Probing the Thermodynamic Properties of Mantle Rocks in Solid and Liquid States

Thesis by

Aaron S. Wolf

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2013

(Defended May 29th, 2013)

“Science has achieved some wonderful things, of course, but I’d far rather be happy than right any day.” —The Hitchhiker’s Guide to the Galaxy

Acknowledgements

Completing my PhD is an accomplishment that I'm lucky to share with a large group of supportive people. I first wish to acknowledge my Advisors Paul Asimow and Jennifer Jackson. They have always allowed me great freedom in my studies and encouraged me to pursue interesting scientific problems, even if they lie somewhat further afield from the core of my work. In particular, Paul showed great faith and patience as I searched for a simplified liquid model. There were many iterations and dead ends, but he stuck with me through it all, providing a crucial sounding board until we made it through to the other side. Jennifer has always indulged my curiosity about research that grew out of the primary study while helping to keep the main project moving toward completion. When I arrived at Caltech, I began my primary work with Dave Stevenson, which focused on studying the interior structure of extrasolar planets. Over time, however, my attentions turned increasingly downward to the interior of our own planet. Nevertheless, Dave has provided invaluable guidance over the years, primarily in the form of off-handed comments or asides that would result in fundamental shifts in my thinking, significantly altering my strategy for solving the problem at hand. I will seriously miss these guiding nudges when I leave Caltech. I am also thankful for the time I have spent with John Johnson, who taught the Bayesian Data Analysis course that solidified my love of the Bayesian approach. I sincerely appreciate

the responsibility that John gave me as a teaching assistant for that class in my final year, encouraging me to develop and give numerous lectures and homework sets. I can happily report that I will be taking this course with me as I continue the tradition of preaching the Bayesian gospel to all who have not yet heard it.

I feel truly fortunate to have had the best graduate peer group imaginable. This allowed me to collaborate on many projects with fellow graduate students, typically providing the statistical expertise necessary to attack the interesting problems that research has presented. This began with a project with my officemate Darin Ragozzine, looking into measuring the core properties of extrasolar hot Jupiters through changes in transit light curves. Later, I also collaborated with Alex Hayes using radar observations to constrain lake depths on Titan and Mike Line developing accurate methods for retrieving atmospheric gases and thermal profiles in the extrasolar planetary atmospheres. These side projects have shown me how much fun it is to collaborate with others and how lucky I have been to be in an environment that allows and encourages such things. I also want to thank the three more senior graduate students who were my officemates through much of graduate school: Darin Ragozzine, Alejandro Soto, and Margarita Marinova. You guys were amazing and I learned so much just talking with you about research, homework, life, and politics. Getting assigned to you as an officemate was an amazingly lucky draw, and I feel so fortunate to have been able to spend those years chatting with and working alongside you. More recently, I began having regular lunches with a few fellow students across the division whose interests all overlapped in thinking about the deep earth: June Wicks, Dan Bower, Miki Nakajima, and Claire Waller-Thomas. You guys were as important as my professors in providing support and encouragement in the final year as I wrapped up my thesis work! And of course, there are many other friends that I've had, both within Caltech and outside of it, who made my time in graduate school great. Thanks especially to Steve Chemtob, Brian Cornelius, Meg Rosenberg, Jon Wolfe, and to everyone else that I fail to mention individually.

Finally, but most importantly, I want to thank my family. To my parents Daryl and Linda Wolf, I couldn't ask for more supportive, encouraging, and understanding people to have raised me. Now we know that all those science experiments over the years growing up have clearly amounted to more than the messes they made. Thanks to my brother Josh Wolf, who is hilarious and incredibly talented, and who I know will always have my back. I was also fortunate that both my parents grew up in Los Angeles, and so most of my extended family lives here. Thanks to my grandparents Estelle and Mel Brunetti and Fred and Pauline Wolf. I have really appreciated all the encouragement you've always been willing to give and it was great to be able to visit you more often over the last few years. A special thanks to my grandfather Marty Bolhower, who unfortunately died early during my time at Caltech. We all miss you so much and I just feel lucky that I got to spend what time I did with you in my first year here at Caltech. Also, many thanks to my aunts and uncles who live nearby and have hosted many Thanksgiving and Passover dinners. I will miss you when I've moved away! The biggest thanks of all goes to my wife Gretchen Keppel-Aleks. She went through this whole process two years ago when she got her PhD, and I hope that I was even a tiny fraction as supportive as she has been for me, especially over the last year. More than the degree that Caltech will give me, I am most thankful to Caltech for bringing me together with Gretchen. We met here, have vacationed to escape here, and subsequently got married here. I look forward to the coming decades with Gretchen and making them as exciting and adventure-filled as the last five years have been.

Abstract

Our understanding of the structure and evolution of the deep Earth is strongly linked to knowledge of the thermodynamic properties of rocky materials at extreme temperatures and pressures. In this thesis, I present work that helps constrain the equation of state properties of iron-bearing Mg-silicate perovskite as well as oxide-silicate melts. I use a mixture of experimental, statistical, and theoretical techniques to obtain knowledge about these phases. These include laser-heated diamond anvil cell experiments, Bayesian statistical analysis of powder diffraction data, and the development of a new simplified model for understanding oxide and silicate melts at mantle conditions. By shedding light on the thermodynamic properties of such ubiquitous Earth-forming materials, I hope to aid our community's progress toward understanding the large-scale processes operating in the Earth's mantle, both in the modern day and early in Earth's history.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

To attempt to understand our planet's origin and subsequent evolution, we must look below the thin skin of the Earth's surface into its deep interior. Comprising over 99% of our planet's mass, the Earth's mantle and core have been actively redistributing heat and matter for the last four and a half billion years, as they respond to and help shape conditions at the surface. The evolutionary path traced by the Earth's interior is strongly linked to the behavior of the rocky material from which it is made. Understanding the physical properties of mantle rocks over a wide range of extreme conditions is therefore crucial to the goal of building a coherent picture of the Earth, both as it is today and as it once was in the distant past. Peering into the deep Earth is no easy task, as it requires that we probe through thousands of kilometers of opaque rock to a region that is both visually and physically inaccessible—instead, we must rely on more indirect methods. In this thesis, I present experimental and theoretical work which helps to expand our understanding of the thermodynamic properties of rocky material at extreme temperatures and pressures, relevant to both the modern-day mantle and to the mantle as it was early in Earth's history.

The Earth is composed primarily of a large thick shell of solid silicate rocks surrounding a liquid iron core. The core is just under half the Earth's radius, leaving almost 85% of the volume to the Earth's mantle. The mantle is in turn split between the upper and lower mantle, where the distinction reflects the different mineral assemblages that comprise each region. The Earth's lower mantle is composed primarily of a mixture of roughly 80% $(\mathrm{Mg, Fe})(\mathrm{Si, Al})\mathrm{O}_3$ perovskite and the remaining 20% is mostly $(\mathrm{Mg, Fe})\mathrm{O}$ ferropericlase. As perovskite dominates the largest volumetric region of the Earth, this mineral likely represents the single most common material in the entire Earth. It also lies in a particularly dynamic region that is actively convecting, carrying heat up and out of the Earth's interior. If we could choose to understand only a single material in the hopes of shedding light on the evolutionary processes that shaped the Earth, silicate perovskite would undoubtedly be the best choice.

Our understanding of the Earth's mantle depends on direct seismic observations of how waves propagate at depth, combined with our knowledge of how terrestrial rocks behave at extreme conditions. Constraining the high-pressure thermal properties of mantle minerals is therefore one of the outstanding challenges in mineral physics, and helps to inform us about the structure and evolution of the deep Earth. The core-mantle boundary region is marked by numerous seismically observed heterogeneities spanning a wide range of length scales, the largest of which are the Large Low Shear Velocity Provinces (LLSVPs) that lie beneath Africa and the Pacific. These $\sim$1000 km-tall structures are made visible by their distinct shear and bulk sound speeds, density contrasts, and seismically sharp boundaries, which are generally interpreted as evidence of combined thermal and compositional differences from the bulk mantle. Likely compositionally distinct and lying at the base of the Earth's mantle, these large intriguing structures prompt many questions. How could such a vast compositional heterogeneity remain chemically isolated from the convecting mantle? What is the cause of their extreme topographic relief and how does this relate to their apparent longevity? What process is capable of forming such large compositionally distinct reservoirs and implanting them at the base of the mantle?

There are a number of geodynamic theories that potentially explain how such vast chemical reservoirs could resist mixing with the bulk mantle, a situation further complicated by the large vertical length scales involved. These theories generally fall into two broad categories: dynamically lifted passive piles and self-supporting metastable domes (*Tan and Gurnis*, 2007). Passive piles are made of material that is everywhere denser than the ambient mantle, causing it to pool on the core-mantle boundary (CMB), where externally imposed forces like down-going slabs dynamically sweep the piles upward giving them their observed morphology. Metastable domes, on the other hand, are less dense than average mantle at the CMB, causing them to internally convect and propping themselves up with a height determined by the height of neutral buoyancy. As these behaviors are a consequence of the material properties of the pile, they can be directly tested by experiments that determine the compression behavior of mantle minerals at high pressures and temperatures.

In Chapters 2 and 3, I report the results of laser-heated diamond anvil cell experiments that are analyzed to constrain the thermal properties of $\mathrm{(Mg, Fe)SiO_3}$ perovskite, the dominant phase in the Earth's lower mantle and therefore a large likely component of LLSVPs. Unlike previous equation of state studies, we investigate the effect of iron (13%) on both the ambient and thermal properties of perovskite in a quasi-hydrostatic neon pressure medium.

We further employ a simplified mixture model in order to estimate the physical properties of perovskite for a wide range of iron compositions and compare these values to bulk mantle properties at a range of P-T conditions. Using the apparent contrast in density and bulk modulus relative to average lower mantle values, we assess the relative likelihood of passive pile and metastable dome theories for perovskite-dominated LLSVPs.

The evolutionary theories of how such chemical piles could have been generated in the first place is another important line of inquiry, one that connects us with the deep history of our planet. As is often the case, finding answers to the question of “why?” are even more challenging than the question of “what?”. Generally speaking, the mechanistic hypotheses for the generation of LLSVPs fall either into the camp of slow accumulative processes or the remnant of a primordial event. The favored class of long-timescale explanation is that these chemical piles represent the residue of billions of years oceanic crust generation and subduction as a consequence of plate tectonics (*Tackley*, 2011). Oceanic crust is chemically dense relative to the ambient mantle and order of magnitude estimates of the total amount of crust generated over the Earth’s lifetime are sufficient to explain the inferred volumes of these reservoirs (*Xie and Tackley*, 2004). The major poorly constrained aspect of the problem is how efficiently the basaltic component of oceanic plates can be de-laminated and segregated from the bulk of the plate to avoid being carried back upward by convective mantle flow. In contrast to this slow and gradual formation mechanism, the leading primordial remnant theories propose a link between the generation of these chemical reservoirs and the evolutionary processes that operated in the magma ocean present early in Earth’s history. The magma ocean theory states that the large accretional energies delivered to the Earth in the form of giant impacts resulted in melting of significant portions of the mantle (discussed in greater detail below). The cooling of the mantle from this initially molten state is accompanied by complex crystallization processes that have the potential to generate and gravitationally segregate the bulk mantle into multiple chemically separated constituents. Great uncertainty exists about the cooling pathways available to the early molten Earth, but likely potential processes that could generate chemically distinct reservoirs include fractional crystallization (*Solomatov and Stevenson*, 1993) and the longer-lived evolution of an insulated and isolated basal magma ocean (*Labrosse et al.*, 2007). Though our understanding of magma ocean processes are fraught with uncertainty, it is an attractive focus of study due to its potential for fundamentally altering the course of deep Earth evolution.

As magma ocean evolution likely represents one of the primary forces that shaped the early development of the mantle, understanding the behavior of silicate liquids at mantle pressures is crucial to reconstructing the early evolution of Earth and other terrestrial planets. While molten silicates play a minor role in the deep mantle today, their material properties directly determined the Earth's evolutionary path as it cooled from a nearly whole-mantle magma ocean to its present-day solid state. Earth's early history was marked by a series a giant and traumatic events. Giant impacts of large planetary embryos were responsible for setting the initial thermal state of the mantle. Since the gravitational energy of accretion for the Earth exceeds the threshold for complete melting by many orders of magnitude (*Stevenson*, 2007), the degree of melting actually experienced was determined by a balance between accretion and cooling timescales, which depend on how effectively energy was delivered to great depth by large impacts. Based on accretion simulations (e.g., *Chambers*, 2004), we expect that the Earth suffered a series of a few to ten giant impacts during formation, the largest of which probably formed the moon. A recent numerical investigation into the moon-forming impact suggests that the entire mantle is likely to have melted (*Nakajima and Stevenson*, 2012); at the same time, new He and Xe isotopic evidence indicates that the Earth must have experienced at least two magma ocean periods in its history while simultaneously maintaining some primordial unmixed reservoirs (*Mukhopadhyay*, 2012; *Mukhopadhyay et al.*, 2012). These isotopic studies underscore the potential links that exist between very early Earth processes and the structures that we can seismically observe today residing in the deep mantle.

Our understanding of magma ocean crystallization and evolution is currently undergoing significant changes, linked directly to new insights on the behavior of silicate melts. Recent experimental and theoretical works (*Mosenfelder et al.*, 2007; *Stixrude et al.*, 2009) show that the thermal properties of silicate liquids upon compression differ significantly from our previous understanding, causing adiabatic thermal profiles that are much steeper than previously assumed—steeper even than the solidus (or melting curve for simple systems) at the base of the mantle. The cooling liquid adiabat adopted by the rapidly convecting magma ocean therefore likely first encounters the solidus at mid mantle depths, rather than at the base of the mantle, indicating a complex center-outwards style of crystallization rather than the classic bottom-up solidification that has dominated our thinking in magma ocean studies (e.g., *Solomatov and Stevenson*, 1993).

Considering these recent advancements in our general understanding of high-pressure silicate liquids and the serious qualitative implications they have for magma ocean evolution, it is clear that we need to take a deeper mechanistic look into the compression behavior of silicate melts. In Chapter 4, I present the development of a general simplified model for oxide liquids. The **C**oordinated **HA**rd **S**phere **M**odel (CHASM) provides a simplified framework for predicting the behavior of oxide and silicate liquids at high pressures and temperatures. This model builds upon the analytic hard sphere liquid model, applied to silicates by *Jing and Karato* (2011), which captures the dominant entropy source for liquids, stemming from structural freedom due to lack of long-range order. On top of the basic hard sphere model backbone, CHASM incorporates an explicit accounting of the variable oxygen coordination states available to each cation in the liquid. The addition of a range of coordination states that are populated according to free energy minimization is crucial to representing gradual coordination evolution, which is the primary compression mechanism of silicate and oxide melts. In addition to presenting the general model framework, we also show how the CHASM approach can be adapted to any chemical system of interest by training its parameters on the energy compression curves for a variety of different crystal structures. We demonstrate this training method and the subsequent calculated results for the MgO system, showing strong agreement with first-principles calculations over a wide range of pressures and temperatures.

Throughout this thesis, I have endeavored to demonstrate the tight links that exist between our knowledge of the thermodynamic properties of rocky materials and our understanding of the Earth's interior. Using a mixture of experimental and theoretical techniques, I have studied the high-pressure and -temperature properties of silicate perovskite and general oxide-silicate melts. Together, these represent the dominant phases of the Earth's mantle, now and in the distant past. Through the work presented in this thesis, I have added to our understanding of the behavior of these important phases under extreme conditions, and thus contributed to our growing knowledge of the deep Earth.

Chapter 2

The Thermal Properties of Iron-bearing Silicate Perovskite and the Implications for Lower Mantle Structures

Abstract

The high-temperature and -pressure equations of state of 0% and 13% Fe-bearing Mg-silicate perovskite are measured using powder diffraction in the diamond anvil cell. The derived equation of state properties are broadly consistent with previous work, though the Mg-endmember shows slightly larger volumes at high pressure than previous studies. The derived equations of state are incorporated into an ideal mixing model used to probe the density stability of perovskite material at a range of temperatures and compositions. This stability analysis shows that the metastable dome theory for Large Low Shear Velocity Provinces (LLSVPs) is possible but unlikely, with a probability of less than roughly 10%, given the narrow range of compositions values consistent with metastability.

2.1 Introduction

The Earth's lower mantle is thought to be composed of primarily aluminous $(\mathrm{Mg},\mathrm{Fe})\mathrm{SiO}_3$ perovskite coexisting with $(\mathrm{Mg},\mathrm{Fe})\mathrm{O}$ ferropericlase. While the exact phase proportions depend on the assumed compositional model for the lower mantle, e.g., pyrolytic vs. picritic, iron-bearing silicate perovskite is thought to dominate making it the most common mineral in the silicate Earth (*Kesson et al.* 1998, *Mattern et al.* 2005). As the primary lower mantle phase, it plays a major role in determining the physical properties of lower mantle material, thereby exerting large control over mantle dynamics and evolution. Characterizing its temperature-dependent equation of state is therefore vital to both interpreting seismic observations and providing accurate input to geodynamics calculations.

Most previous high-pressure experimental studies of silicate perovskite have either focused on the Mg-endmember and neglected the role of Fe (*Fiquet et al.*, 2000), or they have studied a range of Fe compositions but were restricted to ambient temperature (*Lundin et al.*, 2008; *Dorfman et al.*, 2013). There are also numerous studies that have examined natural compositions, including ferric iron and aluminum, which can also have an important impact on the equation of state behavior of perovskite (*Catalli et al.*, 2011). While clearly crucial to our understanding of the deep Earth, such natural compositions add significant complication to interpreting results, and thus it is important to couple them with studies the focus on simpler systems where it is easier to gain an understanding of systematic behavior. The majority of these previously mentioned studies have also used strongly non-hydrostatic pressure media, including NaCl and Ar. Recent work, however, has shown that non-hydrostatic stress states may alter a mineral's compression behavior (e.g., *Takemura*, 2007; *You Shu-Jie and Chang-Qing*, 2009; *Iizuka et al.*, 2010), rendering such experiments perhaps less representative of conditions within the Earth. In this study, we seek to help provide a more complete picture of perovskite by studying a simplified chemical system in great detail, examining iron-bearing and iron-free synthetic samples at both high pressures and temperatures.

In this work, we report the temperature-dependent equation of state for polycrystalline perovskite-structured $(\mathrm{Mg}_{1-x}, \mathrm{Fe}_x)\mathrm{SiO}_3$ samples with two different iron compositions relevant to the lower mantle ($x = 0\%$ and 13%). These samples where probed with x-rays under a wide range of conditions between 30 and 130 GPa and room temperature up to 2500K. These results are then compared to assess the effect of iron composition on silicate perovskite's high-temperature compression behavior. Finally, we assess the plausibility of different competing models to explain perovskite-rich lower mantle structures.

2.2 Methods

Our perovskite samples were made from synthetic $(\mathrm{Mg}_{1-x}, \mathrm{Fe}_x)\mathrm{SiO}_3$ pyroxene starting materials with iron compositions of $x = 0\%$ and 13%. These compositions were verified using micro-probe analysis, and Mössbauer spectroscopy on the Fe-bearing sample showed it to contain undetectable levels of ferric iron, constraining it to less than 3% Fe^{3+} (*Jackson et al.*, 2009; *Zhang et al.*, 2011). The iron-free MgSiO_3 glass starting sample was mixed with roughly 10 wt% Au, which served as an infrared absorber, ensuring good coupling with the laser. To minimize contamination of the sample, the grinding and mixing was performed between two pristine tungsten-carbide cubes. Since their Fe content is sufficient to couple with the laser, the iron-bearing samples were not mixed with gold, avoiding any potential $\mathrm{Fe-Au}$ alloy formation which might affect experimental results; All samples were then loaded into symmetric diamond anvil cells using pre-indented Re gaskets and 250 micron culets. The samples were placed on top of $\sim$10 micron thick insulating dehydrated NaCl plates, which are crucial for reducing thermal gradients within the sample during laser heating. Each sample was also loaded with a few synthetic ruby spheres for offline pressure determination using the ruby fluorescence method (e.g., *Jacobsen et al.*, 2008; *Silvera et al.*, 2007; *Mao et al.*, 1986). The cells were loaded with a Ne pressure medium using the GSECARS gas-loading system at the Advanced Photon Source (APS) of Argonne National Laboratory (*Rivers et al.*, 2008). The samples were pressurized to $\sim$30 GPa where they were then laser annealed within the stability field of perovskite.

