physics* 10.

Kant, Immanuel (1755). *Allgemeine Naturgeschichte und Theorie des HimmelsTitle*.

Karki, Bijaya B. and Renata M. Wentzcovitch (2003). "Vibrational and quasiharmonic thermal properties of CaO under pressure". In: *Physical Review B - Condensed Matter and Materials Physics* 68(22). ISSN: 1550235X. DOI: 10.1103/PhysRevB.68.224304.

Karki, Bijaya B. et al. (2019). "Thermodynamics, structure, and transport properties of the MgO–Al2O3 liquid system". In: *Physics and Chemistry of Minerals* 46(5), pp. 501–512. ISSN: 14322021. DOI: 10.1007/s00269-018-01019-5. URL: http://dx.doi.org/10.1007/s00269-018-01019-5.

Karthika, S. et al. (2016). "A Review of Classical and Nonclassical Nucleation Theories". In: *Crystal Growth and Design* 16(11), pp. 6663–6681. ISSN: 15287505. DOI: 10.1021/acs.cgd.6b00794.

Kaula, W.M. and C.F. Yoder (1976). "Lunar Orbit Evolution and Tidal Heating of the Moon". In: *Lunar and Planetary Science Conference* 7.

Kobsch, Anaïs and Razvan Caracas (2020). "The Critical Point and the Supercritical State of Alkali Feldspars: Implications for the Behavior of the Crust During Impacts". In: *Journal of Geophysical Research: Planets* 125(e2020JE006412).

Kohanoff, J. and N. I. Gidopoulos (2003). *Density functional theory: basics, new trends and applications*. Handbook of Molecular Physics and Quantum Chemistry. ISBN: 0471623741.

Kohn, W. and L. J. Sham (1965). "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Physical Review* 140(4A), A1133–A1138. ISSN: 0031-899X. DOI: 10.1103/PhysRev.140.A1133. URL: https://link.aps.org/doi/10.1103/PhysRev.140.A1133.

Kresse, G. and J. Furthmüller (1996a). "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set". In: *Computational Materials Science* 6(1), pp. 15–50. URL: https://linkinghub.elsevier.com/retrieve/pii/0927025696000080.

Kresse, G. and J. Furthmüller (1996b). "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set". In: *Physical Review B* 54(16), pp. 11169–11186. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.54.11169. URL: https://pubs.acs.org/doi/10.1021/acs.jpca.0c01375https://link.aps.org/doi/10.1103/PhysRevB.54.11169.

Kresse, G. and J. Hafner (1993). "Ab initio molecular dynamics for liquid metals". In: *Physical Review B* 47(1), pp. 558–561. ISSN: 01631829. DOI: 10.1103/PhysRevB.47.558.

Kresse, G. and D. Joubert (1999). "From ultrasoft pseudopotentials to the projector augmented-wave method". In: *Physical Review B* 59(3), pp. 1758–1775. ISSN: 0163-1829. DOI: 10.1103/PhysRevB.59.1758. URL: https://link.aps.org/doi/10.1103/PhysRevB.59.1758.

Kruijer, Thomas S et al. (2015). "Lunar tungsten isotopic evidence for the late veneer". In: *Nature* 520(7548), pp. 534–537. ISSN: 0028-0836. DOI: 10.1038/nature14360. URL: http://www.nature.com/articles/nature14360.

Laplace, Pierre-Simon (1796). *Exposition du système du monde*.

Lennard-Jones, John Edwards (1931). "The proceedings of the physical society". In: *Cohesion* 43(240), pp. 461–482.

Leu, An-Lu et al. (1975). "Properties of Molten Magnesium Oxide". In: *Proceedings of the National Academy of Sciences* 72(3), pp. 1026–1030. ISSN: 0027-8424. DOI: 10.1073/pnas.72.3.1026.

Li, Zhi et al. (2020). "Partial core vaporization during Giant Impacts inferred from the entropy and the critical point of iron". In: *Earth and Planetary Science Letters* 547, p. 116463. ISSN: 0012821X. DOI: 10.1016/j.epsl.2020.116463. URL: https://doi.org/10.1016/j.epsl.2020.116463https://linkinghub.elsevier.com/retrieve/pii/S0012821X20304076.

Lock, Simon J. and Sarah T. Stewart (2017). "The structure of terrestrial bodies: Impact heating, corotation limits, and synestias". In: *Journal of Geophysical Research: Planets* 122(5), pp. 950–982. ISSN: 21699100. DOI: 10.1002/2016JE005239. arXiv: 1705.07858.

Lock, Simon J. et al. (2018). "The Origin of the Moon Within a Terrestrial Synestia". In: *Journal of Geophysical Research: Planets* 123(4), pp. 910–951. ISSN: 21699100. DOI: 10.1002/2017JE005333.

Lock, S.J. et al. (2016). "A NEW MODEL FOR LUNAR ORIGIN: EQUILIBRATION WITH EARTH BEYOND THE HOT SPIN STABILITY LIMIT". In: *47th Lunar and Planetary Science Conference*.

MacDonald, G.J.F. (1964). "Tidal Friction". In: *Reviews of Geophysics* 2(3), pp. 467–541.

Mahoney, Michael W. and Petros Drineas (2009). "CUR matrix decompositions for improved data analysis". In: *Proceedings of the National Academy of Sciences* 106(3), pp. 697–702. ISSN: 0027-8424. DOI: 10.1073/pnas.0803205106. URL: https://pnas.org/doi/full/10.1073/pnas.0803205106.

Manara, D. et al. (2014). "On the melting behaviour of calcium monoxide under different atmospheres: A laser heating study". In: *Journal of the European Ceramic Society* 34(6), pp. 1623–1636. ISSN: 09552219. DOI: 10.1016/j.jeurceramsoc.2013.12.018. URL: http://dx.doi.org/10.1016/j.jeurceramsoc.2013.12.018.

McWilliams, R. S. et al. (2012). "Phase Transformations and Metallization of Magnesium Oxide at High Pressure and Temperature". In: *Science* 338(6112), pp. 1330–1333. ISSN: 0036-8075. DOI: 10.1126/science.1229450. URL: https://www.sciencemag.org/lookup/doi/10.1126/science.1229450.

Mehl, M. J. et al. (1988). "Linearized augmented plane wave electronic structure calculations for MgO and CaO". In: *Journal of Geophysical Research* 93(B7), p. 8009. ISSN: 0148-0227. DOI: 10.1029/JB093iB07p08009. URL: http://doi.wiley.com/10.1029/JB093iB07p08009.

Meijer, R. J. de et al. (2013). "Forming the Moon from terrestrial silicate-rich material". In: *Chemical Geology* 345, pp. 40–49. ISSN: 00092541. DOI: 10.1016/j.chemgeo.2012.12.015. URL: http://dx.doi.org/10.1016/j.chemgeo.2012.12.015.

Melosh, H. J. (2007). "A hydrocode equation of state for SiO2". In: *Meteoritics and Planetary Science* 42(12), pp. 2079–2098. ISSN: 10869379. DOI: 10.1111/j.1945-5100.2007.tb01009.x.

Morgan, John W. and Edward Anders (1980). "Chemical composition of Earth, Venus, and Mercury". In: *Proceedings of the National Academy of Sciences* 77(12), pp. 6973–6977. ISSN: 0027-8424. DOI: 10.1073/pnas.77.12.6973.

Morgan, John W. et al. (1978). "The moon: Composition determined by nebular processes". In: *The Moon and the Planets* 18(4), pp. 465–478. ISSN: 01650807. DOI: 10.1007/BF00897296.

Musella, R. et al. (2019). "Physical properties of MgO at deep planetary conditions". In: *Physical Review B* 99(064110). ISSN: 24699969. DOI: 10.1103/PhysRevB.99.064110.

Nosé, Shuichi (1984). "A unified formulation of the constant temperature molecular dynamics methods". In: *The Journal of Chemical Physics* 81(1), pp. 511–519. ISSN: 00219606. DOI: 10.1063/1.447334.

Ohtani, Eiji (1983). "Melting temperature distribution and fractionation in the lower mantle". In: *Physics of the Earth and Planetary Interiors* 33(1), pp. 12–25. ISSN: 00319201. DOI: 10.1016/0031-9201(83)90003-1.

Pahlevan, Kaveh and David J. Stevenson (2007). "Equilibration in the aftermath of the lunar-forming giant impact". In: *Earth and Planetary Science Letters* 262(3-4), pp. 438–449. ISSN: 0012821X. DOI: 10.1016/j.epsl.2007.07.055. arXiv: 1012.5323.

Payne, M. C. et al. (1992). "Iterative minimization techniques for ab initio total-energy calculations: Molecular dynamics and conjugate gradients". In: *Reviews of Modern Physics* 64(4), pp. 1045–1097. ISSN: 00346861. DOI: 10.1103/RevModPhys.64.1045.

Perdew, John P. et al. (1996). "Generalized Gradient Approximation Made Simple". In: *Physical Review Letters* 77(18), pp. 3865–3868. URL: https://link.aps.org/doi/10.1103/PhysRevLett.77.3865.

Pulay, Péter (1980). "Convergence acceleration of iterative sequences. the case of scf iteration". In: *Chemical Physics Letters* 73(2), pp. 393–398. ISSN: 00092614. DOI: 10.1016/0009-2614(80)80396-4. URL: https://linkinghub.elsevier.com/retrieve/pii/0009261480803964.

R. N. Mcnally et al. (1961). "Laboratory Furnace for Studies in Controlled Atmospheres; Melting Points of MgO in a N¡sub¿2¡/sub¿ Atmosphere and of Cr¡sub¿2¡/sub¿O¡sub¿3¡/sub¿ in N¡sub¿2¡/sub¿ and in Air Atmospheres". In: *Jounal of the American Ceramic Society* 44(10), pp. 491–493.

Reufer, Andreas et al. (2012). "A hit-and-run giant impact scenario". In: *Icarus* 221(1), pp. 296–299. ISSN: 00191035. DOI: 10.1016/j.icarus.2012.07.021. URL: http://dx.doi.org/10.1016/j.icarus.2012.07.021.

Ringwood, A. E. (1991). "Phase transformations and their bearing on the constitution and dynamics of the mantle". In: *Geochimica et Cosmochimica Acta* 55(8), pp. 2083–2110. ISSN: 00167037. DOI: 10.1016/0016-7037(91)90090-R.

Robert, F. et al. (1992). "Oxygen isotopic homogeneity of the Earth: new evidence". In: *Earth and Planetary Science Letters* 108(1-3), pp. 1–9. ISSN: 0012821X. DOI: 10.1016/0012-821X(92)90055-Z. URL: https://linkinghub.elsevier.com/retrieve/pii/0012821X9290055Z.

Root, Seth et al. (2015). "Shock Response and Phase Transitions of MgO at Planetary Impact Conditions". In: *Physical Review Letters* 115(19), pp. 1–6. ISSN: 10797114. DOI: 10.1103/PhysRevLett.115.198501.

Rufu, Raluca et al. (2017). "A multiple-impact origin for the Moon". In: *Nature Geoscience* 10(2), pp. 89–94. ISSN: 17520908. DOI: 10.1038/ngeo2866. arXiv: 1903.02525.