Powder diffraction measurements were taken at the Sector 13-ID-D GSE-CARS beamline at the APS. Diffraction spectra were taken in roughly 2 to 4 GPa steps between 35 and 110 GPa for each sample. The pressure for each spectrum was determined using the Ne pressure medium as the primary pressure marker, *Dewaele et al.* (2008), together with the offline ruby fluorescence measurements. As discussed in Section 2.3.2, great care is taken in the analysis of the neon diffraction lines to assess the effect of deviatoric stress on the measured pressure. In roughly 5 to 10 GPa steps, the samples were laser heated in stages up to a maximum temperature of 2500 K. These temperatures were determined from thermal emission spectra by fitting them to a blackbody curve (*Heinz and Jeanloz*, 1987; *Shen et al.*, 2001).

The raw powder diffraction images were converted to one-dimensional spectra using a suite of routines written in MATLAB. In these routines, the observing geometry is first determined from calibration spectra using an automated statistical method. This has the advantage over the standard 'click-based' method employed in FIT2D that it requires little user input and generates a reproducible result that is determined by the maximum likelihood configuration given the data. Using the derived geometric calibration, it is straightforward to calculate the observed diffraction angle (2θ) corresponding to each pixel on the CCD. The diffraction angle can then be converted into the more physically relevant inverse d-spacing ($1/d$), which describes the spacing between a set of coherently reflecting crystal planes, by simply applying Bragg's law for first-order reflections, $1/d = 2\sin(2\theta/2)/\lambda$, where λ is the wavelength of the monochromatic x-rays used to probe the sample. Each raw image is then 'integrated' assuming Poisson statistics for the uncertainties in number of photons hitting each pixel. For some of the diffraction images, which contain over-exposed pixels, a further preprocessing step is required to produce accurate 1D spectra (for more details on our data reduction pipeline see 2.A). The final step in the integration process is to subtract off a decent guess of the background spectrum so that data sets can be easily examined and fit. We employ the Bayesian background identification and subtraction method presented in *David and Sivia* (2001), which automatically determines a reasonable polynomial background curve assuming the potential presence of large positive deviations due to the presence of as-yet unmodeled diffraction peaks.

After these processing steps, we obtain a set of 1D powder diffraction spectra for a range of pressures and temperatures for both the 13% Fe-bearing perovskite sample and the pure Mg-endmember sample. To gain a good overview of the room temperature data, we construct interpolated cold-compression maps shown in Figures 2.1 and 2.2. To best highlight all aspects of the data, we plot the intensity using a geometric color-scale [1], rendering every line visible to the eye even though their amplitudes vary over many orders of magnitude. In each figure, we also overplot the fitted line positions for each phase, showing that dozens of perovskite peaks are visible in the spectra as well as the high-intensity peaks from neon which are used as in situ pressure markers, as described in Section 2.3.2.

2.3 Analysis

The analysis of the 1D integrated spectra is broken into four parts, each of which are each discussed below. First is the process of peak-fitting, which transforms the continuous spectrum into a discrete set of measured peak positions and associated uncertainties. The next task is to determine the pressure–temperature conditions inside the diamond anvil cell. In this section, particular care is taken to determine the deviatoric stress contribution to the pressure. Thirdly, the set of peak positions corresponding to perovskite in each spectrum are transformed into estimates of the unit cell parameters along with uncertainties. Finally, the estimated volumes, temperatures and pressures are all combined to retrieve the confidence regions for the overall equation of state of each sample.

---

[1] We use the inverse hyperbolic sin transformation to visualize the background subtracted intensities, which generates a roughly geometric color-scale. This transformation is well suited to values that vary over orders of magnitude, since it reduces to a simple log-transformation for large values, while still allowing for negative values that result from background subtraction (*Burbidge et al.*, 1988). Note that this transform is only used for visualization and plays no role in the actual data fitting.

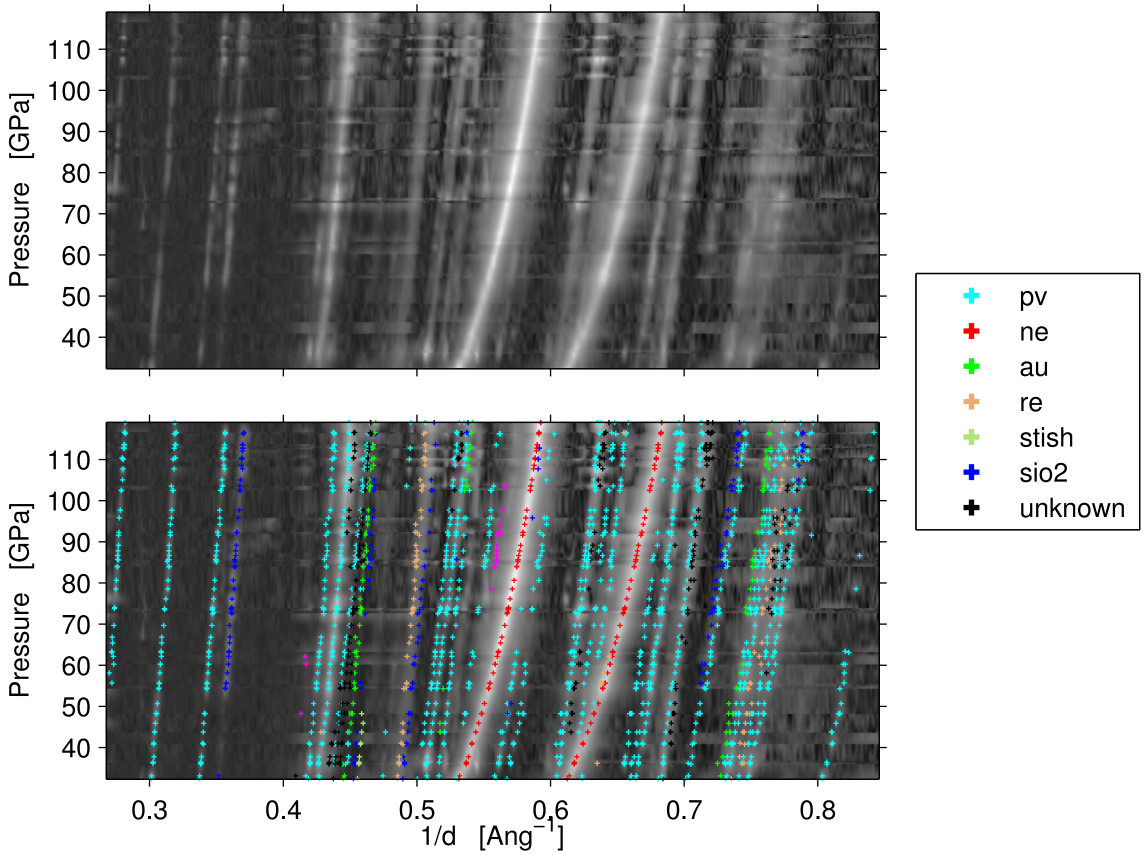

Figure 2.1: Cold compression data map for 13% Fe-bearing perovskite sample showing intensity as a function of pressure and inverse d-spacing. The upper panel shows the raw interpolated intensities, while the lower panel overplots the fitted line positions for the entire data set. Diffraction intensities are shown using a geometric (or roughly logarithmic) scale, allowing even lines of very small amplitude to be visible. The phases with lines apparent in at least some spectra are the perovskite sample (pv), neon(ne), gold (au), the rhenium gasket (re), stishovite (stish), and the post-stishovite $\mathrm{CaCl_2}$ structure of silica (sio2).

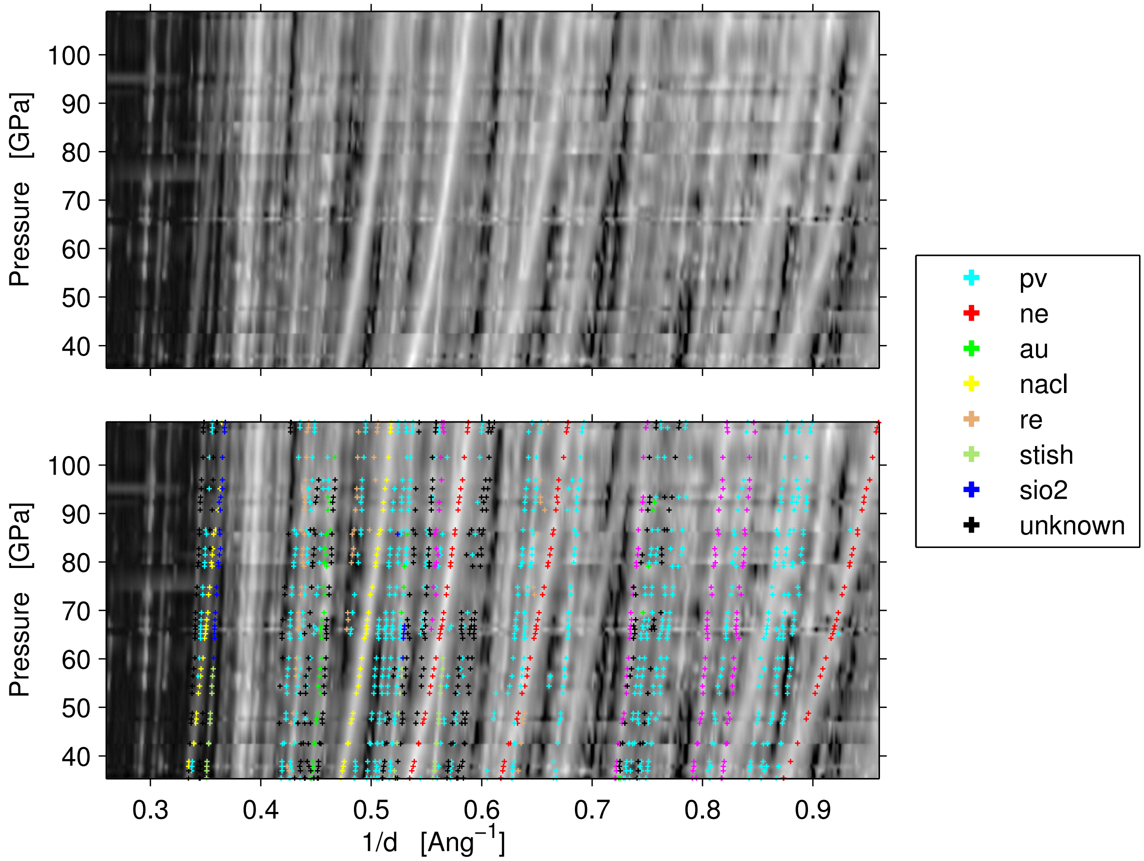

Figure 2.2: Cold compression data map for pure endmember Mg-perovskite sample showing intensity as a function of pressure and inverse d-spacing. The upper panel shows the raw interpolated intensities, while the lower panel overplots the fitted line positions for the entire data set. Diffraction intensities are shown using a geometric (or roughly logarithmic) scale, allowing even lines of very small amplitude to be visible. The phases with lines apparent in at least some spectra are the perovskite sample (pv), neon(ne), gold (au), B2-structured NaCl (nacl), the rhenium gasket (re), stishovite (stish), and the post-stishovite $\mathrm{CaCl_2}$ structure of silica (sio2).

2.3.1 Peak Fitting

The process of obtaining unit cell parameters from a set of integrated 1D diffraction spectra usually follows one of two main strategies: whole pattern refinement or individual peak-fitting. The first method, officially called Rietveld refinement, is carried out by modeling the entire spectrum as a whole, and is thus often referred to as whole pattern refinement. Using this method, every crystal phase that is present in the spectrum is assumed to have a predetermined crystal structure, which dictates the relative line amplitudes. This global model is simultaneously fitted to the entire spectrum, where the free parameters are taken to be the unit cell parameters for each phase (which determine the line positions), the apparent phase fraction determining the relative amplitude of the lines from different phases, and a set of parameters that control the line widths and shapes as a function of diffraction angle. The fitting method follows the standard weighted least-squares, or chi-square minimization approach, which is appropriate to independent normally distributed data as derived using the Bayesian framework in Chapter 3. There exist numerous codes that perform this type of fitting and have graphical user interfaces (GUIs) improving ease of use, including the program GSAS (General Structure Analysis System) with its EXPGUI interface (*Toby*, 2001).

While whole pattern refinement can be quite convenient, it has a number of limitations. As it is typically applied, the user constructs a complete model of the crystal system, which allows the calculation of the entire diffraction pattern including the relative amplitudes of each line in the profile, which are tied to both the overall structure and the relative atomic positions. This is an excellent approach when the data are well represented by an ideal powder diffraction model. In many high-pressure diamond anvil cell experiments, however, the diffraction pattern is seen to evolve with increasing pressure and temperature, causing the relative amplitudes of each line to shift away from their default values expected at ambient conditions. There are many possible explanations for this behavior, including: incomplete sampling of randomly oriented crystallites resulting from small number statistics for crystallites in the diffraction condition; crystal preferred orientation (due perhaps to deviatoric stress in the sample chamber); or perhaps systematic shifts of the atomic positions with increasing pressure. While all of these effects can potentially be directly modeled, they require large amounts of very precise—usually single crystal—diffraction data, making it impractical for the vast majority of experiments.

In the face of this large set of unmodeled physical processes, the whole pattern refinement method is often incapable of accurately representing the relative line intensities in high-pressure experiments. We may choose to ignore such systematic errors in our model, dismissing them as minor effects, but we do so at the risk of inducing strong biases in our results. Therefore, when applying the whole pattern refinement method to high-pressure data, one must be very careful to ensure that the relative amplitudes are well described by the model to within the observational Poisson noise. If they are not, then the best-fit model curves, which may look generally reasonable to the eye, are likely to be heavily weighted toward only a few diffraction peaks within spectrum. In particular, the fit will be depend almost exclusively on only the smallest subset of intense diffraction lines that can constrain the unit cell parameters (representing a range of hkl values). All of the information contained in the clearly visible but low intensity lines can easily be ignored in favor of small yet meaningless improvements to the fit quality for the most intense peaks.

One of the typical ways to deal with this issues is to employ the LeBail fitting method, as implemented in GSAS (*Toby*, 2001). This is an iterative algorithm that allows the relative peak intensities to adjust away from their expected ideal values. It accomplishes this by approximately apportioning the observed intensities amongst the diffraction peaks present at each portion of the spectrum. The problem with this method is that it can often lead to unstable or "runaway solutions", where the fitting procedure must be artificially halted early on, prior to convergence, otherwise the line intensities change so drastically that all of the model parameters, including the dimensions of the unit cell, begin to undertake unrealistically large excursions from reasonable values. This reflects a general property of data fitting in situations where the model parameters are under-constrained—when too many degrees of freedom are given to the model, the fitting algorithm will take advantage of those parameters, adjusting them to reduce the misfit, even if those adjustments are not sensible and lead to unphysical results. The LeBail method is therefore a good choice when the data quality is extremely high and thus there is little potential for bias resulting from over-fitting. Unfortunately, this is often not the case for multiphase high-pressure diffraction spectra, especially when the spectra include low symmetry phases like perovskite.

It is for these reasons that we favor the second method of individual peak fitting. In this approach, the spectrum is split up into different sections that contain diffraction lines of interest. These lines are then fit individually, or in small clustered groups where the peaks are seen to overlap one another. Using this approach, it is simple to leave the relative peak intensities unconstrained, and thus avoid the issues outlined above. This advantage comes with the extra cost of needing to identify each peak in the spectrum to be modeled, which can be fairly time consuming. Of course obtaining reasonably good fits is also correspondingly easier when breaking the spectrum up into smaller sections as compared to fitting the spectrum as a whole. This is typically done by assuming a reasonable baseline model in terms of the unit cell parameters for each spectrum and then locating the diffraction lines that appear to best line up with the predicted peak positions. The major downside to this approach is that some amount of line misidentification is inevitable when observed peaks in the spectrum happen to occur in the expected region of diffraction angle space. In Chapter 3, we discuss in great detail the process of obtaining good estimates of unit cell parameters while avoiding the systematically biasing effects of misidentified peaks.

After obtaining estimates and uncertainties for the crystal dimensions from each spectrum (see Chapter 3), we carry out a careful study of the observed neon diffraction peak positions to determine the pressure inside the diamond anvil cell, account for deviatoric stress effects. Combining the estimates of pressure, temperature, and crystal volume, we estimate the equation of state parameters by fitting the derived unit cell dimensions for each spectrum. Rather than merely reporting the 'best-fit' parameter values as is quite common throughout the literature, we present the correlated uncertainties in our model parameters and demonstrate how important these correlations are crucial to interpreting the difference between different equations of state.

2.3.2 Determining P-T Conditions

Determination of the externally imposed environmental conditions, pressure and temperature, experienced during diamond anvil cell experiments is obviously one of the most important tasks for obtaining accurate and unbiased material properties measurements. Unfortunately, this task is significantly more complicated than one might naively guess. In this section, we discuss how we employ careful statistical methods together with detailed physical models to retrieve the temperature and pressure conditions for each measurement.

In situ temperatures during laser-heating are obtained from measurements of the thermal emission of the sample. The laser heating system at the Sector 13 IDD beamline of GSECARS is equipped with a set of mirrors that simultaneously focus a laser-heating spot on the sample, while carrying the thermally radiated light from the sample back to two independently calibrated spectrometers. The sample's thermal radiation spectrum is then fit at the beamline assuming a gray-body spectrum, enabling estimation of the temperature for both the upstream and downstream sides of the sample. Since the spectrometers on either side are independently calibrated, they provide two independent measurements of the same quantity, assuming small axial thermal gradients. One issue that should be addressed, however, is that the calibration process necessarily leads to some degree of systematic disagreement between the upstream and downstream measurements; as seen in Figures 2.3 and 2.4, the raw temperature measurements for the upstream and downstream sides are systematically shifted from the one-to-one line. This calibration error is readily corrected by plotting the histogram of the difference between the upstream and downstream measurements, which yields a distribution that is reasonably approximated as normal, as shown in the insets in Figures 2.3 and 2.4. The mean and standard deviation of the best-fit normal distribution therefore give the systematic temperature bias and approximate temperature error for each dataset. We subsequently correct the upstream and downstream measurements toward more realistic values by shifting them each by half the typical measurement difference, thereby accounting for any systematic biases in the spectral temperature calibration.

Improving determination of in situ pressures inside the diamond anvil cell has received considerable attention over the past decade, as many workers have focused on increasing the internal consistency amongst the equations of state for numerous pressure markers as well as increasing the absolute accuracy by accounting for deviatoric stresses. Deviatoric stress, which describes the non-hydrostatic component of the stress tensor, arises primarily in response to the uniaxial nature of the compression imposed by the standard diamond anvil cell design, which squeezes the sample only along the axial beam-path direction. The pressure medium surrounding the sample, in this case neon, is therefore relied upon together with the rhenium gasket that forms the walls of the sample chamber, to translate this entirely uniaxial imposed stress into a nearly hydrostatic stress field appropriate to understanding the state of minerals in the deep mantle. The effectiveness of this transformation depends on strength of the pressure medium, hence the preference for ultra-soft media like neon to other harder materials like NaCl.

Based on the uniaxial compression design, it is generally assumed for diamond anvil cell experiments that the dominant part of the deviatoric stress tensor is an excess normal component along the axial direction relative to the two radial directions (*Singh*, 1993), and all other non-hydrostatic contributions are neglected as small high-order corrections. The

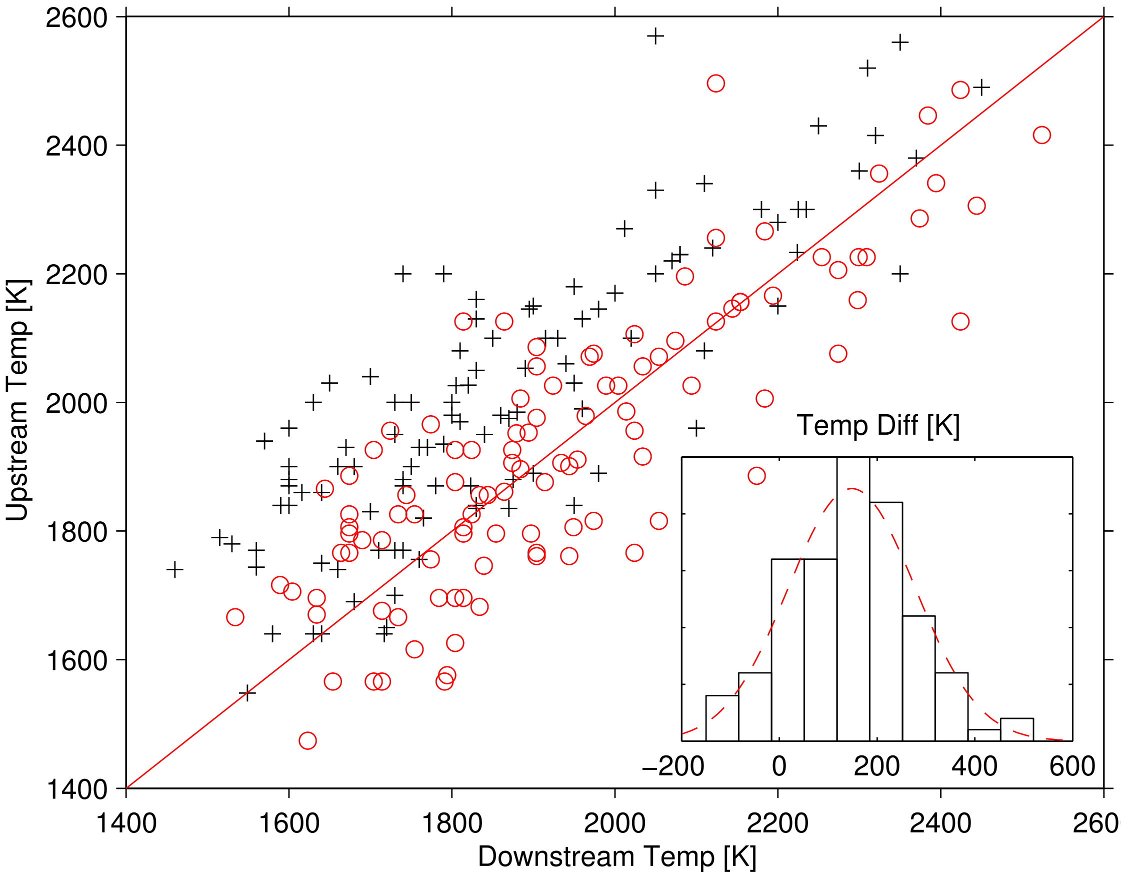

Figure 2.3: Measured spectral temperatures for Fe-bearing sample during laser heating. The main figure plots upstream and downstream measurements against each other together with the one-to-one line in red, with the raw measurements in black crosses and the corrected measurements in red circles. The inset figure shows the distribution of measured temperature differences, consistent with an absolute calibration difference of 150 K between the upstream and downstream spectrometers and a random measurement error for each spectrum of 130 K. This calibration difference is used to correct the raw measurements to their most likely value, shifting the red corrected points toward the one-to-one line.

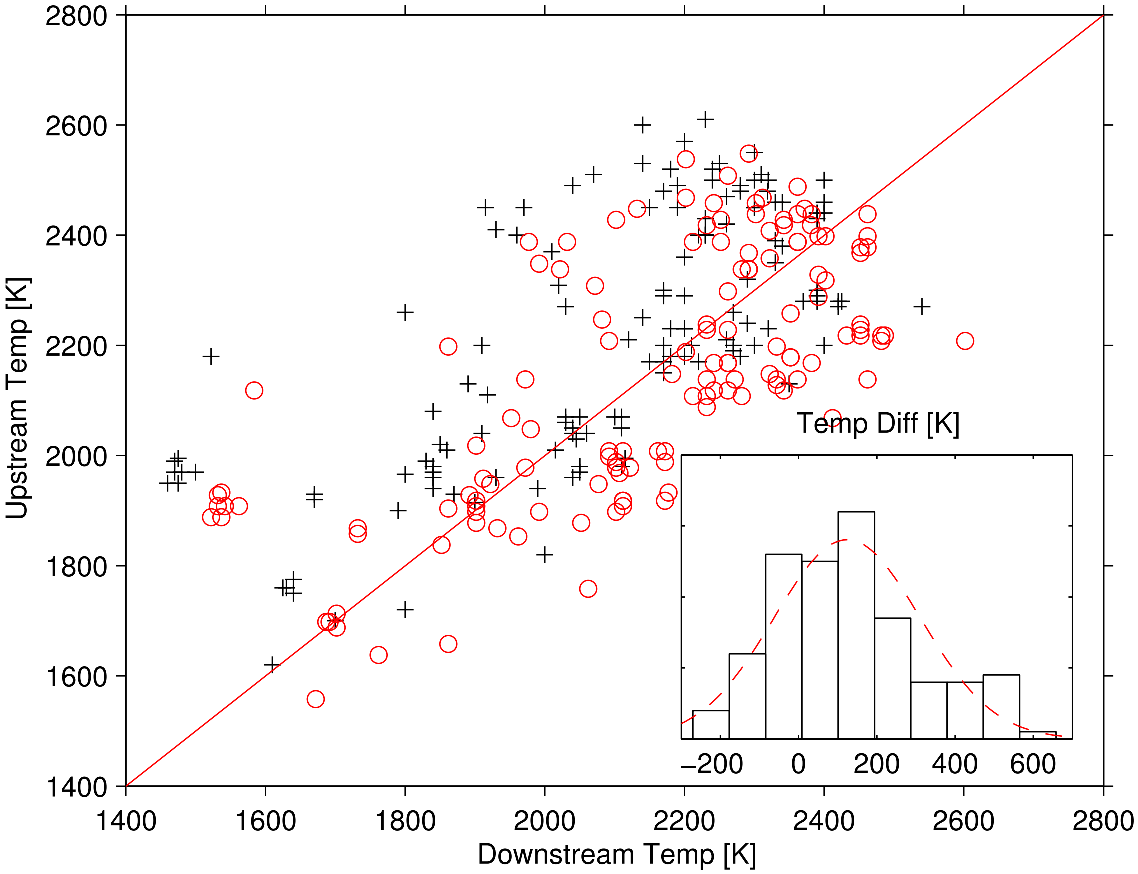

Figure 2.4: Measured spectral temperatures for Mg-endmember sample during laser heating. The main figure plots upstream and downstream measurements against each other together with the one-to-one line in red, with the raw measurements in black crosses and the corrected measurements in red circles. The inset figure shows the distribution of measured temperature differences, consistent with an absolute calibration difference of 125 K between the upstream and downstream spectrometers and a random measurement error for each spectrum of 190 K. This calibration difference is used to correct the raw measurements to their most likely value, shifting the red corrected points toward the one-to-one line.

deviatoric stress experienced inside the sample chamber is bounded above by the yield strength of the pressure medium, which marks the point at which the material will begin to plastically deform and flow, generally allowing the sample chamber to thin along the axial direction, attempting to equalize the stresses along the axial direction (imposed by the diamond anvil surfaces) and along the radial direction (imposed by the edges of the gasket hole).

In this study, we rely primarily on the diffraction lines of the neon pressure medium to determine in situ pressures. While multiple pressure markers were placed inside the sample chamber, including NaCl and Au, the majority of the powder diffraction spectra show relatively weak signals from those markers, making their pressure determinations highly uncertain. In contrast, the diffraction lines from neon represent the most intense reflections in every spectrum.

Using the peak-fitting procedure described above and in Chapter 3, we are able to accurately retrieve peak positions for both the neon 111 and 200 lines in every diffraction spectrum, as well as the neon 220 line in most of the Mg-endmember sample spectra. For a cubic crystal under hydrostatic stress, the crystal plane spacing is given simply as

$$d_P(hkl)^{-1} = a_P^{-1}\sqrt{h^2 + k^2 + l^2} \tag{2.1}$$

where $d_P(hkl)$ is the d-spacing under purely hydrostatic stress and $a_P = V^{\frac{1}{3}}$ is the unit cell length. An estimate for the in situ pressure can then be determined by obtaining the apparent unit cell length for each diffraction line and converting the corresponding volume into a pressure using the well-determined equation of state for the pressure standard in question. In our case, we use the neon equation of state reported in *Dewaele et al.* (2008), which provides a carefully constrained Debye-Grüneisen equation of state using high-cadence room-pressure compression data up to 200 GPa and precise resistive heating high-temperature measurements between 300K and 1000K. We find that the measured neon line positions imply apparent unit cell sizes, and corresponding pressure values, that are systematically shifted relative to one another, reflecting the presence of non-negligible deviatoric strain. By assuming that excess uniaxial stress dominates the deviatoric stress tensor, *Singh* (2009) provide the equations that describe the effect that deviatoric strain has on the measured lattice plane spacings from powder diffraction experiments:

$$\begin{aligned} d_m(hkl) &= d_P(hkl)[1 + (1 - 3\sin^2\theta)Q(hkl)] \\ Q(hkl) &= \frac{t}{6G^X(hkl)} \end{aligned} \tag{2.2}$$

where $d_m(hkl)$ is the measured d-spacing of the hkl line and $Q(hkl)$ is a scaled ratio of the deviatoric stress t to the "diffraction" shear modulus, $G^X(hkl)$.

The diffraction shear modulus expresses the average value for a polycrystalline aggregate averaged over the subset of crystallites that contribute to the measured diffraction intensity. Since only crystallites in certain orientations satisfy the diffraction condition and contribute to the measured diffraction peaks, the diffraction shear modulus varies depending on the set of crystal planes, and is therefore a function of the hkl values. Following the typical approach, the average aggregate elasticity is assumed to be a weighted harmonic mean of the Voigt limit (derived assuming stress continuity along grain boundaries) and the Reuss limit (derived assuming strain continuity along grain boundaries). Accordingly, the diffraction shear modulus is assumed to follow the empirical form:

$$[G^X(hkl)]^{-1} = \alpha[G^X_R(hkl)]^{-1} + (1-\alpha)G_V^{-1} \tag{2.3}$$

where $G^X_R(hkl)$ is the Reuss limit for the diffraction shear modulus, G_V is the Voigt limit for the shear modulus, and α is the weighting factor typically assumed to fall between 0.5 (equal weighting of Voigt and Reuss bounds) and 1 (the Reuss limit). Note that $\alpha = 0.5$ is not the Voigt-Reuss-Hill average, which is simply the arithmetic mean of the Voigt and Reuss bounds, since Equation (2.3) uses a harmonic mean. Nevertheless, *Hill* (1952) points out that the choice of averaging scheme is arbitrary and should be chosen empirically. For diffraction measurements of shear moduli, the harmonic mean has been shown to effectively capture the elastic state of polycrystalline aggregates (*Singh*, 2009). The expressions of these limits for a cubic material are given by *Singh* (2009) as:

$$\begin{gathered} G_V = [C_{11} - C_{12} + 3C_{44}]/5 \\ [G^X_R(hkl)]^{-1} = 2[S_{11} - S_{12} - 3S\Gamma(hkl)] \\ \Gamma(hkl) = (h^2k^2 + k^2l^2 + l^2h^2)/(h^2+k^2+l^2)^2 \end{gathered} \tag{2.4}$$

where $S \equiv S_{11} - S_{12} - S_{44}/2$ is the elastic anisotropy, S_{ij} are the elastic compliances, and C_{ij} are the elastic stiffnesses, which are just related by a matrix inversion ($\mathbf{C} = \mathbf{S}^{-1}$). When written in this form, it is clear how stress continuity implies simple additivity of stiffnesses, which have units of pressure, whereas strain continuity implies simple additivity of compliances.

In order to evaluate the diffraction shear modulus, we must be able to estimate the elastic moduli for each spectrum. A theoretical study of the elastic properties of noble gas solids by *Tsuchiya and Kawamura* (2002) showed that the normalized elastic constants, $C'_{ij} = C_{ij}/K$, of Ne, Ar, Kr, and Xe are all nearly constant over a wide pressure range between 0 and 200 GPa. *Shimizu et al.* (2005) performed Brillouin spectroscopy measurements of solid neon between 5 and 7 GPa and verified that the normalized elastic constants are approximately independent of pressure, though the measured values are systematically shifted from the theoretical calculations. As discussed by *Tsuchiya and Kawamura* (2002), this is likely the result of inadequately accounting for many-body interactions at finite temperatures. We therefore estimate the elastic properties of neon by assuming that the relative elastic constants are independent of pressure and given by the values measured by *Shimizu et al.* (2005): $C'_{11} = 1.13$, $C'_{12} = 0.93$, and $C'_{44} = 0.40$. It is important to note that while vibrational measurements are sensitive to the adiabatic elastic constants, due to the high frequencies involved, deviatoric strain measurements are sensitive to the isothermal elastic constants. Therefore, we can obtain the appropriate unnormalized elastic constants using $C_{ij} = C'_{ij} K_T$ and $K_T = (C_{11} + 2C_{12})/3$.

The approach adopted by most workers—including *Singh* (1993); *Takemura and Dewaele* (2008); *Dorfman et al.* (2012)—is to make a number of approximations in order to render Equations (2.2) to (2.4) in a convenient linear form. These involve dropping what are assumed to be small correction terms. We note, however, that such approximations are potentially undesirable, as well as unnecessary. Instead, we recognize that Bragg's law gives the direct relation between the diffraction angle and inverse d-spacing for each line,

$\sin\theta_m = 0.5\lambda d_m^{-1}$. Substituting this into Equation (2.2) and algebraic manipulation results in a simple cubic equation for the modeled inverse d-spacing, $d_{\mathcal{M}}^{-1}$, of each line:

$$0 = \frac{\lambda^2 t}{8G^X(hkl)} d_{\mathcal{M}}^{-3} + \left(1 + \frac{t}{6G^X(hkl)}\right) d_{\mathcal{M}}^{-1} - \frac{\sqrt{h^2+k^2+l^2}}{a_p} \tag{2.5}$$

recalling that $G^X(hkl)$ depends on the value of α and is given by Equations (2.3) and (2.4). Since this is just a simple cubic equation, it can be solved trivially using any standard polynomial root-finding method, yielding the modeled inverse d-spacing $d_{\mathcal{M}}^{-1}$ as a function of the unknown axial deviatoric stress t and unit cell length a_P.[2]

It is now just a simple matter of fitting the deviatoric stress and neon volumes to each set of measured lines. Since each spectrum contains only 2 or 3 neon lines, fitting these values without imposing further constraints leads to a rather large scatter amongst the determined deviatoric stress values, due to the large correlations between the fitted deviatoric stress and unit cell size. Figure 2.5 shows the unconstrained coupled determinations of deviatoric stress and pressure (evaluated using the determined unit cell size together with the equation of state from *Dewaele et al.* (2008)). While there is large apparent scatter, it is clear that the deviatoric stress is confined to low values of less than a few GPa. Further constraints can be imposed on the stress state retrieval by noting that the relative deviatoric stress fraction, stress over pressure, shows a clearly preferred value of around 0.5%, as indicated in the inset histogram in Figure 2.5. Assuming such a model where the axial deviatoric stress is linearly proportional to pressure is also reasonable given that deviatoric stress

---

[2] It should be noted that while there are 3 potential solutions to any cubic equation, only one of them gives a reasonable value for $d_{\mathcal{M}}^{-1}$, while the other two solutions are unphysical having either extremely large, negative, or imaginary values and are thus easily identified.

typically scales roughly with bulk modulus, which is itself roughly linear in pressure. We therefore employ this model as a prior constraint on the deviatoric stress and refit the set of neon line positions to obtain the best constant for the deviatoric stress fraction for each dataset, yielding values of 0.5% and 0.6% for the Mg-endmember and Fe-bearing samples, respectively, in agreement with the central peaks of the apparent deviatoric stress fraction histograms. Through this analysis, we obtain final determinations of the pressure and its error, accounting for deviatoric stress.

2.3.3 Extracting Crystal Volumes and Cell Dimensions

After obtaining line position estimates and errors from the powder diffraction spectra, the next step is to fit these positions with a crystal reflection model to estimate unit cell dimensions and volumes. We use the Pbnm crystal structure determination for Mg-perovskite from *Sugahara et al.* (2006) to model the line positions for our sample. From this structural refinement, we obtain a list of the hkl values for the reflections expected to be present in the perovskite spectrum[3]. For an orthorhombic crystal symmetry, it is then trivial to calculate line positions from the reflection list using the following equation:

$$\left(\frac{1}{d}\right)^2 = \left(\frac{h}{a}\right)^2 + \left(\frac{k}{b}\right)^2 + \left(\frac{l}{c}\right)^2 \tag{2.6}$$

where hkl are the reflection order parameters, abc are the crystal unit cell parameters, and d is the atomic plane spacing.

By applying the Bayesian mixture model presented in Chapter 3, we obtain reason-

---

[3] We also obtain a list of relative peak intensities at ambient pressure that can be used as a rough order-of-magnitude guide for the high-pressure data, which is helpful in line identification

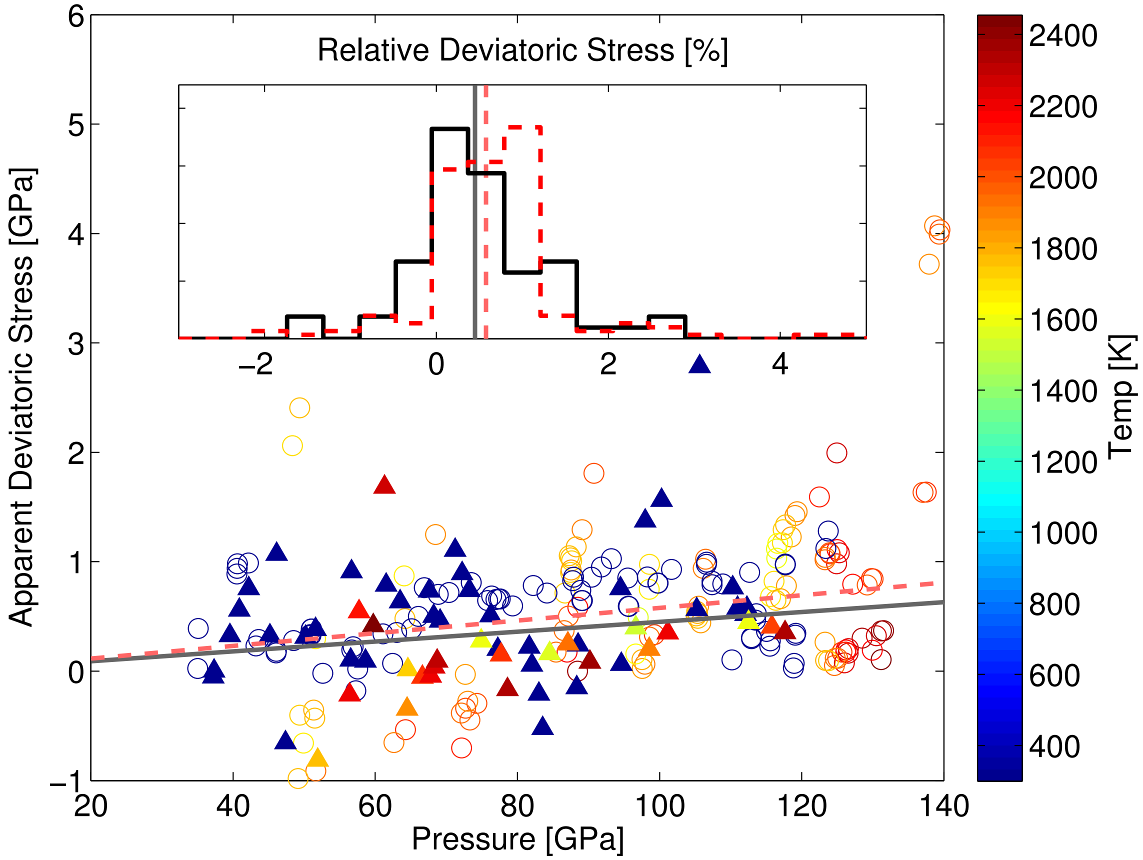

Figure 2.5: The apparent deviatoric stress is plotted against the neon pressure and color coded by temperature, obtained by analyzing the 111, 200, and 220 lines of neon. By fitting the two or three neon lines apparent in each spectrum, we obtain unconstrained simultaneous estimates of the axial deviatoric stress and the neon unit cell size, which are converted into pressure estimates using the neon equation of state from *Dewaele et al.* (2008). While there is large apparent scatter, the relative deviatoric stress, as a fraction of total pressure, shows a preferred value of roughly 0.5%, as seen in the inset histogram. The Mg-endmember sample is shown in solid black and filled symbols, while the 13% Fe sample is shown in dashed red and open symbols.

able estimates for the unit cell dimensions of the perovskite sample for each spectrum. As mentioned in Chapter 3, the choice of parameterization for any statistical analysis problem is important in order to ensure reasonable retrievals, since the parameterization is linked to the prior probabilities. The parameters we choose to fit are the volume of the unit cell $V = abc$, and the normalized axial ratios $(c/a)^* = c/\sqrt{2}a$ and $(b/a)^* = b/a$. Given this parameterization, it is simple to extract the pressure-volume curve as well as track the change in the axial ratios with compression, which are limited to the few percent level. Using the standard uncertainty estimation method, also described in 3, we can obtain reasonable estimates for the measurement errors on volume and the axial ratios.