Sanville, Edward et al. (2007). "Improved grid-based algorithm for Bader charge allocation". In: *Journal of Computational Chemistry* 28(5), pp. 899–908. ISSN: 01928651. DOI: 10.1002/jcc.20575. URL: http://doi.wiley.com/10.1002/jcc.20575.

Savitzky, A. and M. J. E Golay (1964). "Smoothing and Differentiation of Data by Simplified Least Squares Procedures". In: *Anal. Chem* 36(8), pp. 1627–1639.

Solomatova, Natalia V. and Razvan Caracas (2021). "Buoyancy and Structure of Volatile-Rich Silicate Melts". In: *Journal of Geophysical Research: Solid Earth* 126(2), pp. 1–14. ISSN: 21699356. DOI: 10.1029/2020JB021045.

Soubiran, François and Burkhard Militzer (2020). "Anharmonicity and Phase Diagram of Magnesium Oxide in the Megabar Regime". In: *Physical Review Letters* 125(17), p. 175701. ISSN: 10797114. DOI: 10.1103/PhysRevLett.125.175701. URL: https://doi.org/10.1103/PhysRevLett.125.175701.

Stevenson, D.J. (1987). "Origin of the moon-The collision hypothesis". In: *Annual review of earth and planetary sciences* 15, pp. 271–315.

Stewart, S. T. et al. (2022). "A Hydrocode EOS for Pyrolitic Mantles and Magma Oceans". In: *53rd Lunar and Planetary Science Conference*.

Stewart, Sarah et al. (2020). "The shock physics of giant impacts: Key requirements for the equations of state". In: *AIP Conference Proceedings*. Vol. 2272, Article 080003. ISBN: 9780735440005. DOI: 10.1063/12.0000946. arXiv: 1910.04687. URL: http://aip.scitation.org/doi/abs/10.1063/12.0000946.

Stukowski, Alexander (2010). "Visualization and analysis of atomistic simulation data with OVITO-the Open Visualization Tool". In: *Modelling and Simulation in Materials Science and Engineering* 18(1). ISSN: 09650393. DOI: 10.1088/0965-0393/18/1/015012.

Sun, X. W. et al. (2010). "The high-pressure melting curve of CaO". In: *Solid State Communications* 150(37-38), pp. 1785–1788. ISSN: 00381098. DOI: 10.1016/j.ssc.2010.07.015. URL: http://dx.doi.org/10.1016/j.ssc.2010.07.015.

Swedenborg, E. (1734). *Opera philosophica et mineralia*.

Tang, W. et al. (2009). "A grid-based Bader analysis algorithm without lattice bias". In: *Journal of Physics Condensed Matter* 21(084204).

Taniuchi, Takashi and Taku Tsuchiya (2018). "The melting points of MgO up to 4 TPa predicted based on ab initio thermodynamic integration molecular dynamics". In: *Journal of Physics Condensed Matter* 30(114003).

Taylor, S.R. (2001). *Solar System Evolution: A New Perspective*. Cambridge University Press.

Thiemens, Maxwell M. et al. (2019). "Early Moon formation inferred from hafnium–tungsten systematics". In: *Nature Geoscience* 12(9), pp. 696–700. ISSN: 17520908. DOI: 10.1038/s41561-019-0398-3.

Thompson, Samuel L. (1990). "ANEOS Analytic Equations of State for shock physics codes input manual". In: *SANDIA Report*.

Thompson, S.L. (1970). *IMPROVEMENTS IN THE CHART D RADIATION-HYDRODYNAMIC CODE I: ANALYTIC EQUATIONS OF STATE.* Tech. rep. Albuquerque, NM, and Livermore, CA (United States): Sandia National Laboratories (SNL). DOI: 10.2172/4125539. URL: http://www.osti.gov/servlets/purl/4125539/.

Thompson, S.L. and H.S. Lauson (1974). *Improvements in the CHART D radiation-hydrodynamic code III: revised analytic equations of state*. Tech. rep.

Tillotson, J. H. (1962). *Metallic Equations of State For Hypervelocity Impact*. Tech. rep.

Touboul, Mathieu et al. (2015). "Tungsten isotopic evidence for disproportional late accretion to the Earth and Moon". In: *Nature* 520(7548), pp. 530–533. ISSN: 0028-0836. DOI: 10.1038/nature14355. URL: http://www.nature.com/articles/nature14355.

Touma, Jihad and Jack Wisdom (1998). "Resonances in the Early Evolution of the Earth-Moon System". In: *The Astronomical Journal* 115(4), pp. 1653–1663. ISSN: 00046256. DOI: 10.1086/300312.

Townsend, Joshua P. et al. (2020). "Liquid-Vapor Coexistence and Critical Point of Mg2SiO4 From Ab Initio Simulations". In: *Geophysical Research Letters* 47(e2020GL089599).

Trinquier, Anne et al. (2009). "Origin of Nucleosynthetic Isotope Heterogeneity in the Solar Protoplanetary Disk". In: *Science* 324(5925), pp. 374–376. ISSN: 0036-8075. DOI: 10.1126/science.1168221. URL: https://www.nature.com/articles/s41550-019-0948-zhttps://www.science.org/doi/10.1126/science.1168221.

Umemoto, Koichiro et al. (2017). "Phase transitions in MgSiO3 post-perovskite in super-Earth mantles". In: *Earth and Planetary Science Letters* 478, pp. 40–45. ISSN: 0012821X. DOI: 10.1016/j.epsl.2017.08.032. arXiv: 1708.04767. URL: http://dx.doi.org/10.1016/j.epsl.2017.08.032.

Vasisht, Vishwas V. et al. (2011). "Liquid-liquid critical point in supercooled silicon". In: *Nature Physics* 7(7), pp. 549–553. ISSN: 17452481. DOI: 10.1038/nphys1993. arXiv: 1103.3473.

Vassent, J.L et al. (2000). "Thermodynamic analysis of molecular beam epitaxy of MgO(s) II. Epitaxial growth of MgO layers on Fe(001) substrates". In: *Journal of Crystal Growth* 219(4), pp. 444–450. ISSN: 00220248. DOI: 10.1016/S0022-0248(00)00632-1. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022024800006321.

Vega, Lourdes et al. (1992). "Phase equilibria and critical behavior of square-well fluids of variable width by Gibbs ensemble Monte Carlo simulation". In: *The Journal of Chemical Physics* 96(3), pp. 2296–2305. ISSN: 0021-9606. DOI: 10.1063/1.462080. URL: http://aip.scitation.org/doi/10.1063/1.462080.

Ward, William R. et al. (2020). "Analytical Model for the Tidal Evolution of the Evection Resonance and the Timing of Resonance Escape". In: *Journal of Geophysical Research: Planets* 125(6). ISSN: 2169-9097. DOI: 10.1029/2019JE006266. arXiv: 2005.00587. URL: https://onlinelibrary.wiley.com/doi/10.1029/2019JE006266.

Wood, D. M. and A. Zunger (1985). "A new method for diagonalising large matrices". In: *Journal of Physics A: General Physics* 18(9), pp. 1343–1359. ISSN: 03054470. DOI: 10.1088/0305-4470/18/9/018.

Wood, J.A. (1986). "Moon over Mauna Loa-A review of hypotheses of formation of earth's moon." In: *Origin of the Moon*.

Xiao, Bing and Lars Stixrude (2018). "Critical vaporization of MgSiO3". In: *Proceedings of the National Academy of Sciences of the United States of America* 115(21), pp. 5371–5376. ISSN: 10916490. DOI: 10.1073/pnas.1719134115.

Yu, Min and Dallas R. Trinkle (2011). "Accurate and efficient algorithm for Bader charge integration". In: *Journal of Chemical Physics* 134(064111).

Zhang, Junjun et al. (2012). "The proto-Earth as a significant source of lunar material". In: *Nature Geoscience* 5(4), pp. 251–255. ISSN: 17520894. DOI: 10.1038/ngeo1429.

Appendix A

Publications

Physical Reviews B (2022): Critical point and supercritical regime of MgO

Critical point and supercritical regime of MgO

T. F. J. Bögels [1,*] and R. Caracas [1,2,3,†]

[1] *Centre National de la Recherche Scientifique, Laboratoire de Géologie de Lyon, LGLTPE UMR 5276, Centre Blaise Pascal, Ecole Normale Supérieure de Lyon, F-69364 Lyon Cedex 07, France*

[2] *The Center for Earth Evolution and Dynamics (CEED), University of Oslo, 0371 Oslo, Norway*

[3] *Université de Paris, Institut de Physique du Globe de Paris, CNRS, Paris, France*

(Received 31 July 2021; revised 9 December 2021; accepted 13 January 2022; published 18 February 2022)

The position of the critical point determines the top of the liquid-vapor coexistence dome, and it is a physical parameter of fundamental importance in the study of high-energy shocks, including those associated with large planetary impacts. For most major planetary materials, such as oxides and silicates, the estimated position of the critical point is below 1 g/cm^3 at temperatures above 5000 K. Here we compute the position of the critical point of one of the most ubiquitous materials: MgO. For this we perform first-principles molecular dynamics simulations. We find the critical density to be in the 0.45–0.6 g/cm^3 range and the critical temperature in the 6500–7000 K range. We investigate in detail the behavior of MgO in the subcritical and supercritical regimes, and we provide insight into the structure and chemical speciation. We see a change in Mg-O speciation toward lower degrees of coordination as the temperature is increased from 4000 to 10 000 K. This change in speciation is less pronounced at higher densities. We observe the liquid-gas separation in nucleating nanobubbles at densities below the liquid spinodal. The majority of the chemical species forming the incipient gas phase consists of isolated Mg and O atoms and some MgO and O_2 molecules. We find that the ionization state of the atoms in the liquid phase is close to the nominal charge, but it almost vanishes close to the liquid-gas boundary and in the gas phase, which is consequently largely atomic.

DOI: 10.1103/PhysRevB.105.064105

I. INTRODUCTION

Giant impacts are a characteristic of the early stages of the evolution of planets, when chaotic trajectories of planets and planetesimals oftentimes intersect. The impact processes can be so energetic that they can produce partial or even total melting and/or vaporization of the bodies involved. The resulting ejecta gather to form a disk. If the temperatures reached during the peak of the impacts exceed the conditions of the critical point (CP), the constituting materials of those celestial bodies become supercritical. The resulting disks would then be monophasic. Upon cooling of a supercritical disk, or in case the supercritical conditions are not reached in the disk during the impact, the system evolves toward a biphase system along the liquid-vapor equilibrium curve. As the maximum of the liquid-vapor equilibrium is fixed by the CP, the position of the CP itself is of fundamental importance in understanding large and giant planetary impacts. And even though giant impacts may be considered rare events, they can be responsible for the creation of planets and moons, or sometimes for their destruction.

Currently, the most widely accepted hypothesis for the formation of our Moon is one such giant impact: a Mars-sized impactor, called Theia, collided with the proto-Earth. This giant impact melted, vaporized, and rendered supercritical a significant portion if not all of the impactor and of the proto-Earth [1–5], creating a large accretion disk from which the Moon formed. Using constitutive equations of the materials involved in the impact, smoothed-particle hydrodynamics (SPH) simulations [6] can describe some of the complex aspects of these giant impacts, and predict the outcome of such planetary impacts as well as the formation of protolunar disks. These SPH simulations are large consumers of data from shock equations of state, like SESAME [7] or ANEOS [8]. Additional meaningful data consist of supercritical points, equations of state of supercritical fluids, and liquid-vaporization equilibria.