2.3.4 Obtaining the Equation of State from P-V-T Measurements

Having determined the sample volume for a range of pressure conditions both at ambient and elevated temperatures, we are now prepared to evaluate the high-temperature equation of state by fitting the P-V-T data. This process is performed in two steps: first we obtain the cold portion of the equation of state using only ambient temperature measurements. We then fit the remaining laser-heated data to estimate the uncertainty bounds on the hot parameters. In the following section, we present the method we use to fit these data and obtain rigorous uncertainty bounds on the parameters.

In order to convert measured volumes sampled at a discrete set of pressures and temperatures into a form useful for interpolation and extrapolation, we must have a theory-based description of the equation of state. We use the simple thermal pressure approximation, which says that the free energy of a material can be divided into cold and hot components, which lead to separate cold and hot contributions to the pressure since $P(V,T) = \partial F(V,T)/\partial V_T$. For convenience, the cold contribution to the pressure is often defined with reference to ambient temperature conditions, $T_0 = 300K$, rather than absolute zero. This requires that we subtract off the nonzero thermal contribution at the reference temperature, as indicated below:

$$P(V,T) = P_c(V) + P_{th}(V,T) - P_{th}(V, 300\text{K}) \tag{2.7}$$

where P_c is the "cold" portion of the total pressure defined at 300K, described below, and P_{th} is an expression for the thermal contribution, for which we use a Mie-Grüneisen-Debye formulation.

At ambient temperature conditions, and in the absence of phase transitions, most solid materials are well described by a Vinet equation of state (*Vinet et al.*, 1989). In particular, for data spanning a large pressure range, *Cohen et al.* (2000) showed that the Vinet equation of state yields more accurate extrapolation behavior over a large compression range as compared to the more commonly used third-order Birch-Murnaghan equation of state. The Vinet equation is derived from an approximate universal representation of the shape of the inter-particle energy well, which yields the following expression for the "cold" contribution to the pressure as a function of volume:

$$\begin{gathered} P_c(x) = 3K_0(1-x)x^{-2}\exp\left[\nu(1-x)\right] \\ \text{where} \quad x = (V/V_0)^{\frac{1}{3}} \quad \text{and} \quad \nu = \frac{3}{2}(K_0' - 1) \end{gathered} \tag{2.8}$$

where x is the average axial strain, V_0 is the zero-pressure volume, K_0 is the zero-pressure bulk modulus, and K_0' is the pressure derivative of the bulk modulus at zero pressure ($K' \equiv \partial K/\partial P$).

Somewhat more complex is the representation of the thermal component of pressure, which relies upon the simplified Debye crystal model to approximate the energetic contribution of thermal vibrations in a crystalline solid. The Mie-Grüneisen-Debye expression for the thermal pressure is given below:

$$P_{th}(V,T) = \frac{C_{vm}\gamma}{V} T D\left(\frac{\Theta}{T}\right) \tag{2.9}$$

where $C_{vm} = 3RN_{\text{cell}}$ is the Dulong-Petit high-temperature limit for the volumetric heat capacity of the unit cell, γ is the thermodynamic Grüneisen parameter, and Θ is the Debye temperature. The function $D(x)$ is the Debye integral, which together with the Debye temperature determines how rapidly the heat capacity approaches its high-temperature limit, and is defined as

$$D(x) = \frac{3}{x^3}\int_0^x \frac{y^3 dy}{e^y - 1} \tag{2.10}$$

where this integral must be evaluated numerically. What remains is an expression for the Grüneisen parameter, as well as a corresponding expression for the Debye temperature. A common expression for the Grüneisen parameter is in the form of a power law in volume, reflecting the fact that it is roughly independent of temperature, sensitive primarily to compactness of the crystal lattice:

$$\begin{aligned}
\gamma(V) &\equiv -\frac{d\ln\Theta}{d\ln V} = \gamma_0(V/V_0)^q \\
\Theta(V) &= \Theta(V_0)\exp\left[-\int_{V_0}^{V}\frac{\gamma(\nu)}{\nu}d\nu\right] \\
&= \Theta_0\exp\left[\frac{\gamma_0}{q}\left(1-\left(\frac{V}{V_0}\right)^q\right)\right]
\end{aligned} \tag{2.11}$$

where γ_0 is the Gruneisen parameter at zero pressure, Θ_0 is the Debye temperature at zero pressure, and q is the power law exponent.

Given the complete equation of state formulation described above, we can determine the parameter values by fitting the measured perovskite volumes for each spectrum. It should be noted that the errors associated with this fitting procedure appear on both dependent and independent variables P, T, and V. This situation only adds a minor complication to the fitting process, since we can use the “effective variance method” to estimate effective error bars appropriate for use in the standard least-squares method (*Orear*, 1982). This method is based on the assumptions that the model being fit to data behaves reasonably linearly with respect to the model parameters over the “high probability” region of parameter space (i.e., a first-order Taylor expansion of the model about the best fit is a good local approximation). The result of this method (which is derived using a Bayesian maximum likelihood estimation procedure) is a simple and intuitive expression for the effective variance, which can be used in place of the measurement error in a standard least-squares analysis:

$$\sigma_{\text{Tot}}^2 \approx \sigma_P^2 + \left(\frac{\partial P}{\partial T}\right)^2\sigma_T^2 + \left(\frac{\partial P}{\partial V}\right)^2\sigma_V^2 \tag{2.12}$$

where $\partial P/\partial T$ and $\partial P/\partial V$ are simply model derivatives evaluated at the location of the best fit. For the effective variance method, we have the freedom to choose for the dependent variable whichever variable is most convenient, since all possible choices result in expressions that are identical to first order. We chose to consider P as the dependent variable, as it is much more convenient given that the EOS models are written with pressure as functions of volume and temperature, and these functions cannot be analytically inverted. The form of the effective variance is also quite attractive, as it is identical to the expression for the propagation of independent error sources, a rather intuitive way of viewing the combination of these different error contributions.

We combine this procedure for determining total effective errors, with the analysis method described in Chapter 3 to obtain the best-fit equation of state parameter values, along with the associated confidence regions. The best-fit values of the parameters, together with one-dimensional marginalized uncertainties are found in Table 2.1—while the precise values presented in that Table correspond to the global fit to all ambient and heated data, the hot data contain little extra information on the cold parameters and therefore alter their best-fit values by less than the 1-σ uncertainties. The cold compression data are plotted together with the 68$\%$ confidence interval on the equations of state in Figure 2.6. From that figure, it is clear from the narrower confidence bounds on the compression curve that the Fe-bearing data has reasonably higher accuracy than the Fe-free data. Nevertheless, these two compression curves are extremely similar to one another, with the Fe-bearing curve showing only slightly greater compressibility, which is reflected in the slightly smaller values of both K_0 and K_0'. It is important not to overstress this difference, as the data clouds

Table 2.1: Vinet and Mie-Gruneisen-Debye Equation of State Parameters

X_{Fe}	0%	13%
V_0	162.51(19)	163.23(18)
K_0	259.0(56)	254.0(38)
K_0'	4.28(17)	4.060(85)
γ_0	3.19(25)	1.579(93)
q	2.16(43)	0.63(27)

We fix the zero-pressure Debye temperature to the value from *Anderson* (1998), $\Theta_0 = 1100K$, since our high-temperature data all significantly exceed this value and are thus largely insensitive to it.

strongly overlap one another, reflecting only small changes to the "cold" equation of state parameters. Discussion of the comparison with previous equation of state measurements, as shown in the inset figure, is deferred to the Discussion section.

The thermal equation of state parameters are obtained by fitting the laser-heated measurements including the thermal pressure contribution, given by the Mie-Grüneisen formulation in Equation 2.9. This method introduces three additional "hot" EOS parameters, Θ_0, γ_0, and q. However, the sample behavior is insensitive to the zero-pressure Debye temperature, Θ_0, for very high temperatures well above its value, where the heat capacity is well approximated by the Dulong-Petit high-temperature limit. Since all laser-heated measurements exceed its nominal value of $\Theta_0 = 1100K$ (*Anderson*, 1998), we fix it to this value throughout this investigation. Note that this is fundamentally different from fixing data sensitive parameters like $K'0$, since they have a clear effect on the equation of state over the region of interest, whereas Θ_0 has almost no effect over the temperature ranges of either the data or the lower mantle, and therefore negligible correlation with the other equation of state parameters.

The proper method for incorporating the results from the cold parameter fit uses the

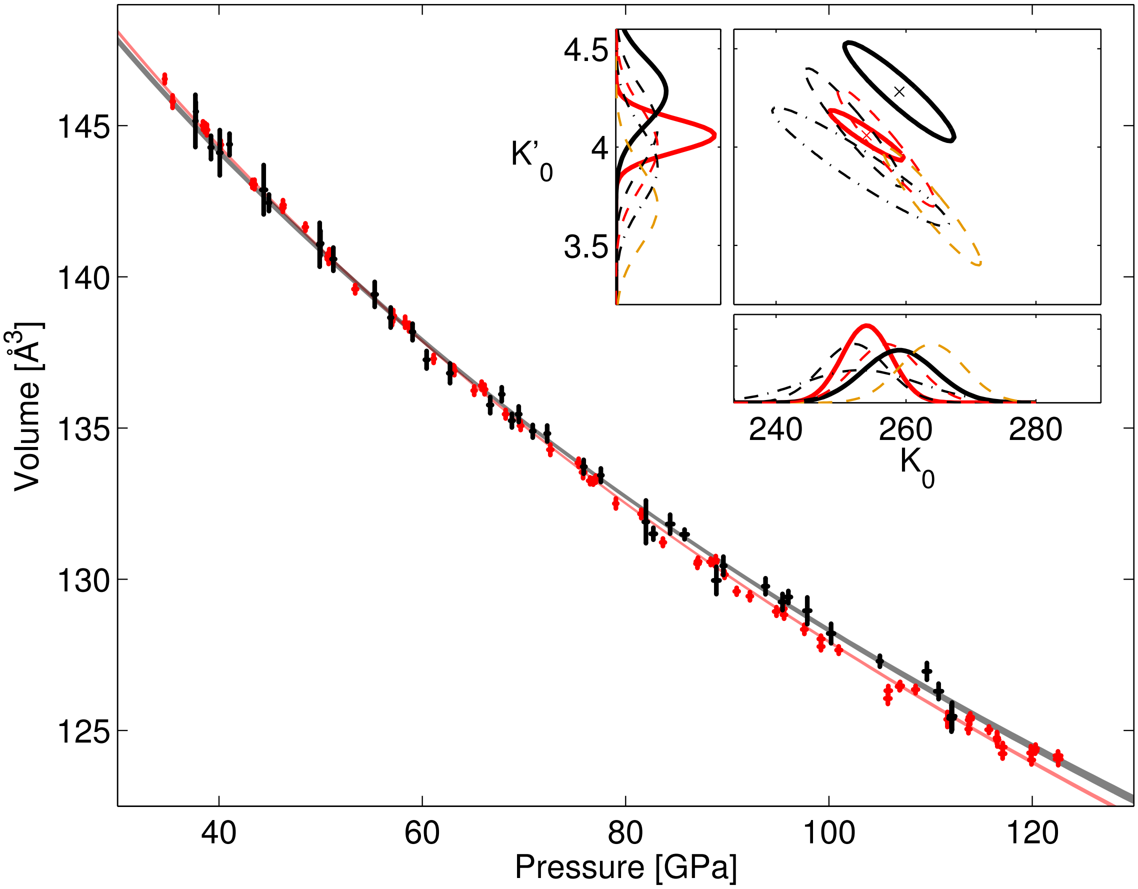

Figure 2.6: The cold compression curves for 0% and 13% Fe-bearing perovskite samples in black and red, respectively. The pressure and volume error bars reflect the 1-σ uncertainties as determined from the measurement errors. The pale shaded regions indicate the 68% confidence bounds on the equation of state, showing the somewhat higher confidence in the Fe-bearing equation of state. The inset figure compares the highly correlated 68% confidence bounds on K_0 and K_0' with past investigations. The 0%, 9%, and 13-15% Fe-contents are shown in black, orange, and red. The uncertainties from this work are shown in heavy solid lines, (*Lundin et al.*, 2008) fits are shown in dashed lines, and (*Fiquet et al.*, 2000) in dot-dashed lines. (It should be noted that this comparison mixes third-order Birch-Murnaghan and Vinet equations of state. Also the confidence regions for previous work are estimated assuming similar parameter correlations across studies).

Table 2.2: Equation of State Parameter Correlations

0%

	V_0	K_0	K_0'	γ_0	q
V_0	$+1.00$	-0.72	$+0.43$	-0.02	-0.07
K_0	-0.72	$+1.00$	-0.93	-0.24	-0.14
K_0'	$+0.43$	-0.93	$+1.00$	$+0.34$	$+0.26$
γ_0	-0.02	-0.24	$+0.34$	$+1.00$	$+0.96$
q	-0.07	-0.14	$+0.26$	$+0.96$	$+1.00$

13%

	V_0	K_0	K_0'	γ_0	q
V_0	$+1.00$	-0.92	$+0.75$	$+0.06$	$+0.06$
K_0	-0.92	$+1.00$	-0.94	-0.27	-0.27
K_0'	$+0.75$	-0.94	$+1.00$	$+0.42$	$+0.45$
γ_0	$+0.06$	-0.27	$+0.42$	$+1.00$	$+0.97$
q	$+0.06$	-0.27	$+0.45$	$+0.97$	$+1.00$

For those unfamiliar, the correlation matrix is just a convenient scaled form of the covariance matrix which enables easy interpretation, where $\mathbf{\Sigma}$ is the covariance matrix and the correlation coefficient is simply $\rho_{ij} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$

Bayesian concept of a prior, which is discussed thoroughly in Chapter 3. The prior simply expresses the probability distribution for the parameters prior to any particular step of the analysis. Thus, we can use the output probabilities from fitting the ambient temperature data as the input to the heated data modeling step. Note that many investigators will fix the cold parameters to their best-fit values when fitting the hot parameters. This is both unnecessary and undesirable, since it underestimates the uncertainties in the hot parameters by neglecting potential covariances between the cold and hot parameters. Instead, we fit the heated data points with a full thermal model, recognizing that nearly all of the information about the cold parameters comes from the cold parameter posterior (converted into a prior for the hot data analysis). As far as the prior on the hot data is concerned, we choose to use uninformed flat priors since we have significantly more simultaneous high-pressure high-temperature data than any previous studies *Fiquet et al.* (2000), and all previous studies chose to fix one or more significant parameters, biasing their uncertainty estimates to unrealistically small values.

The resulting best-fit parameters, with their marginal 1-σ uncertainties, are reported in Table 2.1. To summarize all the information for the posterior, we also report the correlation matrix for both samples in Table 2.2. The residual volumes due to thermal expansion are shown for each sample in Figure 2.7, where the best-fit cold equation of state has been subtracted off for the purpose of visual inspection, leaving only the residual thermal expansion signal. Data are shown with effective error bars in volume, allowing for simple visual inspection, and are determined using the effective variance method with volume as the dependent variable. The volumetric measurement error is determined using the Bayesian method discussed in Ch. 3 and the pressure and temperature contributions to the effective error are derived in Section 2.3.2. Since both panels in Figure 2.7 are show with the same axis limits, its plainly visible that the effect of iron is to dramatically reduce the thermal expansion at high pressures. This observation is directly reflected in the best-fit estimates for the thermal parameters γ_0 and q, which relate to the overall scale for the thermal expansion and its compression (or Volume) dependence. Comparing the figure with the retrieved values, we can see that the Fe-free sample has both a greater average thermal expansion (larger γ_0), yielding wider spacing between isotherms, and a greater compression dependence (larger q), resulting in steeper isotherm slopes. We can therefore see that while the addition of Fe has only a minor effect on the cold equation of state parameters, it has a huge impact on the thermal properties of silicate perovskite.

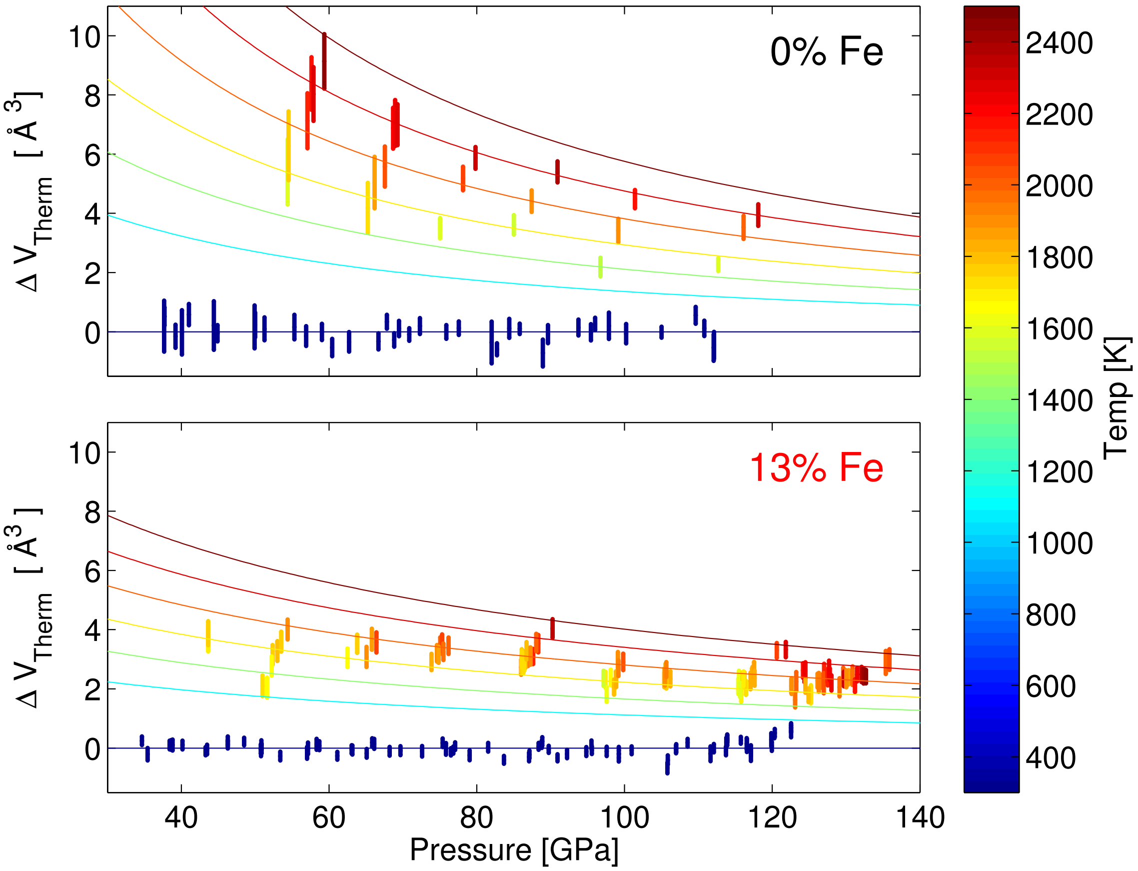

Figure 2.7: Excess volumes due to thermal expansion are shown for the 0% and 13% Fe-bearing samples color coded by temperature, together with the best-fit isotherms. The strong decrease in the apparent thermal expansion due to the addition of iron is clearly visible in the noticeably smaller envelope for the elevated temperature isotherms.

2.4 Discussion

Given the posteriors estimated for both Fe-free and 13% Fe-bearing perovskite, we are now ready to compare our results with previous investigations. Most importantly, we evaluate how apparent differences across studies compare with the correlated uncertainties associated with each study. After that, we then use our posterior estimates to determine the variation of important physical quantities like density, bulk modulus, and thermal expansion for a range of pressures and temperatures. Finally, we incorporate the equation of state determinations into an ideal mixing model in order to evaluate the plausibility of different perovskite-rich compositional models for deep mantle structures.