Magnesium oxide, MgO, is one of the fundamental building blocks of rocky planets, being a ubiquitous component that appears in most complex silicate minerals. It adds up to roughly 38% [9] of the Earth's and Moon's composition. Deep inside rocky telluric planets, a distinct layer may develop where (Mg,Fe)O forms a mineral by itself, namely magnesiowüstite. In the Earth this layer, containing also (Mg,Fe)SiO_3 bridgmanite, corresponds to the lower mantle, which is the largest part of our planet by volume.

Moreover, MgO is an archetype of countless ionic AB diatomic compounds, many of which are minerals or technological materials. Its face-centered-cubic B1 structure, stable at ambient conditions, represents the most symmetric and most common structure of AB materials.

*tim.bogels@ens-lyon.fr
†caracas@ipgp.fr

Because of its relevance for both planetary sciences and materials science, MgO has been studied extensively, both experimentally and theoretically, over a wide range of pressures and temperatures. At ambient pressure, the melting point lies at 3125 K and 3.6 g/cm^3, and the boiling point is at 3870 K. The first numerical predictions of the melting line [10,11] overestimated the temperatures, while the first experimental results [12] underestimated them. Modern molecular-dynamics simulations predicted the melting of the B1 phase to occur at 3100 K and 0 GPa and at 9400 K and 240 GPa [13].

The high-pressure and high-temperature region was investigated extensively in the past [14–19]. A B1-B2 phase transition was confirmed in an experimental setting [16,17] and in more recent numerical studies [19,20]. The pressure-volume equation of state was measured up to 600 GPa, and the temperature and optical reflectivity to beyond 1400 GPa and 50 000 K in shock experiments [16]. These conditions, while not relevant for the Earth, can be reached inside super-Earth exoplanets, where this phase transition might induce further layering in the rocky mantles of these planets [18].

At the other side of the phase diagram, the liquid-vapor equilibrium line and the position of the critical point are so far almost unknown. This is due to experimental difficulties of sampling both low densities and high temperatures. Theoretical [21] and experimental (molecular beam epitaxy and vacuum thermogravity apparatus) [22,23] studies suggest that the vaporization of MgO crystals is a congruent process, the resulting gas obtained from heating MgO crystals being formed of a stoichiometric mixtures of Mg and atomic O gas. This behavior is expected because of the simple chemistry and stoichiometry of magnesia. But these studies stop short at relatively low temperature, and they fail to reach the CP and to explore the supercritical regime. For the large majority of rock-forming minerals, such as feldspars [24], $MgSiO_3$, and Mg_2SiO_4, the vaporization is incongruent [21,24–26].

Here we characterize the high-temperature low-density region of MgO using *ab initio* molecular dynamics (MD) in the density functional theory (DFT) framework. These conditions cover the conditions of the liquid spinodal and the critical point and expand into the supercritical regime. The paper is organized in four main parts. Following the introduction, Sec. II details the methodology, the simulations, and the postprocessing. Section III discusses in detail the results: the position of the critical point, the structure of the fluids, the transport properties, the vibrational spectrum, and the electronic atomic charges. The paper ends with a short discussion and conclusions.

II. METHODOLOGY

A. First-principles molecular dynamics

We study MgO over a broad range of thermodynamic conditions that cover the low pressures and high temperatures characteristic of the liquid side of the liquid-vapor equilibrium dome. For this, we perform first-principles molecular dynamics (MD) calculations based on density-functional theory (DFT) using the Vienna Ab initio Software Package (VASP) implementation [27–31].

We employ the planar augmented wave function (PAW) [32] flavor of the DFT, with standard PAW pseudopotentials for Mg and O, $3s^2$ with a core radius of 1.06 Å for Mg, and $2s^2 2p^4$ with a core radius of 0.80 Å for O. We use the Perdew-Burke-Ernzerhof formalism of the generalized gradient approximation [33] for the exchange-correlation functional. The mass of the thermostat was set such that the temperature fluctuations have approximately the same frequencies as the typical phonon-frequencies of MgO, and it was not adapted to the density. The temperature of the system is controlled with a Nosé-thermostat [34]. We employ a kinetic energy cutoff of the plane waves of 550 eV, and we sample the reciprocal space in the Γ point. These parameters ensure a precision of the calculations for the energy on the order of 5 meV/at, and for the pressure on the order of 2 kbar.

As is customary in molecular-dynamics simulations of fluids, the systems are modeled using cubic simulation boxes, which are periodically repeated along the three directions of the space. We start the simulations from a $3 \times 3 \times 3$ supercell of MgO with B1 structure, with lattice parameter $a = 4.211$ Å [35]. These supercells contain 216 atoms, i.e., 108 formula units. We heat this supercell up to 5000 K using a heating rate of 0.5 K/fs. We monitor the diffusion of the atoms and find that at this temperature and density, MgO is in a fluid state, i.e., the self-diffusivity of atoms is finite and positive. After thermalization for 1 ps at 5000 K, this configuration constitutes the starting point of our simulations. We go from one temperature to another as needed to follow the different isotherms, increasing or decreasing the temperature in steps of 0.5 K/fs. We sample the density space by changing the unit-cell parameter in steps of 1 Å. The production runs at any given point in pressure and temperature are started after allowing for a thermalization period of 0.5–1 ps. The average duration of the production simulations is on the order of 20 ps for the lower temperatures and higher density systems, and it decreases to a minimum of 5 ps for the high-temperature low-density systems, which are considerably more computationally heavy.

To ascertain the magnitude of the finite-size effects, we ran tests on systems of several densities at 5000 K, with 64 and 512 atoms. The results of these tests can be found in the supplemental material. We find that already at system sizes of 64 atoms, the results start to converge. Our simulations with 216 atoms yield almost the same result, in terms of energy, pressure, and equation of states, as the simulations with 512 atoms. This is consistent with the analysis of finite-size effects in forsterite (Mg_2SiO_4) melts [26], where a system size of 56 atoms already approaches the critical temperature within the accuracy range presented in this research. Furthermore, employing the van der Waals correction for the dispersive forces at the smallest density investigated here caused no appreciable difference for bond lifetimes or pressure.

B. Finding the position of the critical point

We compute the variation of the pressure as a function of density along various isotherms. We start with the high-density simulations, where the stable phase is the liquid. As we decrease the density, the pressure continuously decreases.

Under enough stretching, the pressure passes below the liquid-vapor equilibrium value; the liquid becomes metastable. The metastable region is where the vapor is thermodynamically stable, but the liquid is mechanically stable. Under further expansion, the pressure reaches a minimum, which marks the liquid spinodal. At lower densities, the liquid is unstable and a gas fraction spontaneously separates from the liquid. As the density continues to decrease, the pressure builds up due to the gas phase until the system reaches a local maximum of the pressure, which marks the gas spinodal [36–38]. Between the local pressure minimum and the local pressure maximum, the fluid is a mechanical mixture of gas and liquid, as they are both unstable as a single phase. Upon further expansion, at densities lower than that of the gas spinodal, the gas becomes metastable. The pressure starts to decrease, passes again the gas-liquid equilibrium pressure, and then asymptotically decreases to zero under infinite expansion. At densities lower than the gas-liquid equilibrium line, the stable phase is the gas. This behavior is best described using the van der Waals gas-liquid equilibrium model, which employs third-order polynomials. We approximate this model with a standard cubic polynomial least-squares fit relating the pressure to the density in accordance with other studies [3,39] (see the supplemental material for more details [40]). The cubic function allows us to quickly find the local minimum and maximum of the polynomial, which yields, respectively, the liquid and the gas spinodal points. The curvature of the fit is influenced by the finite-size effect [41]. Upon increasing the system size, the amplitude of the curves will decrease, and the pressures will all become positive. Figure 1 illustrates the aforementioned features of the polynomials and the construction of the spinodal lines. The liquid and the gas spinodal lines intersect in the critical point.

As the liquid is stable at high densities, the MD simulations allow for the determination of the liquid spinodals at all temperatures. But because of limited ergodicity, the simulations cannot be reliably run at very low densities. This prevents covering the gas spinodal points at low temperatures. Only close to the critical point we can extend the simulations over a density range that encompasses both the gas and the liquid spinodals. Consequently, we approximate the van der Waals model with a cubic function of the pressure as a function of density. Then, as stated above, the local minimum yields the liquid spinodal, and close to the critical point, the local maximum yields the gas spinodal. Hence the position of the critical point is bracketed in density between the gas and the liquid spinodal. In temperature, it lies between the last isotherm whose pressure shows a local minimum and a local maximum, and the first isotherm that shows a monotonous decrease of pressure. This procedure was applied with success on various other material in the same phase space [3,24,26,42].

C. Postprocessing

We perform all of the postprocessing of the *ab initio* MD runs using the Universal Molecular Dynamics (UMD) software package [43]. We are analyzing structural, transport, vibrational, thermodynamic, and electronic properties in the subcritical and supercritical regimes.

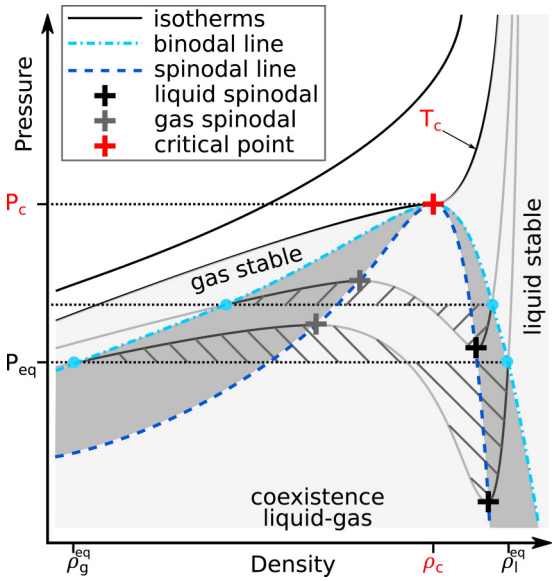


FIG. 1. Construction of the phase stability fields and critical point from pressure-density relations along several isotherms. The maxima and minima of the isotherms coincide with the liquid and gas spinodal points, respectively. The spinodal lines connect these points. The Maxwell construction delimits regions of equal area between the pressure-density curve and the liquid-vapor equilibrium pressure at each temperature. The equal-area regions are represented with hashed fields. The line joining the densities of the vapor, ρ_v^{eq}, in equilibrium with the liquid, ρ_l^{eq}, at all temperatures is the binodal line. Both the spinodal and the binodal line have the critical point as maximum in common. In the area between these lines (dark gray), the liquid and gas phases are metastable. The light gray areas outside of the binodal lines are where the individual phases are stable, and the light gray area between the spinodal lines is the region where the liquid and gas coexistence is stable. The area above the isotherm of the critical temperature is where the supercritical fluid is stable. Figure adapted from Kobsch and Caracas [24].

a. Pair distribution function

The average interatomic bonding and the coordination environment are important structural properties that stem from the analysis of the pair distribution function, commonly referred to as $g(r)$. The pair distribution function describes the relative distribution of atoms as a function of distance. It zeros out at the start since that is the region where atoms are repulsing each other; the distance up to which the distribution remains null defines the interatomic exclusion zone. The first maximum is associated with the most common interatomic distance, oftentimes referred to as the average bond length. The first minimum of the pair distribution function represents the limit of the first coordination sphere. The second minimum is the limit of the second coordination sphere, and so on. In calculations where the liquid is approximated by a periodic box, the applicability of the distribution function is limited by half the size of the edge of the simulation box, in order to avoid artifacts related to the periodicity.