2.4.1 Equation of State Comparison and Uncertainties

Using the general Bayesian approach of estimating covariance matrices from the curvature of the posterior in log-space, we obtain the 1D uncertainty bounds for each parameter together with the correlation matrix presented in Tables 2.1 and 2.2. These results can be directly visualized using a “stair-step” plot, which is presented in Chapter 3, Figure 3.5, which shows the correlated uncertainty bounds on each pair of parameters by plotting every parameter against every other parameter. The stair-step plot is just a way of visualizing the information conveyed by the best-fit and covariance matrix. Looking at the elements of the correlation matrix in Table 2.2, we can get much of the same information, noting in particular the highest correlation of any two parameters exists between the thermal parameters γ_0 and q, reflecting how difficult it is to independently constrain their values. In spite of that difficulty, the large quantity of thermal data for the Fe-bearing sample has very effectively constrained the thermal parameters to a correlated but small region of parameter space, as plainly shown in the lower-right-hand corner of Figure 3.5;

These results should then be compared directly to previous investigations of the perovskite equation of state. Unfortunately, this task is made rather difficult for two reasons. Firstly, none of the previous investigations report the full covariance matrix (or equivalent correlation matrix) for their equation of state parameters. Second, and more important, is the fact that all previous investigations fixed a number of important equation of state parameters in order to improve the apparent constraints on the remaining parameters. The issues with this common but incorrect practices are discussed thoroughly in Chapter 3. Therefore, we must settle for the rough comparison shown in the inset of Figure 2.6 where the correlated 68% confidence regions on the bulk modulus K_0 and its pressure derivative K_0' are shown for this study in solid lines as well as *Fiquet et al.* (2000) in dot-dashed lines and *Dewaele et al.* (2008) in dashed lines. All of the perovskite samples are color coded according to composition as before, with 0% Fe in black, 9% Fe in orange, and 13-15% Fe in red. As the figure indicates, all studies are consistent with bulk moduli near 255 GPa for a wide range of iron contents up to 15%. To give the previous studies reasonable correlations between the parameters (they do not report covariance matrices), we simply apply the average correlation coefficient from this study. Unfortunately, the effect of previous studies fixing parameters remains, where the ellipse sizes are likely significantly smaller than they should be. As a result of the large correlated uncertainties in previous studies, trends of K_0' with iron content are obscured. In this study, we do find a discernible decrease in K_0' with increasing iron, but the magnitude of the change is rather small, as the cold compression data still largely overlap one another.

While considerable attention is always given to comparing EOS parameters, in reality, we are most interested in the behavior at mantle-relevant pressures, rather than the ambient-pressure properties expressed by the equation of state parameters. For this reason, we propagate our posterior uncertainties in order to predict the behavior at lower mantle conditions, by drawing randomly from the posterior and calculating properties at elevated pressure states. In Figure 2.8, we plot the 68% correlated confidence regions for this study, along with rough estimates of the same for previous studies. In that figure, we can see how previous studies have not uncovered any strong evidence for systematics that may exist for perovskite as a function of iron content, If anything, they find hints that the addition of iron causes the high pressure volume of perovskite to increase (shifting the curve upward), but has no discernible effect on bulk modulus—in contrast, we find the opposite, that adding iron can slightly reduce high pressure volumes, though volumes at ambient pressure increase. It should also be noted, however, that true confidence ellipses are larger than depicted for these other studies, implying that their true scatter could be larger than appears.

While these results were obtained using reasonable priors on the zero-pressure volume data, it is important to realize that they inherently rely on the assumption that metastable volume measurements of perovskite are relevant to its behavior within in stability region. It has not been well established how similar the behavior of perovskite is inside and outside its stability range, but there are some indications that the axial ratios may undergo anomalous deviations near the edge of stability. It is also common for minerals to undergo softening as they approach a stability limit, and therefore it is reasonable to expect that metastable measurements provide, at best, only a rough idea of the compression behavior. For this reason, we also refit the non-laser-heated data in order to test how the apparent core-mantle boundary properties of perovskite depend on the usage of metastable data. In the thick dashed lines in Figure 2.8, we plot the 68% confidence regions resulting from uniformed priors on V_0, thereby removing all influence of metastable data. Under this alternate hypothesis, we can clearly see a 5% to 10% increase in the apparent bulk modulus of perovskite at the core-mantle boundary, though there is little change in the volume. Additionally, there is a large inflation of the uncertainties, resulting from the reduced constraint on the equation of state model, resulting in a reduction of the CMB differences between iron-bearing and iron-free samples allowing the confidence regions to partially overlap. This result matches our intuition from Figure 2.6, in which the two data clouds are seen to largely overlap one another. Unlike for the cold parameters discussed here, however, the usage or absence of the metastable data has little effect on the thermal equation of state parameters.

As investigated by other authors, we also compare the evolution of the crystal axial ratios with increasing pressure. Past work (*Lundin et al.*, 2008; *Dorfman et al.*, 2013) have found that the addition of iron causes a noticeable change in the normalized axial ratios, which are generally observed to grow roughly linearly with pressure. The normalized unit cell parameters derived by *Andrault et al.* (2007) are: $a^* = a(V/\sqrt{2})^{-1/3}$, $b^* = b(V/\sqrt{2})^{-1/3}$, and $c^* = c(2V)^{-1/3}$. They are constructed to yield values of exactly one for an ideal cubic perovskite, and deviate progressively with increasing distortion. As discussed in Chapter 3, the use of average crystal length scale in place of pressure for plots like these greatly improves them, by enabling the simultaneous consideration of a range of temperatures, as well as providing a space that is often more closely linear.

In the upper panel of Figure 2.9, we show the normalized axial ratios measured at ambient temperature for both perovskite samples over the entire pressure range. The Fe-bearing sample is shown in solid symbols and the Fe-free sample is shown in open symbols. Furthermore, these axial ratios can be converted into an estimate of the octahedral tilt angle of the corner-sharing silica octahedra within the perovskite structure. It is the progressive tilting of these octahedra with increasing pressure that accommodates the gradual distortion of the unit cell induced by compression. From *O'Keeffe et al.* (1979), we can calculate the octahedral tilt angle as a function of the unit cell parameters:

$$\phi = \cos^{-1}\left(\frac{\sqrt{2}a^2}{cb}\right) = \cos^{-1}\left(\frac{1}{(c/a)^*(b/a)^*}\right) \tag{2.13}$$

where $(c/a)^*$ and $(b/a)^*$ are the normalized axial ratios. In the lower panel of Figure 2.9, we show the linear evolution of the octahedral tilt angle with compression for the two samples. For both the normalized axial ratios and the octahedral tilt angles, there is such a strong similarity between the two samples that we find the past reported resolvable differences as a function of iron content somewhat doubtful, at least for our samples.

2.4.2 Physical Property Profiles

The most scientifically interesting questions that can be addressed using the perovskite equation of state depend on the pressure profiles of various physical properties like thermal expansion and bulk modulus. In order to evaluate these profiles, we must obtain the expres-

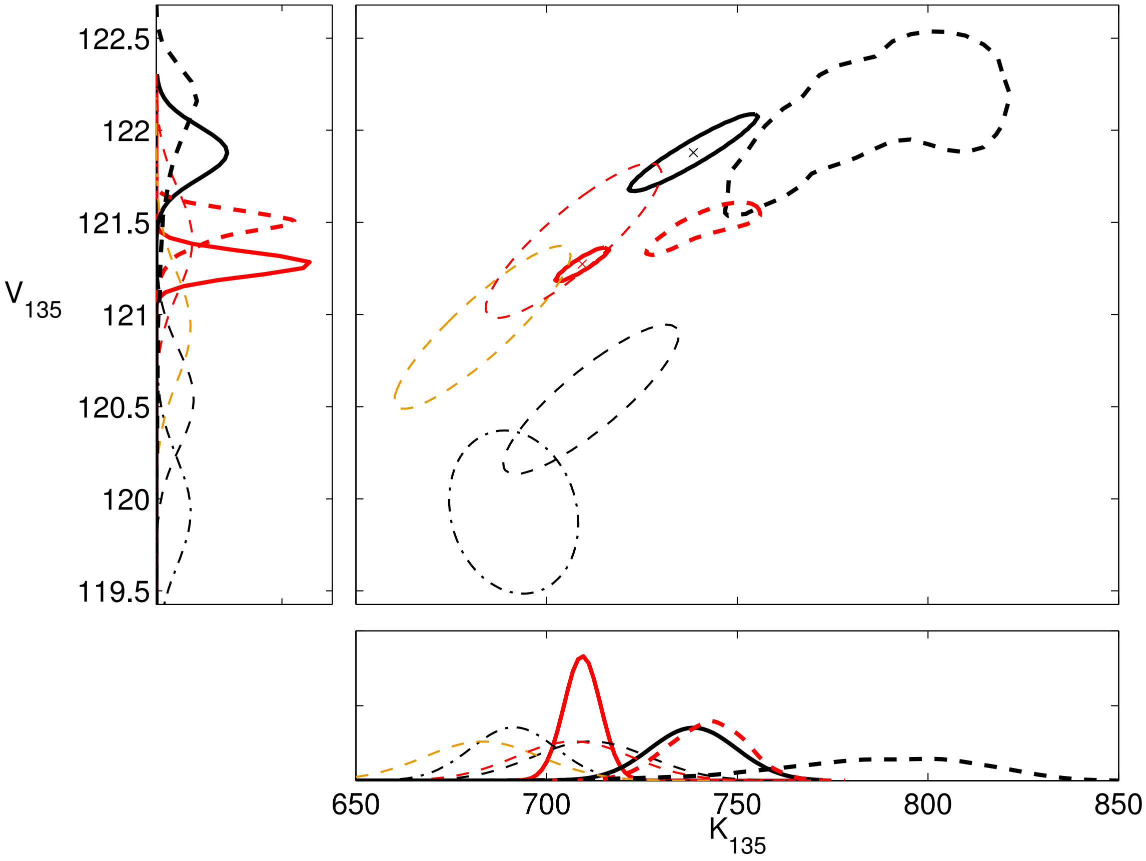

Figure 2.8: Comparison of 68% confidence bounds on local properties at core-mantle boundary pressure conditions are shown for Fe-bearing and Fe-free perovskite. All lines are color coded with Fe-free samples in black, intermediate ∼9% Fe samples in orange, and ∼13–15% Fe-rich samples in red. The thick lines show the results from this study: the solid lines display the results when metastable ambient-pressure measurements are used to constrain V_0, whereas the dashed lines represent the results assuming the metastable behavior of perovskite is a poor predictor high-pressure properties. The thin dashed lines show the three compositions studied by *Lundin et al.* (2008) and the thin dash-dotted line shows the results from *Fiquet et al.* (1998).

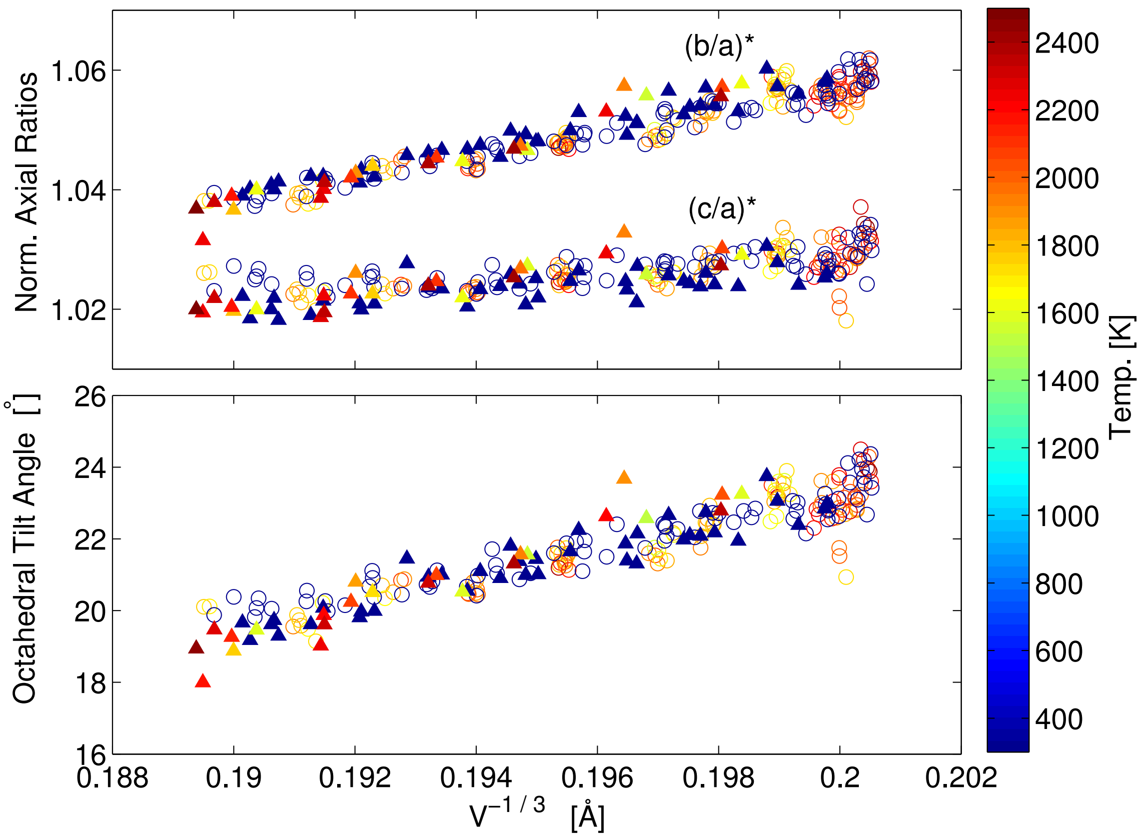

Figure 2.9: Normalized axial ratios and octahedral tilt angles shown for the 0% and 13% Fe-bearing samples. The 13% Fe-bearing sample is plotted in circles and the Fe-free sample is plotted in triangles.

sions for the relevant thermodynamic parameters. Below we derive the expressions for the isothermal bulk modulus by taking derivatives of the pressure, with both a "cold" contribution from the Vinet equation of state and a "hot" contribution from the Mie-Gruneisen-Debye thermal correction:

$$\begin{aligned} K_T(V,T) &= K_c(V) + K_{th}(V,T) - K_{th}(V,300) \\ K_c(V) &= -K_0 x^{-2}\{x[1-\nu(1-x)]-2\}\exp\left[\nu(1-x)\right] \\ K_{th}(V,T) &= \frac{C_{vm}\gamma T}{V}\left[\frac{3\gamma\Theta/T}{e^{\Theta/T}-1} - D\left(\frac{\Theta}{T}\right)(q-1+3\gamma)\right] \end{aligned} \tag{2.14}$$

where $\nu = 3/2(K_0' - 1)$ and $x = (V/V_0)^{\frac{1}{3}}$ is the average axial strain, and recalling that the Debye temperature Θ and the Grüneisen parameter γ are given by Equation (2.11). Combined with the expression for the heat capacity of a Debye solid, a thermodynamic identity can be used to obtain the thermal expansion parameter at any desired volume and temperature:

$$\begin{aligned} C_V(V,T) &= C_{vm}\left[4D\left(\frac{\Theta}{T}\right) - \frac{3\Theta/T}{e^{\Theta/T}-1}\right] \\ \alpha(V,T) &= \frac{\gamma(V)C_V(V,T)}{VK_T(V,T)} \end{aligned} \tag{2.15}$$

Using these expressions, we determine the isothermal profiles of thermal expansion, bulk modulus and volume for the two perovskite samples. In Figure 2.10, we show the propagated 68% confidence intervals for these material properties evaluated at 300K, shown in the light shaded regions, and at 3000K, shown in the dark shaded regions. From these profiles, it is clear that the largest difference between the two perovskite compositions is the sizable affect that the addition of iron has on the thermal expansion. By adding 13% Fe to perovskite, we see a roughly 50% drop in the thermal expansion at the core-mantle boundary, as seen in the top panel. Additionally, we can see that because of this drastic difference in thermal behavior, the bulk modulus and volume see a much reduced change upon increasing temperature for the iron-bearing sample as compared to the Mg-endmember.

It should be noted that we obtain a significantly different thermal behavior for the Fe-free endmember sample than *Fiquet et al.* (2000). A previous investigation by *Mosenfelder et al.* (2009), however, pointed out the none of the previous work on the thermal expansion of $\mathrm{MgSiO_3}$ perovskite is consistent to better than a factor of about two at mantle conditions, and therefore we do not find this level of disagreement particularly surprising. Furthermore, the difference in thermal expansion between the two samples is plainly visible to the eye in the raw data shown in Figure 2.7, further bolstering the robustness of this result.

2.4.3 Perovskite-Dominated Chemical Piles

With a well characterized equation of state model in hand for both Fe-free and 13% Fe-bearing perovskite, we are able to quantitatively assess the behavior of perovskite-dominated lower mantle structures. In this section, we develop a thermodynamic model of Mg-Fe perovskites at arbitrary compositions using an ideal mixing model for combining the equation of state results from our two perovskite samples. We then apply this simple perovskite model to determine the physical properties of perovskite at core-mantle boundary pressure-temperature conditions. Comparison with seismic estimates of bulk mantle physical properties allows us to compare the relative merits of different possible explanations

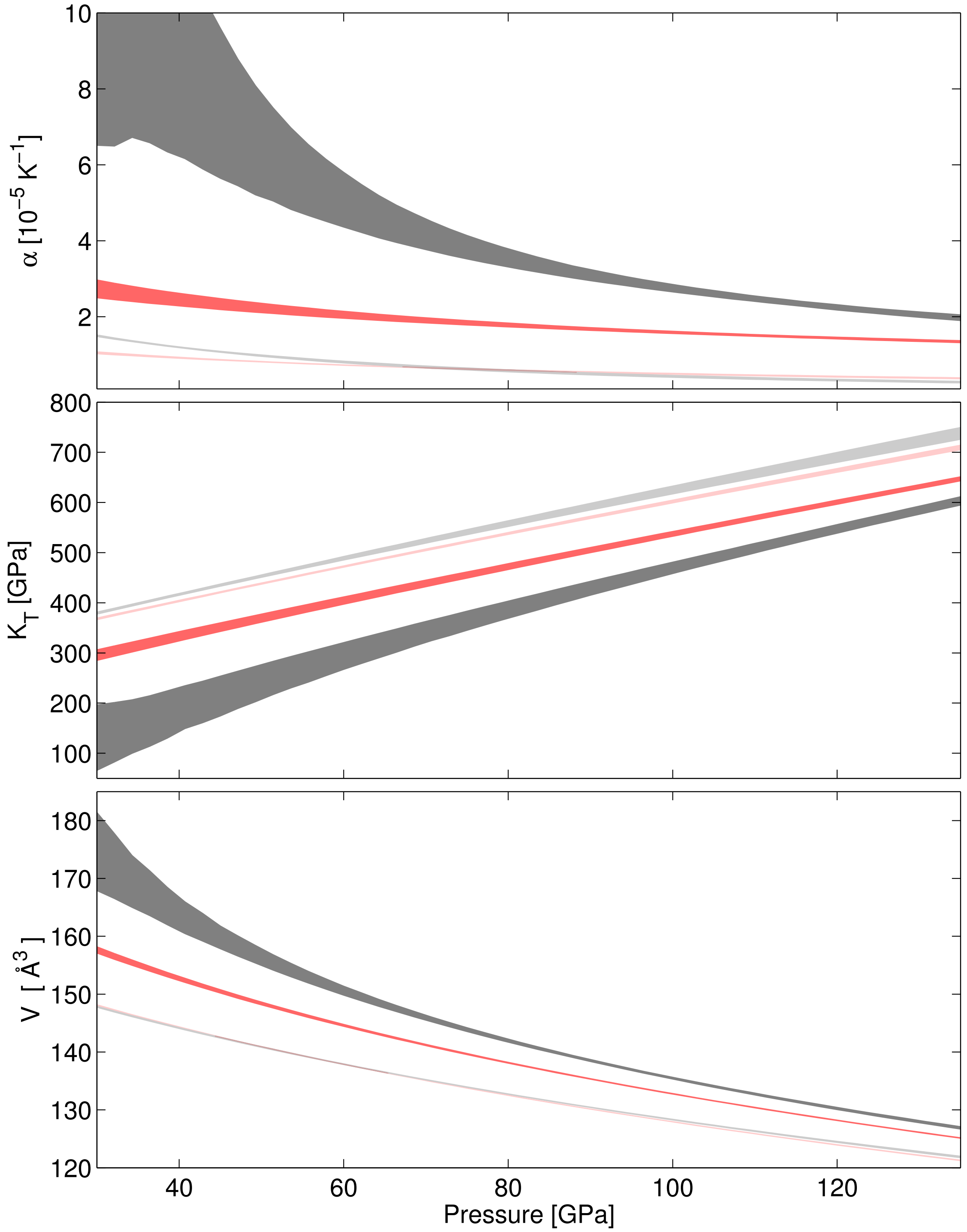

Figure 2.10: Shaded regions show 68% confidence intervals for 300K (light shading) and 3000K (dark shading) isothermal properties of Fe-free perovskite in black and 13% Fe-containing perovskite in red. The effect of additional iron is rather small at room temperatures, with nearly identical volumes and thermal expansions and only slight differences in bulk modulus. The distinctions becomes quite pronounced at elevated temperatures, however, due to the significant difference in thermal expansion. The low-pressure region for the Fe-free sample is rather poorly constrained by the data, but this appears to have little effect on the high-pressure behavior, which retains narrow uncertainty bounds.

for seismically observed core-mantle boundary structures.

Large Low Shear Velocity Provinces (LLSVPs) represent one of the most important first-order features of the lower mantle and may represent the oldest structures in the Earth. It has been proposed that the seismically distinct LLSVPs are likely to represent both chemically and thermally distinct structures on the core-mantle boundary (*Tackley*, 2011). The two LLSVPs, beneath Africa and the Pacific, present a particular challenge for explanation as they appear to have sharp and often steep-walled boundaries and stand roughly 1000 km high off the CMB. Residing at the base of a convecting system, it is a challenge to understand how such structures might remain for geologic time without mixing away through the process of entrainment. The general interpretation of the sharp seismic reflector at the top of the structure is that these features cannot be explained solely by thermal anomalies (*Tackley*, 2011), which would tend to have more diffuse edges, though there is still some disagreement in the literature on this point (*Davies et al.*, 2012). The viewpoint of a chemically distinct pile is further bolstered by an apparent anticorrelation between shear-wave velocity anomalies and both bulk sound velocity and density anomalies within the structures relative to bulk mantle, contrary to the general trends of most heated material (*Ishii and Tromp*, 1999).

Under the umbrella of chemically distinct explanations for these lower mantle structures, there are two broad endmember theories that account for their large topographic relief relative to the CMB, as discussed in *Tan and Gurnis* (2007). The domes can either be explained as passive chemically dense piles that are dynamically propped up by external convective stresses or as free-standing and internally convecting metastable piles. The passive pile explanation is the more intuitive of the two, in which the piles reside at the base of the mantle since their constituent material is denser than bulk mantle material. Fighting their tendency to spread out and pool as thin shallow layers on the CMB, some external force must be invoked to sweep them into domed piles, which is often attributed to cold dense plates descending to the CMB and pinching the sides of these structures in order to dynamically prop them up. This story of course assumes that a sufficient flux and geometry of down-going plates is available at the core-mantle boundary in order to achieve the requisite force to counteract the negative chemical buoyancy of the piles. Under the competing scenario, no external force is required, but rather the chemically distinct piles are made of a material that is less dense that the surrounding mantle at the base, but experiences a density crossover, or height of neutral buoyancy, near the top of the pile about 1000 km above the CMB. Under this explanation, the pile is undergoing slow internal convection, where light material at the base rises to the neutral buoyancy point and then cools and falls back to the CMB, thereby enabling the pile to prop itself up without the help of external stresses. To achieve this behavior, the pile material must be less dense than bulk mantle (on account of elevated temperatures that result from the thermal boundary layer separating the pile from the rest of the mantle) as well as being very stiff compared to the average mantle, enabling a density crossover point. These properties are a function of both the unknown composition and poorly constrained excess temperature of the pile material. Nevertheless, they are a direct consequence of the equation of state properties of the material, and can therefore be directly probed given a thermodynamic model.

While we know very little about the detailed composition of these distinct structures in the deep Earth, one possible model for such structures is in terms of an iron-enriched and perovskite-dominated pile. As perovskite is the primary phase in the lower mantle and given the sizes of the structures, it is also likely the primary component of the LLSVPs. With the equation of state information obtained in this study, we are perfectly positioned to investigate such a possibility. In order to probe a range of iron compositions, we construct an ideal mixing model of the 0% and 13% Fe composition perovskite samples. The ideal mixing model is based simply on the assumption that at constant environmental conditions, temperature and pressure, the behavior of a mixture of two components is given simply assuming that volumes and internal energies mix in proportion to the relative amount of each component, and also that the entropies are also additive together with an additional term that describes the ideal entropy of mixing. Under these assumptions, it is straightforward to derive the combination of thermodynamic parameters that can be simply averaged together in proportion to the quantity of each component. Using this method, we are able to take known or assumed equations of state for two different compositions and mix them under the assumptions of ideality, to obtain the approximate physical properties of a mixture of any desired composition. Since it relies on the assumption of linear additivity, there are clearly limits to this method. We therefore restrict ourselves to examining compositions for the mantle structures that are less than twice the difference between the two mixing members, and thus we consider composition up to 25 mole % Fe.

The results of this mixture model comparison are given in Figure 2.11. In this exercise, use the above described mixture model to calculate the material properties of the mixture at the core-mantle boundary. Imagining that both the iron composition and the excess tem-

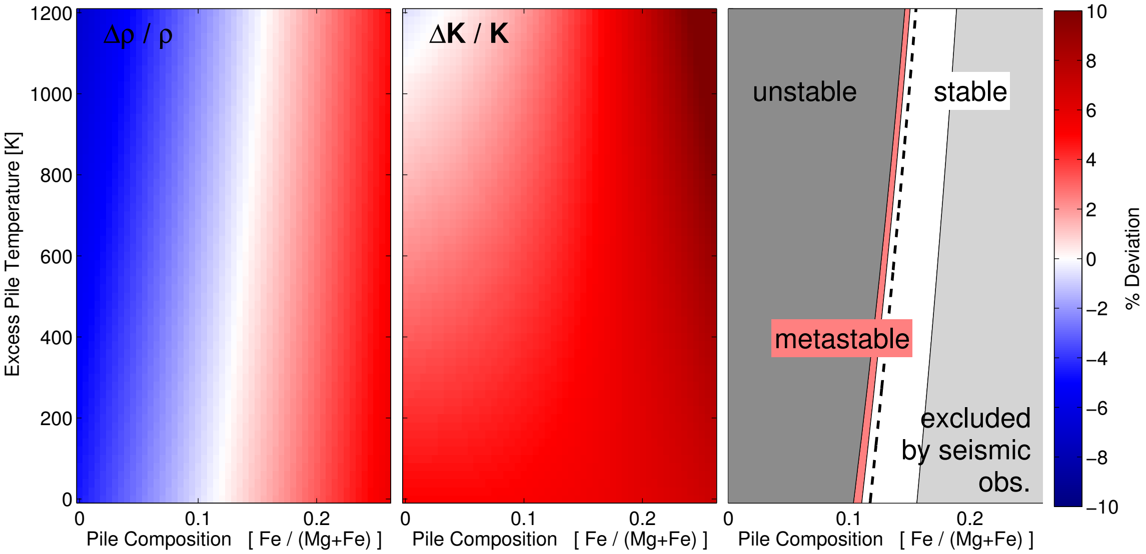

Figure 2.11: The ideal mixture model is used to sample the behavior of perovskite under a range of temperatures and compositions, corresponding to the potential conditions in a perovskite dominated LLSVP. The densities and bulk moduli are compared to the core-mantle boundary values from PREM. This yields the percent contrast maps between the perovskite chemical pile and the bulk mantle shown in the left-hand and central panels. The differences in the local properties are combined in a local linear expansion to estimate the height of the density crossover, which is transformed into the pile stability map shown in the right-hand panel. Densities greater than PREM yield stable passive piles, but excess densities greater than about 1.5% are generally excluded as they would be visible in seismic observations. A small region of space allows for metastable domes between 600 and 1200 km tall, and everything taller is regarded as initially in conflict with observations, and for low enough densities, entirely unstable at the core-mantle boundary.

perature of the LLSVP are unknown, we search a grid over those variables to determine the possible densities and bulk moduli of the mixture. In the left-hand and central panels of the figure, we show the percentage contrast between the mixture model (as a function of composition and temperature) and the bulk mantle as represented by PREM (*Dziewonski and Anderson*, 1981). Using a simple linear expansion of the density and bulk modulus difference at the core-mantle boundary, we can obtain rough estimates for the height of the density crossover. Comparing those predicted heights to a generous range of values corresponding to reasonable estimations for the observed heights of the actual LLSVPs, lying somewhere between 600 and 1200 km above the CMB, we can map out the parameter space by noting the relative stability of the structure as a function of composition and temperature. This result is shown in the right-hand panel of Figure 2.11. In the figure, the dark dashed line represents the equal density line, where to the right of that line, the chemical pile material is denser than ambient mantle and thus stable. At very large iron compositions, however, the excess density becomes so high that it should be plainly visible in seismic observations, especially normal mode inversions like those done by *Ishii and Tromp* (1999). We therefore restrict the allowable compositions to those with excess density contrasts of less than 1.5%, which is a somewhat conservative estimate. Immediately to the left of the equal density line is material that is slightly buoyant and thus the red shaded sliver represents material with a height of neutral buoyancy lying within the reasonable limits of 600 and 1200 km above the CMB, representing solutions consistent with a metastable dome explanation for the LLSVPs. For lower iron compositions, the pile material becomes even less dense until it is entirely unstable escaping from the base of the mantle. As is obvious from the figure, there is only a tiny sliver of allowable phase space that corresponds to the delicate balance required by the metastable dome hypothesis. In contrast, passive piles are extremely insensitive to composition and temperature. We can thus assess the relative plausibility of the passive pile and metastable dome explanations by calculating the fraction of allowable parameter space occupied by material behaviors consistent with the two theories—this is given simply by the relative area of the red shaded region as compared to the white region. This fraction is just below 10%, indicating that while metastable domes are possibly consistent with our current understanding of the perovskite equation of state, they imply very constraints on the temperature-dependent composition of the LLSVP material and are thus not very likely.

2.5 Conclusion

A series of high-pressure and -temperature laser-heated diamond anvil cell experiments were performed in a nearly hydrostatic neon pressure medium to determine the equations of state of 0% and 13% Fe-bearing Mg-perovskite. Using neon as the primary pressure indicator, we develop a new statistical technique for estimating deviatoric stresses in the diamond anvil cell. These stresses are confined throughout the experiments to values of about half a percent of the pressure. Using a series of Bayesian data analysis techniques that are detailed in Chapter 3, we obtain estimates of the unit cell volumes which are then fit with thermal equations of state.

The equation of state parameters found for the Fe-bearing and Fe-free samples are generally consistent with previous measurements perovskite measurements, the Fe-free sample exhibits slightly elevated volumes at high pressure as compared to previous values. The derived equation of state parameters for the Fe-bearing sample are significantly more precise than any previous perovskite measurements, providing tight constraints on both the ambient-temperature and high-temperature properties. The most dramatic results are the large thermal expansions measured for the Mg-endmember sample, corresponding to the high values determined for the thermal equation of state properties.

Finally, we incorporate the pair of equation of state determinations into an ideal mixture model enabling the calculation of approximate equation of state properties for any arbitrary iron composition. Using this mixture model, we examine the range of plausible values in temperature-composition space to explain the presence of LLSVPs. Using plausibility arguments, we show that metastable perovskite domes are possible, given our knowledge of the equation of state, but represent a very precise balance between the iron content and temperature of the chemical piles on the CMB, and are not therefore particularly likely. We instead find the passive pile explanation more compelling, as it does not require that the Earth occupy a special narrow region of parameter space merely by chance.

— Appendices —

2.A Data Reduction Pipeline

Our data pipeline, written in MATLAB, transforms raw powder diffraction images into properly normalized one-dimensional spectra. This data processing procedure involves first obtaining the observation geometry from a diffraction image of a known calibration standard. If necessary, the data image is corrected for saturation effects due to CCD pixel blooming. Intensity weighting factors, determined by the geometry, are applied to the data transforming them to idealized detector measurements. These transformed data are integrated producing a 1-D spectrum with error bars resulting from counting statistics. Finally, the background signal is estimated and subtracted from each spectrum, preparing it for convenient spectrum comparison or fitting.

We determine the detector geometry using a method similar to that reported in (*Hinrichsen*, 2006). After determining an approximate beam-center location, the diffraction data for a known calibration standard–such as $\mathrm{CeO_2}$ or $\mathrm{LaB_6}$–is sliced radially using bi-cubic interpolation to obtain a set of 1-D radial spectra. Peak fitting with pseudo-Voigt profiles is then used to extract the 2θ locations of every line in each radial slice. These calibration line positions (with associated uncertainties) are then fit by varying the experimental geometry parameters controlling the orientation and position of the detector: detector distance, beam-center location, and detector tilt and rotation. The geometric calibration parameters are chosen as the values with the maximum likelihood given the measured line positions using the standard least-squares method (as is appropriate to simple data fitting with Gaussian uncertainties and in the absence of strong prior information).

For diffraction images containing both strong and weak x-ray scatterers, it is often impossible to obtain high quality spectra that do not suffer from over-exposure in certain regions of the image. Due to the basic properties of CCD detectors, exposure of a pixel beyond its full-well depth causes 'blooming' in the final image, where electrons spill over into neighboring pixel wells causing full-intensity streaks to emanate from the excessively bright points in the image The typical approach to this problem is to adjust exposure times to limit its occurrence. Unfortunately, in many cases this method is either ineffective (e.g., in the presence of very weak scatterers) or impractical (e.g., during high-temperature measurements). It would be highly advantageous to be able to use these data while minimizing the impact of the erroneous intensities caused by blooming–such an approach is made possible by the conservation of electrons within the affected region. Since the total number of electrons, equal to the number of photons registered by the CCD, remains constant as electrons spill into neighboring pixel bins, simple summation in an overexposed region will give an approximate total intensity for that region. In order to determine how to reasonably distribute this total amongst the affected pixels, we leverage the angular symmetry inherent to powder diffraction by setting the relative intensity of each pixel to the values from integrated 1-D spectrum. Iterating this procedure then results in reasonable pixel intensities which will no longer induce wild bias into the final integrated spectrum.

The final data reductions step is the integration procedure itself, which operates on the powder diffraction images (pre-corrected for saturation if necessary) using the geometric parameters derived from the calibration image. Using the equations presented in *Hinrichsen et al.* (2008), we determine the proper intensity weighting factors required to transform each measured pixel value into an equivalent ideal pixel intensity for a perpendicular and hemispherical detector. This factor is combined with the standard 2-D Lorentz and polarization correction factors (see, e.g., *Hinrichsen et al.* (2008)), forming a single overall weighting factor for each pixel on the detector. Since this intensity weighting map is independent of the data collected, it need only be calculated once for each detector configuration. To obtain a reasonably smooth one-dimensional spectrum, as remarked by *Hammersley et al.* (1996), measured pixel intensities are divided amongst sub-pixel regions according to standard bi-cubic interpolation. These sub-pixels are then sorted into bins according to their diffraction angle 2θ. The sub-pixel intensities are combined together as weighted observations of a Poisson process, using the intensity weight map described above, resulting in a one-dimensional spectrum of intensity as a function of diffraction angle. As a final post-processing step, the method derived by *David and Sivia* (2001) is used to estimate and subtract a robust Chebyshev polynomial background from the spectrum. This acts as a good initial guess of the background, which can later be refined, and provides a simple flat spectrum ready for analysis and visualization.

Chapter 3

A Bayesian Approach to Determining Equations of State in the Diamond Anvil Cell

Abstract

In this work, we apply the general Bayesian statistical approach to analyzing powder diffraction data from diamond anvil cell experiments. Statistical problems over a wide range of complexities arise in this effort, and we show how the Bayesian statistical framework provides all of the tools necessary to adeptly meet these challenges. In particular, we demonstrate how this method for data analysis naturally leads to probabilistically rigorous solutions that can range from standard least-squares methods, which are fully consistent with non-Bayesian methods, to uniquely Bayesian solutions to problems that require a more sophisticated application of Bayesian priors. This proves especially useful in cases where the data do not fully constrain important physical parameters, but where outside knowledge provides the necessary information to yield reasonably well-constrained solutions. These include situations such as the use of previous studies of the composition-dependence of zero-pressure crystal volumes as well as the usage of smoothness criteria for the com-

pression evolution of axial ratios. In this chapter, we therefore demonstrate through many examples the flexibility and power of Bayesian data analysis as applied to experimental measurements in mineral physics.

3.1 Introduction

Fitting models to data is one of the most common and routines tasks in scientific investigations. It is nevertheless extremely common for model-fitting to be carried out in ad-hoc manner. In this chapter, we present the Bayesian approach to data analysis and specifically apply it to the interpretation of x-ray diffraction data. We hope to show that rather than giving a compendium of specific recipes for analyzing numerical data, Bayesian statistics instead provides a general framework for data inversion problems which transform information, in the form of noisy data, into useful knowledge. Once absorbed, this enables practitioners to develop data processing methods that are grounded in the rules of probability.

3.2 Intro to Bayesian Statistics

The Bayesian approach to statistics essentially boils down to the idea that every data analysis problem can be cast in terms of evaluating the probability of each data point given some model of reality. Any description of Bayesian statistics must start with Bayes' theorem, which is the basic foundational principle that underlies everything else:

$$
\begin{gathered}
prob(\vec{M}|\{d\}) \propto prob(\{d\}|\vec{M})prob(\vec{M}) \\
posterior \propto likelihood \times prior
\end{gathered}
\tag{3.1}
$$

which can be read as: The probability of the model values $\vec{M}$ given the data, called the posterior, is proportional to the probability of the data $\{d\}$ given the model values, called the likelihood, times the initial probability of the model values to begin with, called the prior.[1]

The derivation of Bayes' theorem is deceptively simple, relying only on the rule of conditional probability $prob(A, B) = prob(A|B)prob(B)$, which states that the probability of two things both being true, $prob(A, B)$, is equal to the probability of one alone being true $prob(B)$ times the conditional probability of the other given that the first is true $prob(A|B)$. But the choice of which variable to condition on is arbitrary, so we can write an equivalent expression for $prob(B, A)$. Simple algebraic manipulation then leads directly to Bayes' theorem. The real power of Bayes' theorem is not the expression itself, which is an irrefutable expression of basic probability, but the willingness to interpret A and B as the model and data sets vectors. Once this is accepted, however it leads to countless insights—it is not unreasonable to think of it as the equivalent of Newton's second law for probability studies, where at some level it represents an assertion that if accepted has almost boundless explanatory power.

As we will refer regularly to each of these terms throughout this chapter, and they can

---

[1] The missing constant of proportionality is often called the evidence, but it is unimportant for parameter estimation problems and therefore we will not discuss it here.

often be confused, we will attempt to make their meanings plain. The posterior is the final answer of any statistical data analysis problem and is often quoted as a best-fit along with associated errors on the best-fit parameters. While perfectly correct, this statement of the results can be somewhat misleading. In actuality, the end result of parameter estimation is not a single answer, but rather a *probability distribution*, describing the relative probability of different combinations of parameter values after analyzing the data. In this way, the posterior summarizes our final belief about the parameter values after performing our analysis. The practice of quoting best-fit values with associated uncertainties is a concise way of conveying this information for posteriors which are approximately described by normal (Gaussian) distributions. We will discuss this point further in Section 3.3.3 when we obtain the posterior covariances for the powder diffraction analysis.

The likelihood is the most familiar of the three terms in Bayes' theorem, conveying the relative "goodness-of-fit" of the data to a particular set of model parameters. As we will show later, the most familiar form of the likelihood uses weighted least-squares or "chi-square" minimization, which seeks to minimize the difference between the data and a model. Since the prerequisite of any data analysis problem presupposes that we can generate a model for any particular set of model parameters $\vec{M}$, the likelihood then uses probability distributions for the data points $\{d\}$ to determine the probability of the data given the model values. The likelihood can therefore be thought of as asking the data, "What is the probability of this datum?"[2], and then combining the probability of every data point into an overall probability factor.

---

[2] These words are taken directly from John Johnson. They succinctly convey the root question that the likelihood addresses, and so we quote them here.

Lastly, the prior is simply the term that encapsulates our knowledge about the system before performing the analysis. This term is unfortunately the one that people tend to have the most difficulty with accepting, but it is actually rather straightforward. The prior simply mathematically encodes our “belief” about the model parameter values *before* the data is analyzed, just as the posterior conveys our belief *after* the data is analyzed. The prior therefore explicitly acknowledges that we usually have some previous or outside knowledge about a problem that is influencing our analysis. The most straightforward prior is just the posterior from a previous investigation, though we must be careful that we are comparing two directly analogous studies and be sure that the previous investigator properly reported their final uncertainties. Oftentimes, when lacking a directly applicable previous investigation, we might choose to use weakly informed priors, typically in the form of very wide normal distributions for each parameter. In some cases, we may choose to disregard all outside knowledge and therefore favor totally “uninformed” priors, which are wide and entirely flat. This option seems attractive to many at first, as they deem it the most “objective”; it is easy to recognize that this is the result of flawed thinking, however, since the choice of model parameterization is largely arbitrary and a flat prior for one parameterization can be a very biased prior in another. Assuming a total lack of prior knowledge—which is almost never the case—there is no “objective” method available to choose which parameterization is best. Therefore, it is always important to make informed choices about the combination of parameterization and prior, since together they directly affect the posterior. That said, we needn’t worry too much about such issues given enough good quality data, since the likelihood will always win out over the prior because its affect grows as more data is added while the prior remains fixed.