We take the radius of the first coordination sphere as the threshold value of the interatomic bonds: if two atoms lie closer than this radius, they are considered *bonded*. All the bonded ligands to a central atom define the coordination

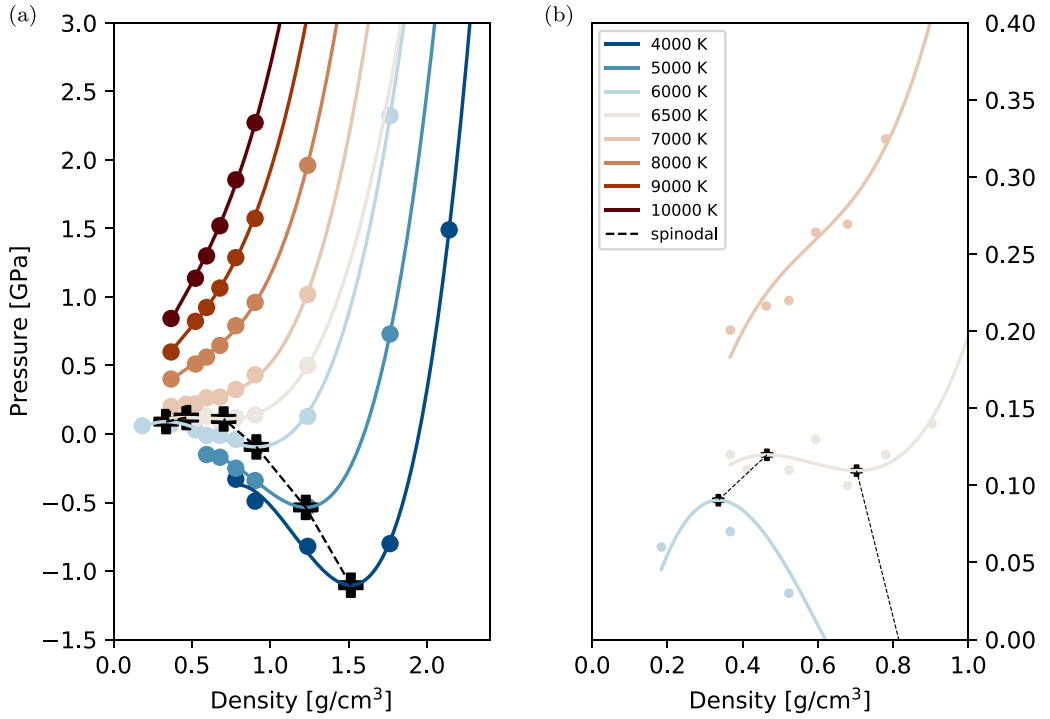


FIG. 2. Variation of the pressure as a function of density for various isotherms (a). The solid lines are cubic function fits. Their local minima and maxima yield, respectively, the liquid and the gas spinodal points, represented with thick black crosses. The spinodal lines are represented with thin dashed black lines. The critical point lies between the liquid and gas spinodals in density, and between the last isotherm where the pressure still reveals local extremes, and the first isotherm where the pressure is monotonously decreasing the decreasing density. For MgO, this places the critical point at 0.45–0.6 g/cm^3 in the density range and between 6500 and 7000 K in the temperature range. The corresponding pressures are on the order of 0.1–0.2 GPa (b).

polyhedra. All the bonded atomic pairs define a connectivity graph, building polymers that describe the structure of the liquid.

b. Mean-squared displacement

The mean-squared displacement (MSD) is the square of the average distance that an atom or cluster of atoms travels as a function of time.

It is calculated using Eq. (1), where N_α is the number of atoms of type α, T is the total time of the simulation, N_{init} is the number of initial times (the number of displacements measured), and τ is the width of the time window,

$$\text{MSD}_\alpha(\tau) = \frac{1}{N_\alpha}\frac{1}{N_{\text{init}}}\sum_{i=1}^{N_\alpha}\sum_{t=0}^{T/2}[r_{\alpha,i}(\tau+t) - r_{\alpha,i}(t)]^2. \tag{1}$$

The slope of the MSD yields the self-diffusion coefficient,

$$D_\alpha = \lim_{\tau\to\infty}\frac{1}{nt}\text{MSD}_\alpha, \tag{2}$$

where $n = 2, 4, 6$ for one, two, and three dimensions, respectively. A positive slope of the MSD is a clear

TABLE I. Computed pressure values for MgO obtained for each isotherm at various densities.

	Pressure (GPa)							
Density	4000 K	5000 K	6000 K	6500 K	7000 K	8000 K	9000 K	10 000 K
3.29 g/cm^3	29.94	35.95	41.56		47.27	52.41	57.59	62.77
2.63 g/cm^3	9.34	13.20	17.22		21.20	24.95	28.55	32.16
2.14 g/cm^3	1.49	4.06	6.77		9.38	12.05	14.66	17.28
1.76 g/cm^3	−0.80	0.73	2.32	3.36	4.17	6.10	8.10	9.93
1.24 g/cm^3	−0.82	−0.53	0.13	0.50	1.02	1.96	3.07	4.10
0.90 g/cm^3	−0.49	−0.34	−0.08	0.14	0.43	0.96	1.57	2.27
0.78 g/cm^3	−0.33	−0.25	−0.04	0.12	0.32	0.79	1.29	1.85
0.68 g/cm^3		−0.17	−0.01	0.10	0.27	0.65	1.07	1.52
0.59 g/cm^3		−0.15	−0.01	0.13	0.26	0.56	0.92	1.30
0.52 g/cm^3			0.03	0.11	0.22	0.51	0.82	1.14
0.37 g/cm^3			0.07	0.12	0.20	0.40	0.60	0.84

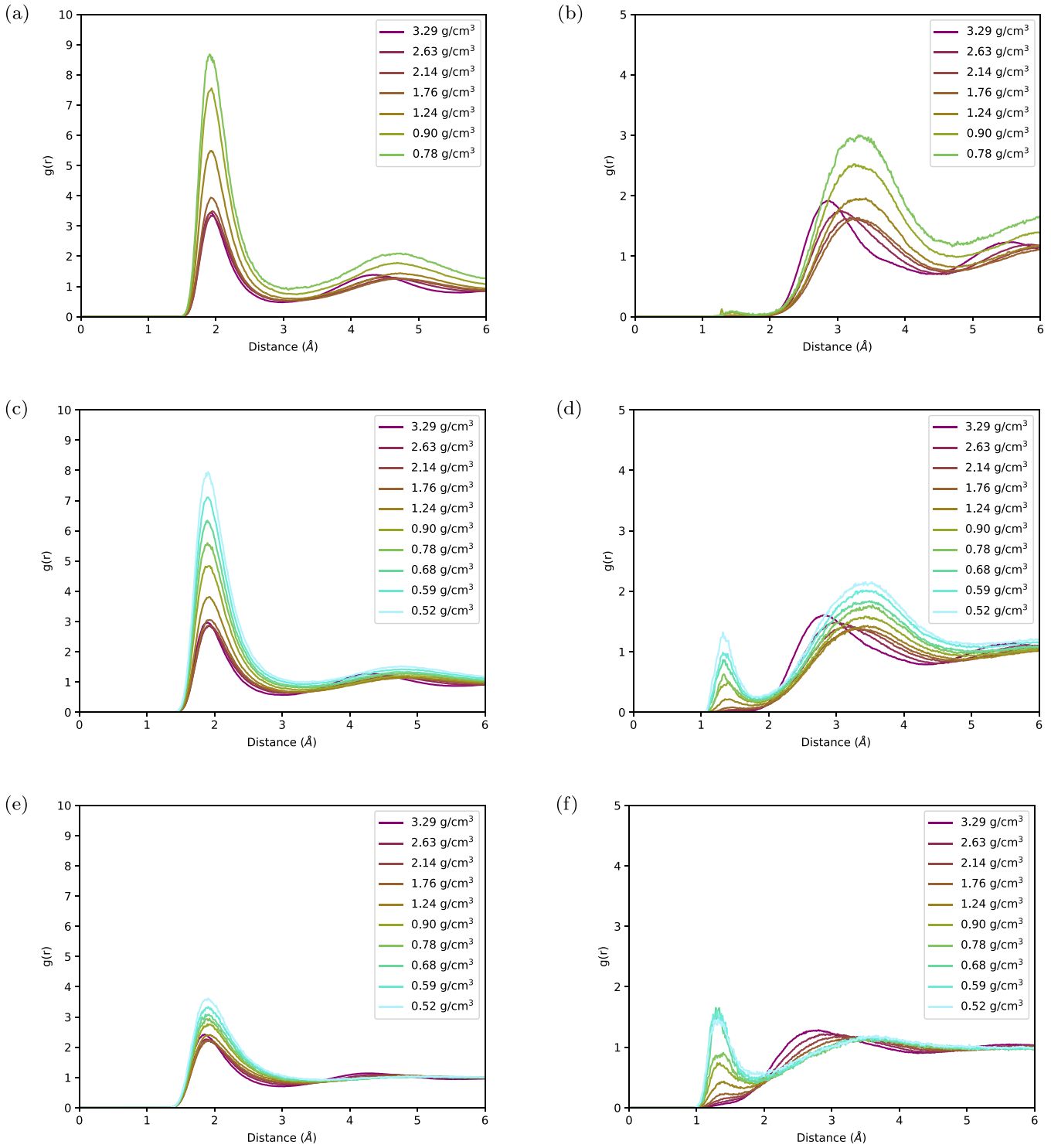


FIG. 3. The pair distribution function for Mg-O (a),(c),(e) and O-O (b),(d),(f) at three isotherms: 4000 (a),(b), 6000 (c),(d), and 10 000 (e),(f) K, and several densities. The first maxima yield a good approximation of the average bond distances. The first minima yield the radius of the first coordination sphere. We use this radius further in the manuscript to define the threshold for interatomic bonding.

indication of the fluid nature of the system studied in the simulation.

c. Velocity autocorrelation function

The general expression for a time correlation function, such as the velocity self-correlation function, is shown in Eq. (3):

$$C(\tau) = \frac{1}{\tau}\sum_{t=0}^{T/2} A(\tau + t) * A(t), \tag{3}$$

where τ is the width of the time interval, T is the total time of the simulation, and A is a time-dependent variable. Here

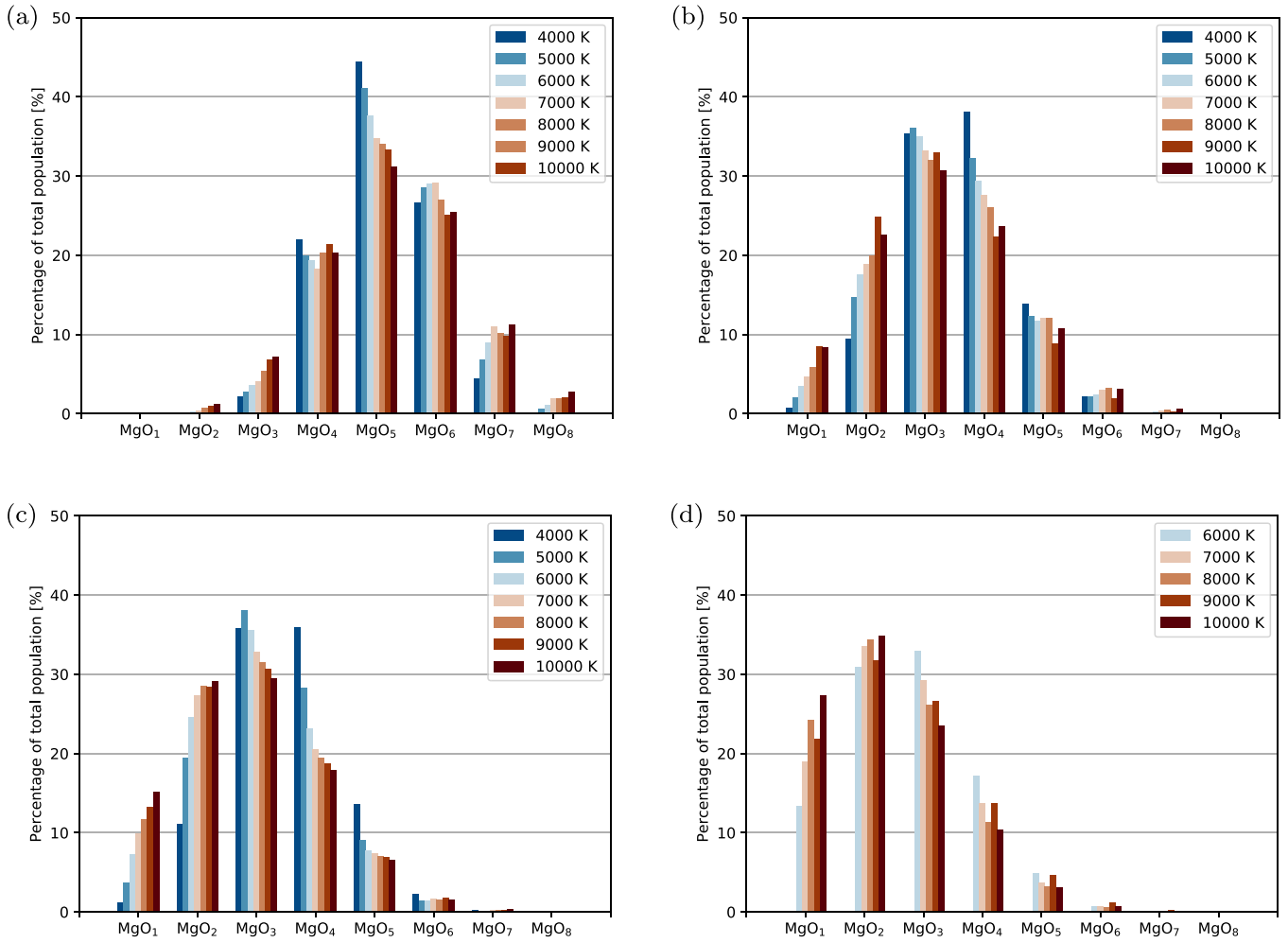


FIG. 4. The speciation of MgO_x polyhedra at several densities and temperatures. The coordination polyhedra around each atom are obtained using the analysis of the pair distribution function (Fig. 3). The MgO fluid is dominated by MgO_5 and MgO_6 at high densities. The coordination decreases sharply as the density decreases toward and passes the spinodal density. Parts (a), (b), (c), and (d) correspond, respectively, to 2.63, 1.24, 0.90, and 0.52 g/cm^3 densities.

we study the self-correlation function of the atomic velocities. Just like with the MSD, we can use the velocity correlation function to the determine the diffusion coefficient by taking its integral, shown in Eq. (4), where again $n = 2, 4, 6$ for one, two, and three dimensions, respectively,

$$D_\alpha = \frac{1}{nm_iN}\int_{-\infty}^{\infty} C(t)dt. \tag{4}$$

d. Bader charge analysis

We apply the atoms-in-molecule approach of the Bader analysis [44] to obtain the static atomic volumes and charges, using a postprocessing code from the Henkelman Group [45–48]. The procedure finds the saddle points of the total electronic charge distribution around each atom, which, when connected, build the zero flux surface of the charge. These surfaces delimit the parts of the volume of the structure that are assigned to each atom. The integrals of the electronic density inside the atomic volumes yield the total negative atomic charge. The atomic charges are obtained after subtracting the positive charge of the nucleus from the total negative charges.

III. RESULTS AND DISCUSSION

A. Critical point

We monitor the variation of the pressure as a function of density at several isotherms in the 4000–10 000 K temperature range (see Table I). These temperatures are above the melting point at ambient pressure conditions, and they extend into the gas and supercritical domains. Depending on the isotherm, we cover the 0.37–3.29 g/cm^3 density range. We approximate the van der Waals model with a cubic function fit to the pressure-density points, whose local extrema yield the two spinodals.

Figure 2(a) shows the pressure-density relation at all the isotherms considered here. The liquid spinodals are also indicated on the diagrams. The last two isotherms at which a local minimum can be identified are the 6000 and 6500 K isotherms [Fig. 2(b)]. For these two isotherms, the calculations can reliably sample low enough densities, i.e., the local maxima in the pressure variation can be identified, corresponding to

the gas spinodals. In particular, the 6500 K isotherm shows a local maximum around 0.4 g/cm^3 and a local minimum around 0.7 g/cm^3 density for pressures of about 1 kbar. On the contrary, there is no local minimum or local maximum along the 7000 K isotherm, but only a monotonous decrease of the pressure with decreasing density. Consequently, the position of the critical point can be constrained in temperature by the isotherms 6500 and 7000 K, and in density by the gas and liquid spinodals at 6500 K, i.e., in the 0.45–0.6 g/cm^3 density interval. The pressure range corresponding to these temperatures and densities is 1–2 kbars.

The computed position of the critical point of MgO lies at higher temperatures than any phase in the MgO-SiO_2 phase space, which is relevant for the bulk composition of rocky planets. The critical points calculated for $MgSiO_3$ [25], Mg_2SiO_4 [26], and SiO_2 [38] lie in the 6200–6500 K temperature range and around 0.50 g/cm^3 density. The higher temperatures of MgO confirm its refractory character, while its smaller range of density corresponds to the lighter mass of MgO.

B. Structure of the fluids

Figure 3 shows the pair distribution functions as a function of density at 4000, 6000, and 10 000 K. The average Mg-O bond length, i.e., the first peak of the distribution function, is on the order of 1.96 Å, weakly dependent on temperature and density. Under compression up to 63 GPa, the bond length decreases by only 0.02 Å along the same isotherm. The radius of the first coordination sphere, which is chosen as the *bonding* criterion, shows a larger variability with both density and temperature. Increasing the density reduces the bonding threshold. The range of the threshold increases considerably with increasing temperature. As a reference, at 30 GPa and 4000 K, the Mg-O bond length in the MgO fluid is slightly larger than the Mg-O bond length in pyrolite [49].

At all temperatures, at 2.63 g/cm^3 density, the MgO fluid is dominated by MgO_5, with the second most abundant coordination being MgO_6. Decreasing the density changes the dominant species toward MgO_4 and MgO_3, and down to MgO_2 in the supercritical fluid at 0.5 g/cm^3. This decrease in coordination is natural, as it accompanies the decompression of the fluid. Increasing the temperature broadens the distribution of the coordination polyhedra. In the liquid, the coordination number of Mg by O is similar to the one encountered in liquid pyrolite [49]. Figure 4 shows the chemical speciation in the MgO fluid at all densities at several temperatures.

The O-O bond distances are sensitive to both density and temperature. At subcritical temperatures, the first maximum of the pair distribution functions lies around 3.5 Å for densities below about 2 g/cm^3. At higher densities, there is a clear decrease of the O-O bond distance, which can be directly related to the increase in coordination of the Mg-O polyhedra from $MgO_{2\text{-}3}$ to $MgO_{4\text{-}5}$. At high temperatures and at densities below 1.24 g/cm^3, the O-O pair distribution function reveals the presence of a peak around 1.3 Å, which corresponds to the characteristic bond length of the O_2 molecule. As observed in various silicate systems, such as feldspars [24] or silica [38], oxygen molecules are present right below the critical temperature as well as in the supercritical fluid.

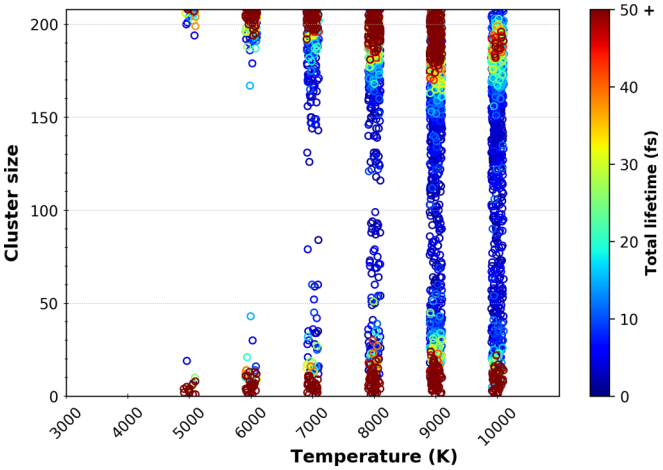


FIG. 5. The speciation of fluid MgO as a function of temperature at 0.68 g/cm^3. Each dot represents one Mg_xO_y cluster, and the vertical axis indicates its size, i.e., $x + y$. At this density, increasing the temperature takes the system from inside the liquid-vapor dome to the supercritical state. The bimodal distribution of cluster sizes is characteristic for a gas + liquid mixture, while a continuous distribution characterizes the supercritical state. The gap between the cluster sizes closes as the system approaches supercritical temperatures; the most stable clusters are always found at the two extremes.

The analysis of the $\cdots$Mg-O-Mg-O-Mg$\cdots$ polymerization in the fluid allows us to separate the gas phase from the liquid phase. Indeed, the fluid is characterized by largely connected $[MgO_x]_n$ clusters, which represent branched polymers of alternating cations, i.e., Mg, and anions, i.e., O. The gas phase shows isolated clusters, of very limited size. Figure 5 shows the population distribution of all the $[Mg_xO_y]$ polymers and gas clusters, at a density of 0.68 g/cm^3 as a function of temperature. At the lowest temperatures, i.e., 5000 and 6000 K, there is a clear separation between two groups of cluster sizes. The highest values of $x + y$ build the liquid MgO phase, and the lowest values build the incipient gas species escaping from the fluid. As the temperature increases, the gap in the distribution of the cluster sizes closes up. This is indicative of the continuous character of the gas to liquid transition as the system reached the supercritical regime. The supercritical feldspars show a similar behavior [24].

The simulations suggest that the dominant species that form the incipient gas of our system are atomic Mg, atomic O, and some MgO and O_2 molecules. However, further quantification of the gas phase, in terms of relative amounts of stable component species, requires a better sampling of the configuration space, which requires both considerably longer and larger simulations and exploring considerably lower densities.