Bayes' theorem is the very heart of the inversion method, allowing us to transform data into knowledge about model parameter values. This process is not at all trivial, since the causal arrow points in the opposite direction: we start with some underlying physical process, as represented by a model, and through that process, the data is generated. In the Bayesian framework, we are required at this point to somewhat broaden our definition of "model", to include both the physical picture that dictates the "ideal" noise-free values of the data coupled with a representation of how noise or observational error are introduced to the data. Therefore, a generic expression for the observed data values is given by:

$$\{d\}^{\text{obs}} = \{y\}^{\text{phys-mod}} + \{\epsilon\}^{\text{noise-mod}} \tag{3.2}$$

where the observed data values, $\{d\}^{\text{obs}}$, result from the combination of both a physical model, generating values of $\{y\}^{\text{phys-mod}}$, and outside noise source, represented by $\{\epsilon\}^{\text{noise-mod}}$. In this way, we can see that the overall "data model" necessarily includes all terms that contribute to the measured values.[3] In fact, it is often the case that the noise model term is taken as a collection of all of the unmodeled physics in a particular analysis problem. Viewed in this way, we are given the freedom to adjust our representation of both the physical model and the error model in order to best represent the data in hand, together parameterized by the model vector $\vec{M}$.

As mentioned above, the most common method of data analysis uses "chi-square" minimization, which it is easy to show is a direct outcome of Bayes' theorem for straightforward

---

[3] It is usually, but not always the case, that the unpolluted signal and the noise are additive. In cases when they are not, the form of Equation 3.2 is simply adjusted to reflect their true relationship.

analysis problems. Also known as least-squares regression, chi-square minimization relies on the use of a normally distributed error model:

$$\begin{gathered} d_i^{\text{obs}} = y^{\text{mod}}(x_i, \vec{M}) + \epsilon_i^{\text{noise}} \\ \text{with} \qquad \epsilon_i^{\text{noise}} \sim \mathcal{N}(0, \sigma_i^2) \end{gathered} \tag{3.3}$$

where the physical model y^{mod} is a function of both an independent variable x_i and the model parameters $\vec{M}$, and the observational noise is normally distributed about the physical model value, $\sim \mathcal{N}(0, \sigma_i^2)$, with a standard deviation corresponding to the known observational errors σ_i. We can evaluate the total likelihood, $\mathcal{L}$, as simply the product of a set of normal distributions:

$$\mathcal{L} = \prod_i \mathcal{N}(d_i - y_i^{\text{mod}}, \sigma_i^2) \tag{3.4}$$

where the data points are assumed to be independent of one another, giving a total probability that is just the product of the individual probabilities for each data point. Taking the log of both sides and substituting in the expression for a normal probability distribution, we can immediately see the relation to least-squares minimization:

$$\begin{aligned} \log \mathcal{L} &= \log \left\{ \prod_i \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left[-0.5 \left(\frac{d_i - y_i^{\text{mod}}}{\sigma_i} \right)^2 \right] \right\} \\ &= -\frac{1}{2} \sum_i \left(\frac{d_i - y_i^{\text{mod}}}{\sigma_i} \right)^2 - \sum_i \log \sqrt{2\pi}\sigma_i \\ &= -\frac{1}{2}\chi^2 + \text{const}, \qquad \text{with} \quad \chi^2 \equiv \sum_i \left(\frac{d_i - y_i^{\text{mod}}}{\sigma_i} \right)^2 \end{aligned} \tag{3.5}$$

where we introduce here the concept of χ^2, which represents error-weighted sum-square residuals. From this equation, it is clear that minimizing χ^2 yields the maximum for the value of the likelihood, and is hence referred to as the Maximum Likelihood Estimate (MLE). If we choose to adopt wide and flat priors, then the prior is just a multiplicative constant for the posterior that combines with the second term in the log-likelihood. According to Bayes' theorem, the posterior is proportional to the product of likelihood and prior, and therefore constants have no effect on the final answer and can thus be dropped. Therefore, we can clearly see that the use of standard weighted least squares fitting is exactly equivalent to the case of fitting data with normally distributed errors and wide and flat priors. Additionally, we can easily include the effect of normally distributed priors under the same framework by merely adding on an additional "prior penalty" term for each model parameter:

$$\chi^2_{\text{tot}} = \chi^2 + \sum_j \left(\frac{M_j - M_j^{\text{prior}}}{\delta M_j^{\text{prior}}} \right)^2 \tag{3.6}$$

where M_j is the j^{th} model parameter value, and M_j^{prior} and $\delta M_j^{\text{prior}}$ give the mean and standard deviation of the prior for that parameter. We can thus easily see that this "workhorse" of frequentist statistics, which describes the more common standard statistical viewpoint, is in fact just a simple application of Bayesian statistics given a few common simplifying assumptions. By remaining in the Bayesian mindset, however, we have the flexibility to adjust this generic data model as needed to tailor the analysis to the specific properties of the data we wish to analyze.

Of course, we cannot stop at obtaining a best fit, since the posterior is actually a probability distribution, summarizing our final knowledge of the model parameters. We therefore must obtain the shape of the posterior to go along with the location in parameter space expressed by the best fit. We will find here, just as before, that the standard method is nothing more than the application of Bayes' theorem under the assumption of Normality. In this case, we must also assume that not just the data errors, but the posterior itself adopts a normal distribution. This is often a good assumption, especially in the presence of large quantities of good data, and depends on how non-linear the physical model is with respect to the model parameters. We can imagine constructing a local first-order Taylor expansion of the physical model in the region of parameter space near the best fit. As long as the physical model behaves approximately linearly over the highly probable region of parameter space, closely matching the Taylor expansion, the posterior will then be well described by a normal distribution *Sivia and Skilling* (2006)[4].

This method of estimating a normal posterior distribution is often called Optimal Estimation, since it uses function evaluations only in the local region around the best fit to approximate the posterior (*Sivia and Skilling*, 2006). In general, a normal posterior is described by a multivariate normal distribution (in multidimensional parameter space):

$$\log \mathcal{P}(\vec{M}) = -\frac{1}{2}\chi^2 + \text{const} \approx -\frac{1}{2}(\vec{M} - \vec{\mu}_M)^T {\mathbf{\Sigma}_M}^{-1}(\vec{M} - \vec{\mu}_M) + \text{const} \tag{3.7}$$

where $\log \mathcal{P}$ is the log-posterior, and $\vec{\mu}_M$ and $\mathbf{\Sigma}_M$ are the mean and covariance of the multivariate normal distribution, corresponding to the best-fit and shape of the posterior distribution for the model parameters M. As before, additive constants are unimportant for the log-posterior, since they merely represent the normalization constant. Though we

---

[4] It is important to recognize that such a nearly normal posterior only behaves normally over the approximately linear region of parameter space, and thus we will be unable to accurately capture the "tails" of the distribution. This is rarely an issue for parameter estimation, since we are usually interested in where the bulk of the probability lies.

must use matrix notation, since the posterior exists in multidimensional parameter space, this is merely the expression of a quadratic function in many dimensions, analogous to $\log \mathcal{P} \approx \text{const} - \frac{1}{2}(M - \mu_M)^2/\sigma^2$ in one dimension. Given the derivative properties of a quadratic, we can relate obtain the covariance matrix Σ_M by evaluating the curvature of log-posterior space:

$$\mathbf{\Sigma}_{M,ij}^{-1} \approx -\frac{\partial^2 \log \mathcal{P}}{\partial M_i \partial M_j} \tag{3.8}$$

where $\mathbf{\Sigma}_M^{-1}$ is the inverse of the covariance matrix. We will discuss how to interpret the covariance matrix in greater detail later, but for now it should be noted that it can be represented by an "confidence-ellipsoid", which contains the most probable parameter values and has a particular size and orientation. The orientation relates the correlation that exists between different pairs of parameters, indicating how well we can independently constrain their values, and the width scales the size of the region of high confidence. This is again an example of Bayesian methods directly corresponding to standard frequentist results under certain simplifying assumptions. Throughout this chapter, we will assume roughly normal posterior distributions, as the nonlinearities in our models remain rather small over the regions of high confidence. We thus do not make use of the many useful, yet complex and time-consuming, Bayesian methods that can handle non-normal posteriors.

3.3 Bayesian Analysis of Powder Diffraction Data

In applying the general Bayesian approach to analyzing powder diffraction data from diamond anvil cell experiments, a wide range of statistical problems arise of varying complexity. In this and following sections, we show how the flexible Bayesian framework is particularly well suited to developing individualized analysis routines that are well suited to the task. We also demonstrate how while always remaining within the Bayesian framework, we arrive at solutions that range from standard methods, directly equivalent to the more familiar frequentist approaches, to entirely Bayesian approaches that lie outside the bounds of the standard frequentist statistical toolbox.

The primary goal of many high-pressure powder diffraction studies is to obtain the equation of state of the material of interest. This thermodynamic function describes how the volume of the crystal unit cell varies as a function of some or all of the thermodynamic state variables: pressure, temperature, and composition. Furthermore, for non-cubic crystals, it is also important to obtain the variation of the unit cell dimensions with changing environmental conditions. In order to accomplish this task, the investigator must carry out a large, many-step inversion process, that converts powder diffraction spectra into estimates of the thermodynamic properties. In this study, we develop and present a number of general data analysis techniques for high-pressure (and high-temperature) diffraction experiments. These methods are demonstrated on powder diffraction data for pure endmember and 13% Fe-bearing Mg-silicate perovskite, obtained in the study presented in Chapter 2.

3.3.1 Estimating Peak Positions and Uncertainties

We begin with the task of obtaining unit cell parameters from a set of integrated 1D diffraction spectra. These integrated spectra were obtained using the technique described in Chapter 2. As discussed in the previous chapter, due to its greater insensitivity to spectral irregularities, which often arise in high-pressure multiphase experiments, we favor the use of individual peak identification and fitting over the whole pattern refinement method. In the process of peak fitting, the spectrum is split up into different sections that each contain sample diffraction lines. These lines are then fit individually, or in small clustered groups where the peaks are seen to overlap one another. For the peak-fitting step, we employ a standard least-squares fitting approach to obtain the positions of each peak from the integrated 1D diffraction profile (e.g., *Press et al.*, 2007). As discussed in the previous section, this least-squares method is entirely consistent with the Bayesian framework under the assumptions of normally distributed errors.

Each spectral peak is assumed to be well described by a pseudo-Voigt line profile, which is generally considered the standard for most powder diffraction applications (e.g., *Toby*, 2001).

$$\begin{aligned} y_i^{\text{mod}} = & \sum_j^{N_{\text{line}}} A_j \left[(1-s_j) \exp\left(-\ln 2 \left(\frac{x_i - p_j}{w_j} \right)^2 \right) + \frac{s_j}{1 + \left(\frac{x_i - p_j}{w_j} \right)^2} \right] \\ & + f_{\text{bkg}}(x_i, \vec{c}_{\text{bkg}}) \end{aligned} \tag{3.9}$$

where the modeled intensity, y_i^{mod}, is a function of the observed inverse d-spacing values x_i [$\mathring{A}^{-1}$]. This equation represents the total intensity as a sum of contributions from a set of N_{line} profiles together with an added polynomial background function, f_{bkg}, described by a vector of roughly 3 to 5 background coefficients $\vec{c}_{\text{bkg}}$. The attributes of each line profile are described by its position p_j, half-width w_j, amplitude A_j, and shape s_j (which smoothly transforms the profile between a perfect Gaussian line shape at $s_j = 0$ to a perfect Voigt line shape at $s_j = 1$).

We start the fitting process with an initial guess for the peak positions, based on the positions of a few of the most intense and easily identified lines (prior knowledge from previous investigations is also useful to obtain a reasonable initial guess). After obtaining an initial guess for the peak position model, the peaks are fit with an iterative procedure, where the identification of each additional line helps to better constrain the model, thereby increasing the ability to identify further lines. This is particularly important for data with many phases visible within each spectrum as well as data that contains a low symmetry phase like perovskite, which has a significant number of partially overlapping peaks. Once the peaks are generally identified, they are fit using the pseudo-Voigt profile model given in Equation 3.9 assuming that the error bars are given by simple photon counting statistics, which is approximately normally distributed for large photon counts. The spectrum fitting is carried using a combined user-driven and automated minimization routine written in MATLAB, finding the set of parameters that maximizes the posterior probability by minimizing chi-square. The resulting individual best-fit peak positions for the Fe-free and Fe-bearing datasets are shown in Figures 3.1 and 3.2.

Once each region has been well modeled to extract the best-fit line positions, the next challenge is to obtain errors for these line measurements, as it is those line position uncertainties that propagate into the errors on the extracted crystal volumes and unit cell dimensions (and eventually into the equation of state parameters). As there are potentially many unmodeled physical effects present in the spectra, it is not practical to use optimal estimation to obtain errors in the peak positions, and thus we seek an alternate way to assess

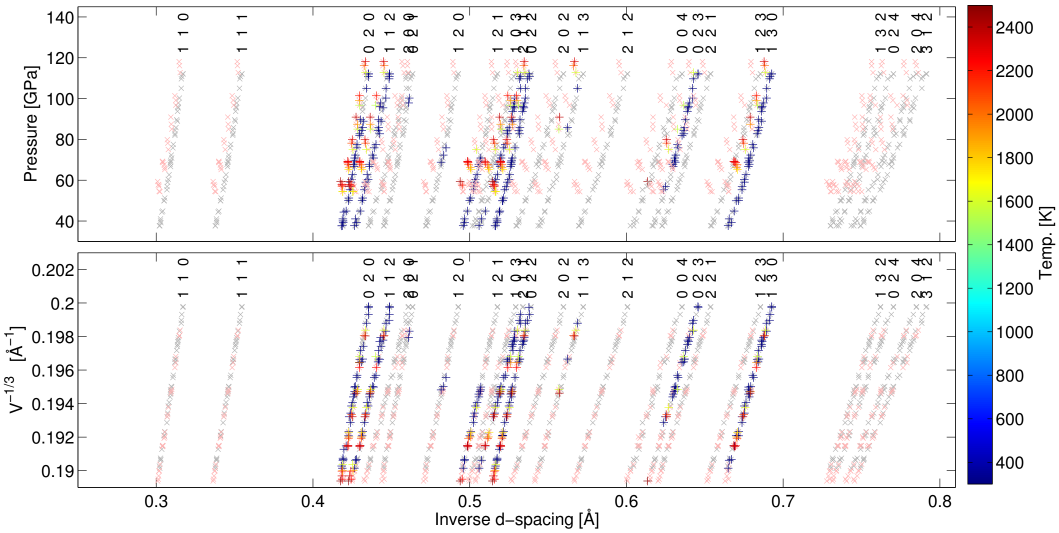

Figure 3.1: Peak positions for 0% Fe perovskite sample are plotted against pressure in the top panel and against $V^{-1/3}$ in the bottom panel. Plotting against pressure is difficult to interpret since it includes effects from the equation of state; in contrast, the purely geometric space leads to linear trends with minimal scatter and no systematic offsets due to temperature. The hkl value of each line is labeled at the top and in pale gray and pale red, the best-fit model peak positions are shown for the cold and heated data points.

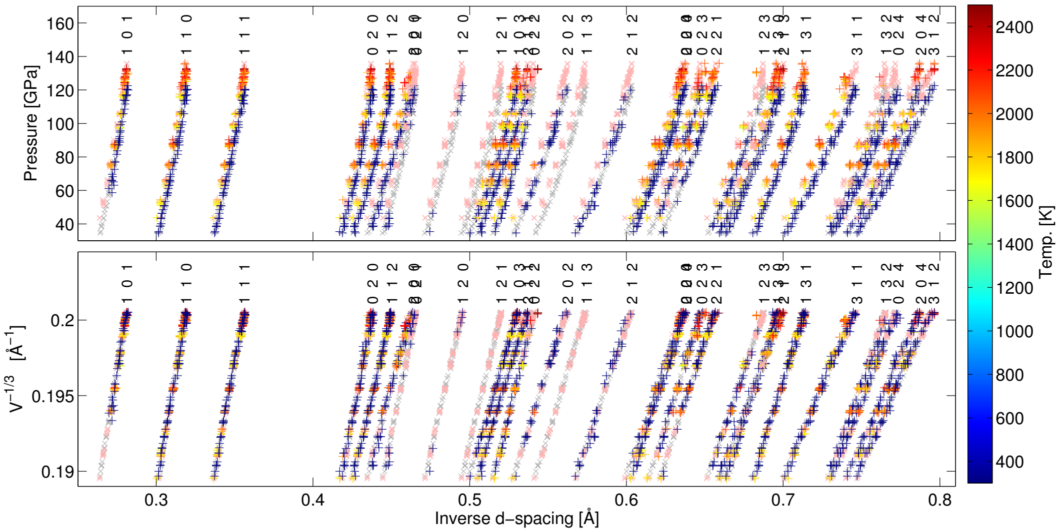

Figure 3.2: Peak positions for 13% Fe perovskite sample are plotted against pressure in the top panel and against $V^{-1/3}$ in the bottom panel. Plotting against pressure is difficult to interpret since it includes effects from the equation of state; in contrast, the purely geometric space leads to linear trends with minimal scatter and no systematic offsets due to temperature. The hkl value of each line is labeled at the top and in pale gray and pale red, the best-fit model peak positions are shown for the cold and heated data points.

reasonable uncertainties. Thinking back to the peak parameters, we note that the primary variable controlling the ability to determine the location of a peak is its width assuming the amplitude is sufficient to make it clearly visible. This reflects the fact that the peak width directly controls the relative curvature of the spectrum in the region of the peak. We therefore suppose that the line position error can be modeled as roughly proportional to the peak width, where the constant of proportionality is specific to each line. If we can identify this constant of proportionality by examining the observed scatter in the measured line positions, we will then have a straightforward way or assigning uncertainties to each position measurement.

The most obvious space in which to characterize the scatter is simply pressure vs inverse d-spacing, since the peak positions should follow a systematic trend with increasing pressure. This choice is not ideal, however, since the peak positions are sensitive to both pressure and temperature, where thermal pressure causes the heated data points to “lift” up off of the 300K curves, making it impossible to define a single curve that the data points are scattering about. This effect is plainly visible in the upper panels of Figures 3.1 and 3.2. In addition, by plotting against pressure, whose value is uncertain, we inadvertently add an additional source of scatter. In searching for a better alternative, we note that the evolution of the relative axial ratios is primarily a geometric phenomenon, and is thus only a function of volume. The planar spacings within the crystal, which are determined by the crystal geometry, should also therefore be nearly independent of temperature. Furthermore, since the cube root of the volume roughly sets the overall length-scale for the crystal, inverse d-spacings for each line should be roughly linear in $V^{-1/3}$. The lower panels of Figures 3.2

and 3.1 show the evolution of the line positions with changing inverse crystal length-scale, where the symbols are color coded by temperature. The inverse length scale can be thought of as a proxy for pressure, where the line positions are found to be conveniently nearly linear in this variable and independent of temperature.

We can therefore use the observed scatter relative to a best-fit linear trend as a way to obtain the line position uncertainty constants. For simple normally distributed data, it is well known that the sample standard deviation is a good estimate of the width of the normal distribution—it is in fact the Maximum Likelihood Estimator. This familiar result comes from the fact that the expectation value of the error-weighted square residuals is simply equal to one for a normal distribution, $\langle(\Delta x_i/\sigma_i)^2\rangle = 1$. Thus we can trivially determine the approximate error on each measured position, σ_i, by substituting in our assumption that the error is proportional to the line width and solving for the proportionality constant yielding:

$$\sigma_i = \alpha w_i, \quad \text{where} \quad \alpha \approx \sqrt{\langle(\Delta x_i/w_i)^2\rangle} \tag{3.10}$$

Using this method, we can easily obtain accurate estimations of the uncertainties of each peak position measurement, completing the first stage of the statistical analysis.

3.3.2 Estimating Unit Cell Parameters and Accounting for Misidentified Lines

After each spectrum is reduced into a list of measured peak positions and uncertainties, the next step in the analysis is to fit these peak lists with a crystal reflection model. The details of this method are given in the previous chapter. Here it is only important to note that for orthogonal perovskite, we have a model that is a function of three parameters, the lengths of the three crystal axes a, b, and c. These three parameters must then be fit to the 10 to 25 measured peak positions.