C. Vibrational spectrum

At ambient conditions, the solid B1 phase of MgO has only one infrared active phonon mode, whose transverse optical (TO) component lies around 380 cm^{-1}, and the longitudinal optical (LO) one is around 700 cm^{-1}, with the bulk of the spectrum in the 300–500 cm^{-1} frequency range. Infrared reflectivity of B1 [50] shows the presence of a shoulder around

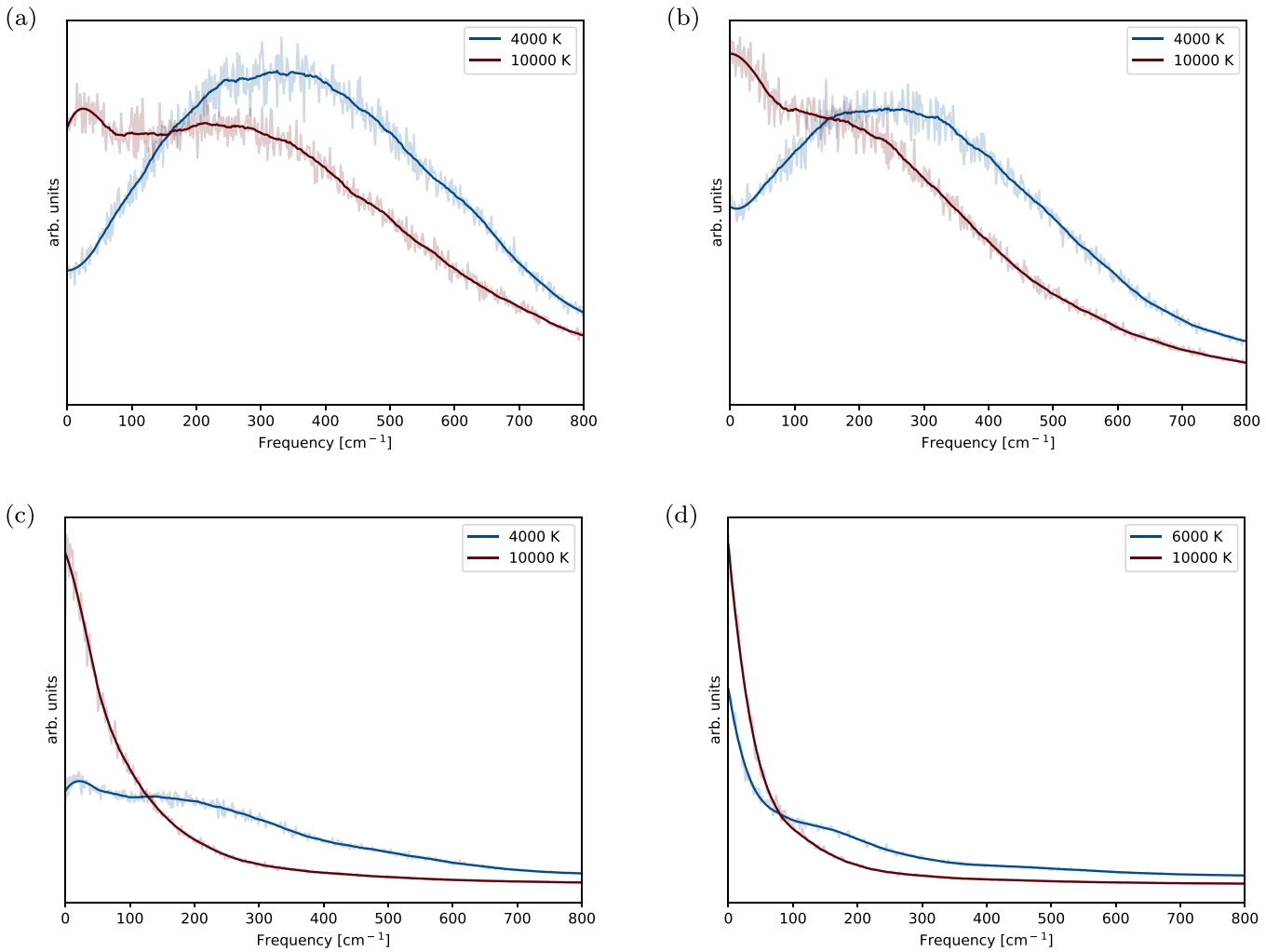


FIG. 6. Total vibrational spectra at several densities of fluid MgO are obtained from the velocity-velocity self-correlation function. Only the values computed at the lowest and the highest isotherms are shown. The noisy data, shown in the background, were filtered using the Savitzky-Golay filter [51]. At high density and low temperature, the spectrum shows a broad peak around 300–500 cm^{-1}. The peak is smoothed out with decreasing density and increasing temperature. Parts (a), (b), (c), and (d) correspond, respectively, to 3.29, 2.63, 0.90, and 0.52 g/cm^3 densities.

620 cm^{-1}, corresponding to the tail of the LO-TO splitting. Increasing temperature makes this shoulder disappear and shifts the entire region toward lower frequencies.

The computed vibrational spectrum for the fluid at 4000 K and a density of 3.29 g/cm^3 (Fig. 6), corresponding to 30 GPa, shows similarities with the B1 phase. There is a broad peak around 300 cm^{-1} with a broad shoulder at higher frequencies. Compared to the solid, the spectrum of the fluid is, as expected, more smoothed out, with no detailed features, because of the temperature and the variety of the local coordinations. Another important feature is the nonzero component at zero frequency, which is due to the diffusion.

Decreasing the density to 0.5 g/cm^3 and increasing the temperature to 10 000 K smoothes out the main vibrationally active region, albeit shifted toward lower frequencies. The details in the spectrum become less and less pronounced. Eventually, at even lower densities and/or higher temperatures, the spectrum should asymptotically become featureless and approach that of an atomic gas.

D. Transport properties

We determine the MSDs at all volume and temperature points. The resulting MSDs for oxygen and for magnesium at several isochores are shown in Fig. 7. At all conditions, the MSDs of both atomic types show positive slopes, the systems being in a fluid state.

The O atoms travel for longer distances than the Mg atoms over the same amount of time. The differences between the two atoms depend on both temperature and density, being much more pronounced at higher temperatures and densities. At 2.63 g/cm^3, i.e., 9 GPa and 4000 K, the Mg and O atoms travel, respectively, 10 and 10 Å over 10 ps. At the same density and 6000 K they travel, respectively, 14 and 15 Å over 10 ps, and at 10 000 K they travel, respectively, 20 and 24 Å over 10 ps. Inside the liquid-gas dome, at 4000 K and 0.90 g/cm^3, the Mg atoms travel 13 Å over 10 ps, and the O atoms travel 14 Å over the same amount of time. The net increase is due to the decompression associated with the opening of the nanobubbles. There is also a clear increase

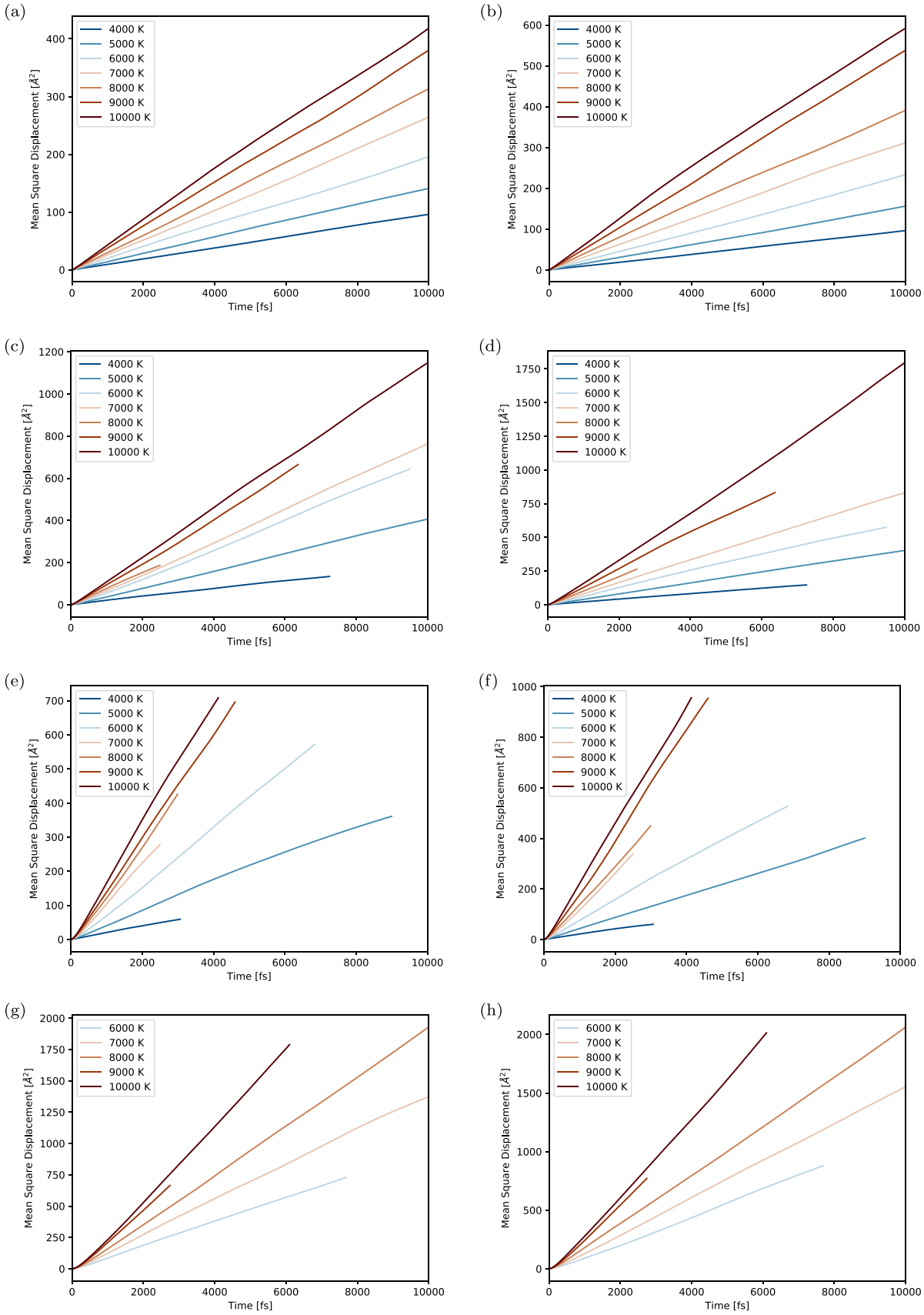


FIG. 7. Mean-square displacements calculated at four isochores. (a),(b) 2.63 g/cm^3; (c),(d) 1.24 g/cm^3; (e),(f) 0.78 g/cm^3; (g),(h) 0.52 g/cm^3. The differences in lengths are the result of differences in the length of simulations, the shortest being 5 ps and the longest 20 ps. MgO is fluid at all conditions studied here.

TABLE II. Comparison between the diffusion coefficients at 0.78 g/cm^3 as a function of temperature, estimated from the slope of the mean-square displacements as a function of time (MSD) and obtained from the velocity autocorrelation function (VA).

	Diffusion coefficient (m^2/s)			
Temperature	Mg (VA)	O (VA)	Mg (MSD)	O (MSD)
4000	4.87×10^{-8}	4.34×10^{-8}	3.22×10^{-8}	3.25×10^{-8}
5000	6.36×10^{-8}	8.59×10^{-8}	6.69×10^{-8}	7.43×10^{-8}
6000	1.51×10^{-7}	1.35×10^{-7}	1.39×10^{-7}	1.29×10^{-7}
7000	2.24×10^{-7}	2.69×10^{-7}	1.85×10^{-7}	2.26×10^{-7}
8000	2.83×10^{-7}	3.32×10^{-7}	2.37×10^{-7}	2.50×10^{-7}
9000	2.92×10^{-7}	3.65×10^{-7}	2.52×10^{-7}	3.46×10^{-7}
10000	2.98×10^{-7}	4.09×10^{-7}	2.86×10^{-7}	3.85×10^{-7}

in distance traveled by the atoms from 6000 to 7000 K for 1.24 g/m^3 and 0.78 g/cm^3 [Figs. 7(b) and 7(c)], which corresponds to the passage to the supercritical fluid.

The slope of the MSD with respect to time yields the diffusion coefficients. The results obtained from integrating the velocity self-correlation function confirm these results. Table II lists the values obtained from both methods for comparison. Figure 8 shows the diffusion coefficients as a function of density or different isotherms, as obtained from the self-correlation function.

Below the critical temperature, the dependence of the diffusion coefficients displays a clear separation into two regimes; a linear trend on a log scale at higher densities, and a roughly constant diffusion at lower densities. The point of the slope change corresponds to the density at which the first nanobubbles start to nucleate in the system. As the volume of the simulation box increases, it is the density of the entire system, liquid + gas, that decreases. But the density of the liquid is roughly constant. As the liquid is the dominant phase in these systems at these conditions, the diffusion of both Mg and O atoms reflects their behavior in the liquid phase.

As the temperature increases above the supercritical point, the system is monophasic, so the diffusion coefficient reflects the behavior of the atoms in the total homogeneous system. Here the increase of the volume of the simulation box induces a decrease of the density of the entire system. As the atoms lie farther apart, their diffusion continues to increase with decreasing density.

Consequently, the temperature variation of the diffusion coefficients yields another way of quantifying the transition towards the liquid + vapor dome, and/or the passage to supercritical conditions.

E. Atomic charges

Finally we analyze the atomic charges of all the atoms in our simulations using the atoms-in-molecule approach, as mentioned in Sec. II. We select several snapshots inside the liquid-vapor dome, which show atoms in both liquid and gas phases, and one snapshot from the supercritical phase.

The values of the Bader charges correlate with the coordination number for both Mg and O. This correlation is visible in Fig. 9. The trend indicates that decreasing the coordination numbers makes the atoms more neutral.

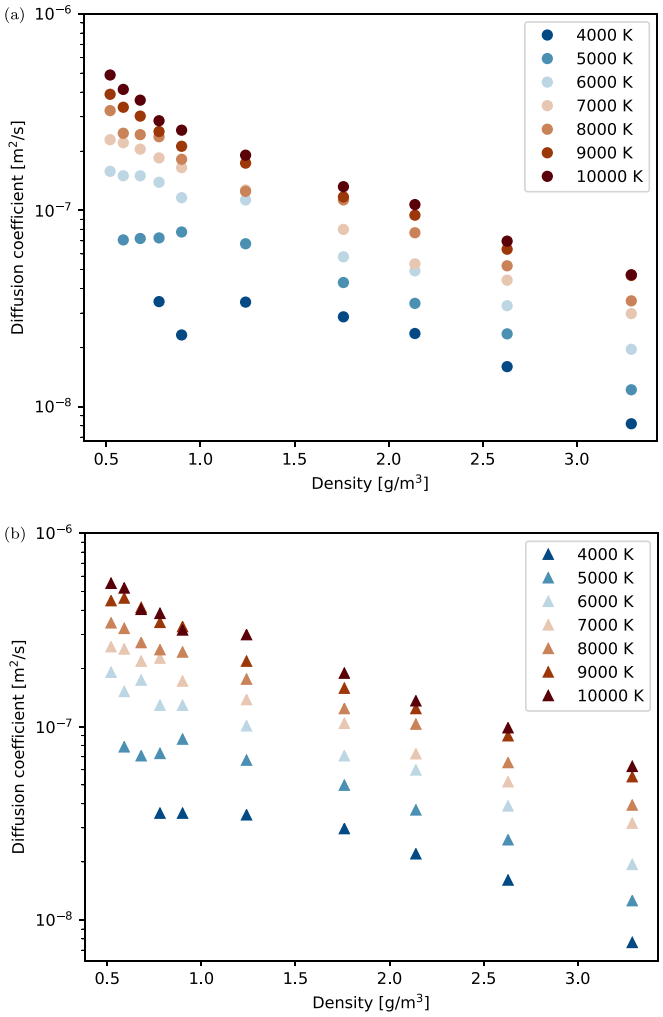


FIG. 8. Diffusion coefficients for Mg atoms (a) and O atoms (b) as calculated from the velocity autocorrelation function plotted in log scale against the density. Below the critical temperature, the diffusion coefficients exhibit a change of slope corresponding to the passage inside the liquid-vapor dome. Above the critical temperature, the diffusion changes monotonously with density.

The isolated atoms in the gas phase all have charges approaching zero. This suggests that the gas is close to an atomic-gas model and is not ionized. Indeed, these temperatures are far below the first ionization energy of monatomic magnesium, while the monatomic oxygen does not carry a supplementary electronic charge. The atoms in the liquid, which lie on or close to the interface with the cavities, are ionized, but to a significantly lesser extent than the atoms in the bulk. This suggests that the surface of the bubbles tends to become neutral, and does not carry dipoles. Both Mg and O atoms that lie inside the bulk liquid phase have large charges, between 1 and 2 in absolute values, negative for O and positive for Mg. This suggests an ionic liquid.

The charges in the supercritical fluid show a smaller spread than for the subcritical conditions. Their values are close to the nominal values for Mg^{2+} and O^{2-}. This suggests that the supercritical fluid preserves the ionic character of the homogeneous liquid.

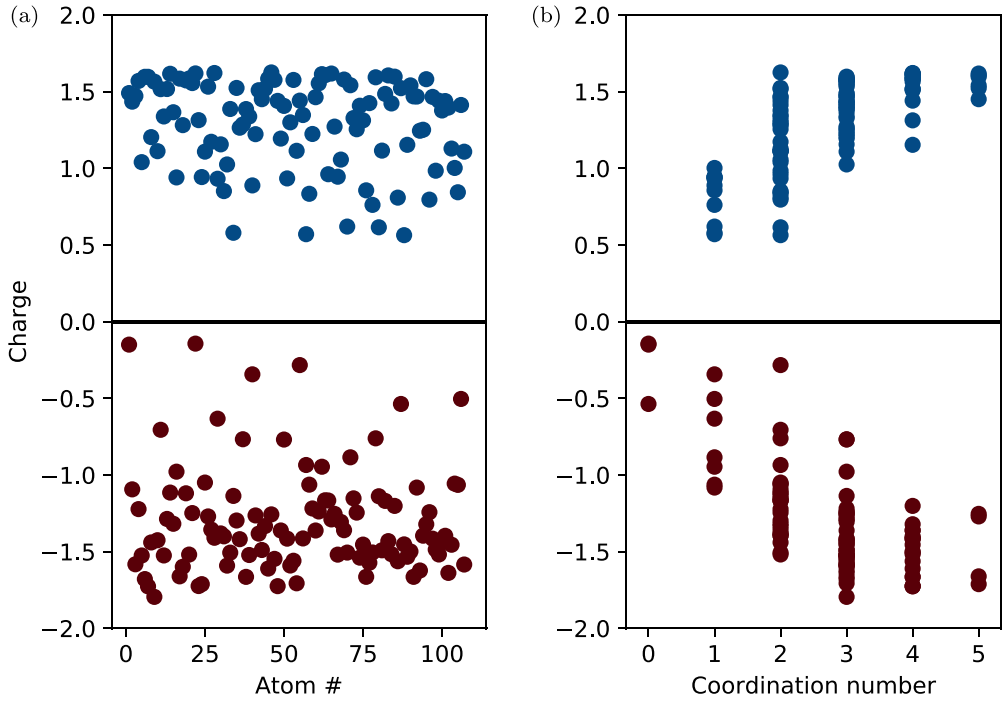


FIG. 9. Atomic charges for all Mg (blue) and O (red) atoms for a representative snapshot inside the liquid-vapor dome (a). The order of the coordination polyhedron around each atom (b). Charge values close to the nominal correspond to highly coordinated atoms, which lie in the bulk liquid. Charge values close to zero correspond to low coordination numbers, as encountered on the liquid-gas interface and in the gas.

IV. CONCLUSION

We explore an as-yet uncharted area of the phase space of MgO. We provide a thorough analysis of the behavior of fluid MgO at the low-density and high-temperature conditions typically occurring around the supercritical point, using *ab initio* molecular dynamics. We apply a wide range of postprocessing tools to describe different facets of the MgO system in the phase space.

We determine the critical point to be in the density range of 0.45–0.6 g/cm^3 and between 6500 and 7000 K. That puts it in the refractory category when compared to other major rock-forming compounds. We characterize a series of transport and structural properties and find similarities to other Mg-rich natural fluids [49]. From the atomic charge analysis we ascertain that the bulk liquid and the supercritical fluid are ionic. For conditions inside the liquid-vapor dome, we find that the surface of the gas bubbles does not carry charge, and that any incipient gas is not charged.

Future studies should address the process of vaporization at much larger time and space scales, and they should determine the liquid-vapor equilibrium curve. For this, more computationally efficient methods will need to be applied, such as machine learning potentials, as they are beyond the scope of the present work. Comparison to other B1-B2 diatomic phases would be interesting as well. Our work contributes to the characterization of an archetypal material, which is also one of the primary constituents of all rocky planets, at conditions typically encountered during planetary formation.

ACKNOWLEDGMENTS

We acknowledge support from the European Research Council under EU Horizon 2020 research and innovation program (Grant Agreement No. 681818–IMPACT to R.C.), the Research Council of Norway, Project No. 223272, and access to supercomputing facilities via the eDARI stl2816 grants, the PRACE RA4947 grant, and the Uninet2 NN9697K grant.

[1] R. M. Canup, Dynamics of lunar formation, Annu. Rev. Astron. Astrophys. **42**, 441 (2004).

[2] S. Stewart, E. Davies, M. Duncan, S. Lock, S. Root, J. Townsend, R. Kraus, R. Caracas, and S. Jacobsen, in *The Shock Physics of Giant Impacts: Key Requirements for the Equations of State*, AIP Conf. Proc. No. 2272 (AIP, New York, 2020), p. 080003.

[3] Z. Li, R. Caracas, and F. Soubiran, Partial core vaporization during Giant Impacts inferred from the entropy and the critical point of iron, Earth Planet. Sci. Lett. **547**, 116463 (2020).

[4] R. G. Kraus, S. T. Stewart, D. C. Swift, C. A. Bolme, R. F. Smith, S. Hamel, B. D. Hammel, D. K. Spaulding, D. G. Hicks, J. H. Eggert, and G. W. Collins, Shock vaporization of silica and the thermodynamics of planetary impact events, J. Geophys. Res. E **117**, E09009 (2012).

[5] R. G. Kraus, S. Root, R. W. Lemke, S. T. Stewart, S. B. Jacobsen, and T. R. Mattsson, Impact vaporization of

planetesimal cores in the late stages of planet formation, Nat. Geosci. **8**, 269 (2015).

[6] J. J. Monaghan, Smoothed particle hydrodynamics and its diverse applications, Annu. Rev. Fluid Mech. **44**, 323 (2011).

[7] J. D. McHardy, An introduction to the theory and use of SESAME equations of state, Tech. Rep., Report No. LA-14503 (Los Alamos National Laboratory, Los Alamos, NM, 2018).

[8] S. L. Thompson, ANEOS—Analytic equations of state for shock physics Codes—Input Manual Report SAND89-2951 (Sandia National Laboratories, 1990), https://www.osti.gov/biblio/6939284.

[9] A. E. Ringwood, Phase transformations and their bearing on the constitution and dynamics of the mantle, Geochim. Cosmochim. Acta **55**, 2083 (1991).

[10] I. Jackson and R. Liebermann, Melting and elastic shear instability of alkali halides, J. Phys. Chem. Solids **35**, 1115 (1974).

[11] E. Ohtani, Melting temperature distribution and fractionation in the lower mantle, Phys. Earth Planet. Inter. **33**, 12 (1983).

[12] A. Zerr and R. Boehler, Constraints on the melting temperature of the lower mantle from high-pressure experiments on MgO and magnesioüstite, Nature (London) **371**, 506 (1994).

[13] T. Taniuchi and T. Tsuchiya, The melting points of MgO up to 4 TPa predicted based on ab initio thermodynamic integration molecular dynamics, J. Phys.: Condens. Matter **30**, 114003 (2018).

[14] A. J. Cohen and R. G. Gordon, Modified electron-gas study of the stability, elastic properties, and high-pressure behavior of MgO and CaO crystals, Phys. Rev. B **14**, 4593 (1976).

[15] M. J. Mehl, R. E. Cohen, and H. Krakauer, Linearized augmented plane wave electronic structure calculations for MgO and CaO, J. Geophys. Res. **93**, 8009 (1988).

[16] R. S. McWilliams, D. K. Spaulding, J. H. Eggert, P. M. Celliers, D. G. Hicks, R. F. Smith, G. W. Collins, and R. Jeanloz, Phase transformations and metallization of magnesium oxide at high pressure and temperature, Science **338**, 1330 (2012).

[17] F. Coppari, R. F. Smith, J. H. Eggert, J. Wang, J. R. Rygg, A. Lazicki, J. A. Hawreliak, G. W. Collins, and T. S. Duffy, Experimental evidence for a phase transition in magnesium oxide at exoplanet pressures, Nat. Geosci. **6**, 926 (2013).

[18] K. Umemoto, R. M. Wentzcovitch, S. Wu, M. Ji, C. Z. Wang, and K. M. Ho, Phase transitions in $MgSiO_3$ post-perovskite in super-Earth mantles, Earth Planet. Sci. Lett. **478**, 40 (2017).

[19] F. Soubiran and B. Militzer, Anharmonicity and Phase Diagram of Magnesium Oxide in the Megabar Regime, Phys. Rev. Lett. **125**, 175701 (2020).

[20] R. Musella, S. Mazevet, and F. Guyot, Physical properties of MgO at deep planetary conditions, Phys. Rev. B **99**, 064110 (2019).

[21] B. Fegley, N. S. Jacobson, K. B. Williams, J. M. C. Plane, L. Schaefer, and K. Lodders, Solubility of rock in steam atmospheres of planets, Astrophys. J. **824**, 103 (2016).

[22] J. Vassent, A. Marty, B. Gilles, and C. Chatillon, Thermodynamic analysis of molecular beam epitaxy of MgO(s) II. Epitaxial growth of MgO layers on Fe(001) substrates, J. Cryst. Growth **219**, 444 (2000).

[23] N. Jacobson, N. Ingersoll, and D. Myers, Vaporization coefficients of SiO_2 and MgO, J. Eur. Ceram. Soc. **37**, 2245 (2017).

[24] A. Kobsch and R. Caracas, The critical point and the supercritical state of alkali feldspars: Implications for the behavior of the crust during impacts, J. Geophys. Res.: Planets **125**, e2020JE006412 (2020).

[25] B. Xiao and L. Stixrude, Critical vaporization of MgSiO3, Proc. Natl. Acad. Sci. (USA) **115**, 5371 (2018).

[26] J. P. Townsend, G. Shohet, and K. R. Cochrane, Liquid-vapor coexistence and critical point of Mg_2SiO_4 from ab initio simulations, Geophys. Res. Lett. **47**, e2020GL089599 (2020).

[27] G. Kresse and J. Hafner, Ab initio molecular dynamics for liquid metals, Phys. Rev. B **47**, 558 (1993).

[28] G. Kresse and J. Hafner, Norm-conserving and ultrasoft pseudopotentials for first-row and transition elements, J. Phys.: Condens. Matter **6**, 8245 (1994).

[29] G. Kresse and J. Furthmüller, Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set, Comput. Mater. Sci. **6**, 15 (1996).

[30] G. Kresse and J. Furthmüller, Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, Phys. Rev. B **54**, 11169 (1996).

[31] G. Kresse and D. Joubert, From ultrasoft pseudopotentials to the projector augmented-wave method, Phys. Rev. B **59**, 1758 (1999).

[32] P. E. Blöchl, Projector augmented-wave method, Phys. Rev. B **50**, 17953 (1994).

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

[34] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, J. Chem. Phys. **81**, 511 (1984).

[35] R. M. Hazen, Effects of temperature and pressure on the cell dimension and X-ray temperature factors of periclase, Am. Mineral. **61**, 266 (1976).

[36] P. Aursand, M. A. Gjennestad, E. Aursand, M. Hammer, and Ø. Wilhelmsen, The spinodal of single- and multi-component fluids and its role in the development of modern equations of state, Fluid Phase Equilib. **436**, 98 (2017).

[37] P. Carlès, A brief review of the thermophysical properties of supercritical fluids, J. Supercrit. Fluids **53**, 2 (2010).

[38] E. C. Green, E. Artacho, and J. A. Connolly, Bulk properties and near-critical behaviour of SiO_2 fluid, Earth Planet. Sci. Lett. **491**, 11 (2018).

[39] V. V. Vasisht, S. Saw, and S. Sastry, Liquid-liquid critical point in supercooled silicon, Nat. Phys. **7**, 549 (2011).

[40] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.105.064105 for more details on the finite-size effect, van der Waals interaction, and EOS fitting.

[41] K. Binder, B. J. Block, P. Virnau, and A. Tröster, Beyond the Van Der Waals loop: What can be learned from simulating Lennard-Jones fluids inside the region of phase coexistence, Am. J. Phys. **80**, 1099 (2012).

[42] G. Faussurier, C. Blancard, and P. L. Silvestrelli, Evaluation of aluminum critical point using an ab initio variational approach, Phys. Rev. B **79**, 134202 (2009).

[43] R. Caracas, A. Kobsch, N. V. Solomatova, Z. Li, F. Soubiran, and J.-A. Hernandez, Analyzing melts and fluids from ab initio molecular dynamics simulations with the UMD package, J. Visual Exp. (175), e61534 (2021).

[44] R. F. Bader, S. G. Anderson, and A. J. Duke, Quantum topology of molecular charge distributions, J. Am. Chem. Soc. **101**, 1389 (1979).

[45] W. Tang, E. Sanville, and G. Henkelman, A grid-based Bader analysis algorithm without lattice bias, J. Phys.: Condens. Matter **21**, 084204 (2009).

[46] E. Sanville, S. D. Kenny, R. Smith, and G. Henkelman, Improved grid-based algorithm for Bader charge allocation, J. Comput. Chem. **28**, 899 (2007).

[47] G. Henkelman, A. Arnaldsson, and H. Jónsson, A fast and robust algorithm for Bader decomposition of charge density, Comput. Mater. Sci. **36**, 354 (2006).

[48] M. Yu and D. R. Trinkle, Accurate and efficient algorithm for Bader charge integration, J. Chem. Phys. **134**, 064111 (2011).

[49] N. V. Solomatova and R. Caracas, Pressure-induced coordination changes in a pyrolitic silicate melt from ab initio molecular dynamics simulations, J. Geophys. Res.: Solid Earth **124**, 11232 (2019).

[50] J. R. Jasperse, A. Kahan, J. N. Plendl, and S. S. Mitra, Temperature dependence of infrared dispersion in ionic crystals LiF and MgO, Phys. Rev. **146**, 526 (1966).

[51] A. Savitzky and M. J. E. Golay, Smoothing and differentiation of data by simplified least squares procedures, Anal. Chem **36**, 1627 (1964).

Appendix B

Alternative critical point approximation

Instead of giving a temperature and density range for the critical point, we can use the method described in Townsend et al. (2020) to get a single critical point with our equation of state (EOS). Here we make use of the fact that we have a cubic function as an EOS, eq. B.1. We can find the critical point by finding the point where the first and second derivative of this EOS are both 0, eq. B.2.

$$P(\rho, T) = (a_0 + a_1 T)\rho + (b_0 + b_1 T)\rho^2 + (c_0 + c_1 T)\rho^3 \tag{B.1}$$

$$\left(\frac{\delta P}{\delta \rho}\right)_T = \left(\frac{\delta^2 P}{\delta \rho^2}\right)_T = 0 \tag{B.2}$$

We can perform orthogonal distance regression on all of our P,T,ρ data (including the standard deviation) to obtain a single critical point for each of our materials. If we plot the first and second derivatives for MgO, CaO, and $MgSiO_3$ we get the plots shown in figure B.1. The associated critical points we obtain this way are summarized in table B.1. We can see that the critical temperatures for MgO and CaO fall within the range we obtained with the spinodal point analysis method, but the $MgSiO_3$ critical temperature falls below the range we obtained previously.

	T_c [K]	ρ_c [g/cm^3]
MgO	6576	0.46
CaO	6383	0.43
$MgSiO_3$	5867	0.52

Table B.1: Critical points of MgO, CaO, and $MgSiO_3$ determined by the point where the first and second derivative of the fit EOS are equal to 0.

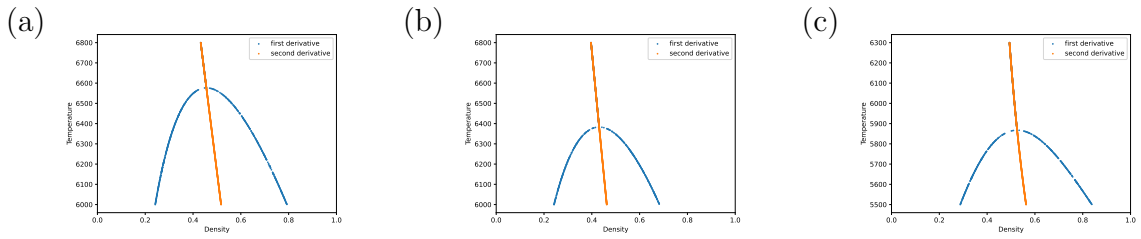


Figure B.1: First and second derivatives of the EOS as fit using eq. B.1 on MgO (a), CaO (b), and $MgSiO_3$ (c).