If every line in a powder diffraction spectrum is correctly identified, we could simply apply the standard least-squares approach given by Equation 3.5. Unfortunately, this is extremely unlikely when fitting lower symmetry phases like perovskite—which have many potentially visible and overlapping reflections—or when many different phases are present in the diamond anvil cell resulting in overlapping peaks—which frustrate accurate peak identification and fitting. Since the primary criterion used for peak identification is that the observed line appears at a position reasonably close to where it is expected, we must concede that misidentified peaks are bound to exist within our peak position list. To decrease the incidence of misidentified peaks, we should first examine the linear peak position maps in Figures 3.1 and 3.2. In addition to plotting the measured peak positions in dark colors, we also show the best-fit model positions for each line in pale gray and red for the cold and hot spectra. By showing the measured and modeled positions together in this well-behaved linear space, it is often easy to pick out strong outliers that are otherwise difficult to identify. We can then go back to the original spectrum to see if there was a simple and obvious mix-up between neighboring peaks, which are readily rectified. In other cases, it can be plainly seen that not just one, but all of the lines identified with a particular peak deviate systematically from the modeled position, indicating that the observed peak must belong to a different mineral phase. Worst of all identification errors are the lines that coincidentally lie near an expected peak position, but never deviate sufficiently far to draw attention to the mistake. These lines will therefore remain within the fitted data set, quietly and systematically biasing the measurements. In order to address this issue, we use a simple Bayesian mixture model approach which is robust against moderate degrees of contamination by peak identification errors.

A Bayesian Mixture Model is a general statistical tool that is useful in analyzing "polluted" datasets, where there are a number of different data models from which each data point can be generated. We start with the general data model description given by Equation 3.3. In writing that simple expression, we implicitly assumed that every data point was generated by the same physical process. When our dataset contains data from a variety of sources, such as correctly identified perovskite peaks and misidentified peaks, that expression is no longer appropriate. We therefore need to generalize Equation (3.5) to account for some fraction of misidentified peaks, which can be done by adding an additional term to the likelihood that accounts for the different possible sources of each data point. To simplify this process, we assume that each position measurement is drawn at random from one of two possible populations: either it is properly identified and drawn from the true sample line population or it is misidentified and draw from a population of confused lines. Just as before, the sample line population is represented merely by a normal distribution about the measured value. Based on how lines are identified, we expect that the confused line population is reasonably modeled by a simple flat distribution centered on the expected position value with a width of Δp, where the value of the width corresponds to how close an observed line must be to the expected position in order to be counted as an identified peak. We estimate a reasonable value for the width of $\Delta p \approx 0.02$, which is roughly a few times larger the typical uncertainties on line position determined from Equation (3.10). Therefore, the total likelihood for each data point is just a "mixture", or a weighted average, of these two distributions (*Sivia and Skilling*, 2006): $\mathcal{L} = \prod_i \left((1-f)N(p_i - p_i^{\mathrm{mod}}, \sigma_i) + \frac{f}{\Delta p}\right)$, where f is just the expected fraction of the data points that are incorrectly identified. When there are no misidentified peaks, $f = 0$ and we recover the standard least-squares approach described above. Taking the log of both sides as before, we obtain the final expression for the Bayesian mixture model:

$$\begin{aligned}\log\mathcal{L} &= \log\left\{\prod_i\left(\frac{(1-f)}{\sqrt{2\pi}\sigma_i}\exp\left[-0.5\left(\frac{p_i - p_i^{\mathrm{mod}}}{\sigma_i}\right)^2\right] + \frac{f}{\Delta p}\right)\right\} \\ &= \sum_i \log\left(\frac{(1-f)}{\sqrt{2\pi}\sigma_i}\exp\left[-0.5\left(\frac{p_i - p_i^{\mathrm{mod}}}{\sigma_i}\right)^2\right] + \frac{f}{\Delta p}\right) \end{aligned} \tag{3.11}$$

In order to use this expression, we need only choose reasonable values for the width of the line identification window and the line fraction like $f \approx 0.1$. In truth, we can generally use values as large as $f = 0.5$ with the only downside being that we will increase the uncertainty in our determined best-fit values. In deriving this generalization of the least-squares method in so few steps, we can see the general power behind the Bayesian framework, which reduces everything to the evaluation of probabilities.

We apply this Bayesian mixture model to obtain reasonable estimates for the unit cell dimensions of the perovskite samples for each spectrum. The results of the crystal model fits are summarized in Figure 3.3, showing the variation of the normalized axial ratios with compression. These normalized axial ratios, defined in the figure caption, represent a simple scaling of the unit cell parameters by the average crystal length scale $V^{-1/3}$. Just as before, we plot these against inverse length scale, rather than pressure, since it reveals the purely geometric behavior of the crystal, rendering the plots essentially independent of temperature. In the left panel of the figure are the results for the 13% Fe-bearing sample. These are obtained using a direct application of the mixture model described above. Notice that the normalized axial ratios appear to evolve nearly linearly with compression, and show relatively small amplitude scatter that remains roughly constant over the compression range. In contrast, the modeling results for the 0% Fe sample is shown in the right panel of the figure. In light "plus" symbols, we show the modeling results using the same method as for the Fe-bearing sample. As can be clearly seen at the left end of the figure, the fitted axial ratios diverge from their expected trends rather dramatically at for a number of spectra, especially at low compressions. Unfortunately, the Fe-free sample has a rather limited number of fitted peak positions, due to poorer spectrum quality resulting from small sample size. While there are enough peaks to fit a model, the reduced number is significantly more susceptible to the deleterious effects of peak misidentification. Examination of the peak position map in Figure 3.1 does not show any obvious trends indicating poor peak identification, however. In order to deal with the issue, we therefore turn to the use of priors, which can provide useful outside constraints to any statistical analysis problem.

Though we have presented peak identification and crystal modeling in two separate sections, in truth they exists large overlap between the two. As alluded to earlier, peak identification and fitting is inherently an iterative process, where peak identification improves as the model is refined with the addition of each new peak. At the same time, however, the addition of an incorrect peak at such an early stage can be quite detrimental

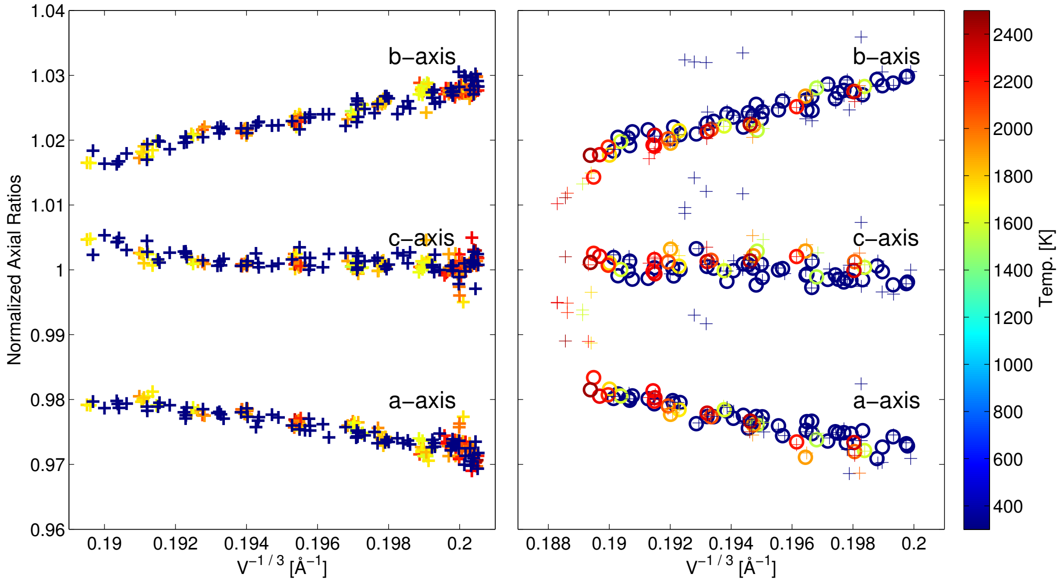

Figure 3.3: The normalized axial ratios are plotted in a purely geometric space against $V^{-1/3}$, where 13% Fe-bearing sample is on the left and the Fe-free sample on the right. For the 13% Fe sample, the data appears just as it should, being nearly linear, without temperature-dependent systematics, and with minimal scatter. The same unconstrained fit to the peak positions of the 0% Fe sample shows strong deviations from expected physically reasonable behavior, as shown in the "+" symbols. We therefore re-perform the analysis constraining the axial parameters by applying an appropriate prior, thereby suppressing this unphysical behavior, as show in the open circle symbols.

when using the standard least-squares approach. The Bayesian mixture model significantly reduces the effect of the misidentified peaks on the overall fit, making it useful both for obtaining final estimates as well as early on in the peak identification and fitting process.

As discussed initially in Section 3.2, it is common to use uninformed, or flat priors, when in the data dominated regime. As just discussed, however, the Fe-free data set does not contain enough visible spectral lines to yield fits that conform with our prior knowledge about the system. We know that the large deviations from linear behavior exhibited in the best-fit values at low pressure are not realistic, and therefore must result from an inability of the data to properly constrain all the crystal dimensions simultaneously. We can reasonably aide the fitting process that imposes the prior of linear behavior onto the crystal model. To do this, we simply fit the initial retrieved axial ratios with a line using a standard robust polynomial-fitting routine in MATLAB (this method also uses a Bayesian technique, but we will not discuss the details here). In doing so, we can estimate the expected linear trend of each of the normalized axial ratios. The prior is then constructed as a normal distribution centered on the robust-fitted line, whose width is chosen to be sufficiently small as to constrain the retrieved axial ratios. Given the scale on the y-axis, it is not unreasonable to expect that we should need to constrain the fit to variations of about half a percent from the linear trend. The results from that prior-constrained fit are shown in think circles over-plotted in the same figure. As is clearly visible, the problem has been fixed using this outside constraint. Also, we can see that the axial values are still able to vary from one spectrum to the next. This reflects how the prior can act as additional pseudo-measurement, affecting the fitted results but without totally eliminating the freedom necessary for the model to best explain the data. Though we do not show it here, it should come as no surprise that the unphysical variations in unit cell dimensions also induced large offsets in the apparent crystal volumes. Upon applying this prior-based method for constraining the crystal dimensions, all of these large deviations were seen to disappear, also providing useful verification for the approach.

3.3.3 Obtaining Unbiased Estimates and Uncertainties for Equation of State parameters

The last step of the inversion process is to obtain the equation of state (EOS) from the measurements of the crystal volumes at each pressure and temperature. The EOS expresses the dependence of crystal volume on pressure and temperature. This is captured using a set of parameters, including the zero-pressure volume, V_0, the resistance to compression given by the bulk modulus at zero pressure K_0, the pressure dependence of the bulk modulus K_0', as well as parameters that relate to the thermal properties of the crystal. While the details of the equation of the state model are given in Chapter 2, we focus here on the more general aspects of the fitting procedure and interpretation of the resulting posterior distributions.

To start with, we should begin with a discussion of priors for the parameter values. While we might wish to have enough data such that we needn't rely on any outside information, this is very often not the case. In fact, it is oftentimes not merely a question of the quantity or quality of the data, but an issue of the diversity of the data. A physical model for any complex system always relies on information about many aspects of that system's behavior. We are limited, however, in what types of data we can collect, and how directly that data speaks to different properties of the system. There are always properties that are poorly constrained, no matter how much data we collect of a particular type, and therefore we must rely on outside or prior information to constrain those values. The equation of state of perovskite is a perfect example of this. In thermodynamic equilibrium, perovskite is stable over nearly the entire range of lower mantle pressures and temperatures, but it becomes unstable at low pressures, (transforming back to pyroxene, a stable lower pressure phase of the same composition). This means that within its stability field, it is impossible to directly measure zero-pressure properties for perovskite, since the transition occurs between around 25 GPa. Unfortunately, the equation of state models rely explicitly on the zero-pressure properties. Because of this, it is generally not practical to collect enough high-pressure measurements to provide strong independent constraints on the zero-pressure properties. This makes intercomparison of the results of different studies particularly difficult, since there exist strong correlations between the zero-pressure parameters, especially the values of the bulk modulus K_0 and its pressure derivative K_0' (*Angel*, 2000; *Bass et al.*, 1981).

A typical, but systematically biasing, approach is to fix some subset of the parameters to particular values, and then fit the remaining parameters as normal. This is a logical, but probabilistically incorrect step to take. A common example of this is, for data that constrain the slope but not the curvature of the P-V compression curve is to fix the value of K_0' to 4. From a physics standpoint, there is nothing special about a value of 4—rather its special nature is related to the particular details of a commonly used equation of state. If the desire is merely to estimate the volume over the range of the data, there is nothing particularly wrong with this approach, since we have already supposed that the curvature is so small in the data region as to be non-noticeable. If, as is usually the case, we would like to take our derived equation of state and then extrapolate to higher pressures, perhaps the core-mantle boundary, then the practice of fixing parameters is a serious mistake. It will severely (and artificially) limit the allowable range of volumes predicted outside the data region. This is because uncertainties in K_0' spill over into uncertainties in other correlated parameters, and such correlated uncertainties are entirely ignored when parameters are fixed—we obtain unrealistically narrow uncertainty regions that bear little relation to our true knowledge about the system given the data. By ignoring the known trade-offs between K_0 and K_0' (the slope and the curvature), we are doomed to wildly underestimate the uncertainties resulting from extrapolation.

Luckily, this challenge is very simple to address in Bayesian statistics, since it merely reflects the desire on the part of the investigator to use outside knowledge or *prior information* to help constrain the parameter estimation. While fixing parameters is generally a poor idea for the reasons discussed above, applying a probability density function that encodes a prior belief about the parameter values is a simple and explicit way to properly constrain the fitting parameters without paying the heavy costs associated with fixing parameters. The simplest informative prior is a normal distribution for each model parameter, which therefore suggests the range of reasonable parameter values without insisting on a particular value. For example, we might replace the assumption that $K'_0 = 4$ with instead a weakly informed prior of $K'0 = 4 \pm 1$. In order to incorporate this information, we turn to Bayes' theorem with says the that posterior probability is proportional to the product of the likelihood and the prior (since independent probabilities multiply). Therefore the log-posterior probability is just the sum of the two independent pieces $\log p = \log \mathcal{L} + \log \Pi$, where a simple normally distributed prior adds a squared residual term for each parameter

$$\log \Pi = -\frac{1}{2} \sum_i \left(\frac{M_i - \bar{M}_i}{\sigma_{M_i}} \right)^2 \tag{3.12}$$

where $\bar{M}_i \pm \sigma_{M_i}$ is the prior information about the i^{th} model parameter M_i. Based on its mathematical form, it is clear that such a normal prior distribution acts as an "apparent" measurement of each model parameter.

For our study of perovskite, the issue of priors on the equation of state parameters arises when considering the zero-pressure volume. While it is not possible to measure the equilibrium value of V_0, since it does not exist, perovskite is metastable at ambient pressure and temperature. The kinetics of the backward transition are so slow that the transformation is inhibited, Thus, there are ambient measurements of metastable perovskite. It is an open question as to how relevant the metastable behavior is to the behavior within the stability field. Many phases undergo softening and changes in their axial ratios as they approach a phase transition. It is usually assumed that these differences are negligible and thus we can make use of metastable measurements. Presuming that is the case, we are faced with a perfect opportunity to make use of priors.

In the absence of zero-pressure measurements of the sample volume, the proper way to constrain the parameter values is to impose a reasonable composition-dependent prior for V_0. It has long been known that many crystals that exist as solid solutions show nearly linear variation of the volume as a function of composition. *Kudoh et al.* (1990) showed that the measurements of metastable perovskite volume are roughly linear in ferrous iron composition. We use a similar analysis of the previous ambient pressure perovskite volume measurements, shown in Figure 3.4, to obtain our prior on V_0 as a function of composition. Based on the large observed scatter in that figure, as compared to the small individual measurement errors, it is clear that there is some source of additional scatter inducing large sample-to-sample variation. We therefore fit these data with a straight line, also allowing for an additional scatter term. The solid line shows the best-fit linear trend, while the dashed lines show the 68% confidence intervals, which are dominated by the intrinsic scatter term. We can summarize these results by representing them as a normal distribution whose mean is linear in composition and whose standard deviation is nearly independent of composition.

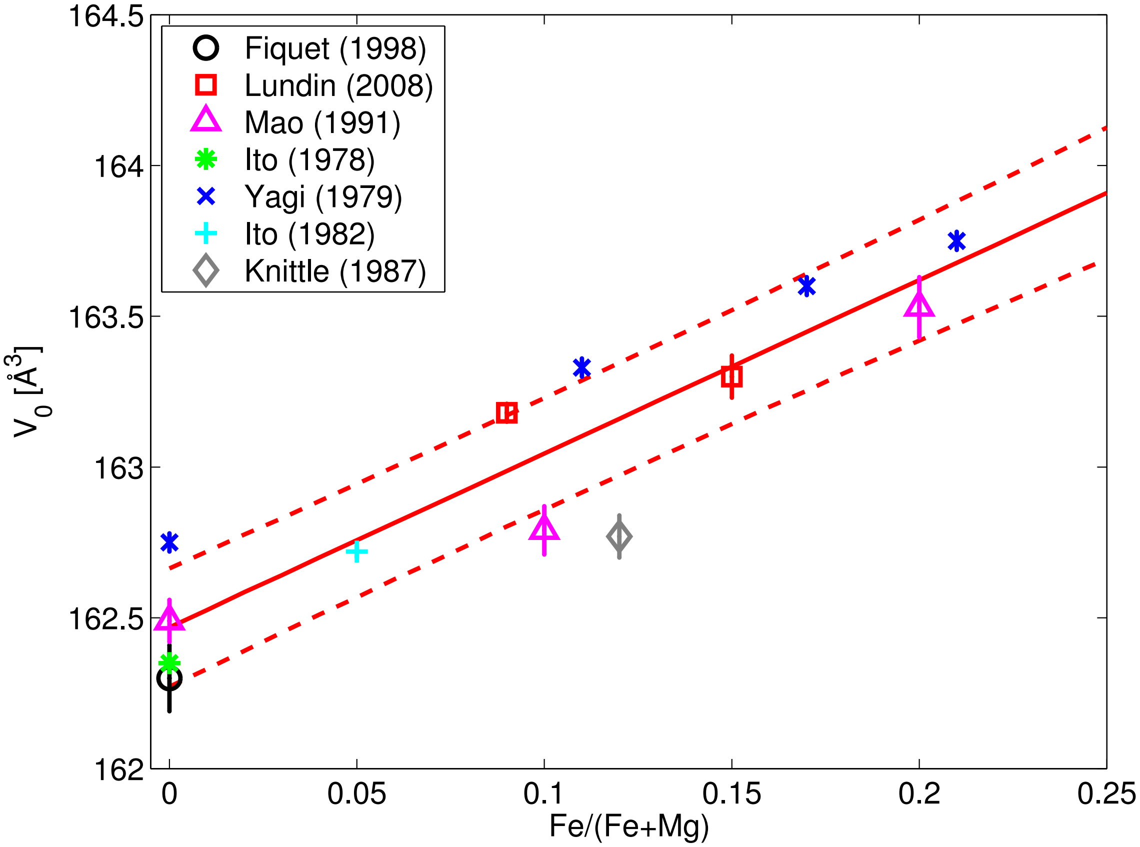

Figure 3.4: Examination of previous zero-pressure volume measurements as a function of iron content, using a linear fit combined with an intrinsic scatter due to sample-to-sample variation. 68% contours are shown.

We therefore use this curve, sampled at 0% and 13% Fe to provide the prior on $V0$ for our fit. As far as the priors on the other parameters, they are much less important and unnecessary, as the combination of the data and the V_0 prior are more than adequate to provide constraints on the parameters values.

With a reasonable prior in hand, we can then fit both the ambient temperature and laser-heated data to determine the equations of state of the two perovskite samples. The details of that fitting procedure were already discussed in Chapter 2, and so we will focus here on interpreting the resulting posteriors. After obtaining the best-fit, we used the general Bayesian approach of optimal estimation, calculating the covariance matrices from the curvature of the posterior in log-space, as expressed by Equations 3.7 and 3.8. These results can be directly visualized using a “stair-step” plot, see Figure 3.5, which shows the correlated uncertainty bounds on each pair of parameters by plotting every parameter against every other parameter. The nested 68% and 95% confidence ellipses are shown in the central region for each parameter pair, with the 0% and 13% Fe samples in black and red lines, respectively. These plots show 2-D marginalized confidence regions, which means that the variations in all other parameters (not the 2 being plotted) are integrated over. This means that these are *projections* from a five-dimensional space down into a two-dimensional space. Note that they are not slices, which would happen when certain parameters are fixed, rather than marginalized over. The one-dimensional marginal probability distributions for each parameter are shown on the edges of the figure. These correspond directly to the best-fit and 1-D uncertainties on each parameter. The subfigures are color coded with cold parameters shaded blue and hot parameters shaded red. They are also organized such that correlations amongst cold parameters are at the top, the correlations between the two hot parameters are at the bottom right, and the cross-correlations are shaded gray in the lower central region. From these plots, it is clear that the correlations are strong between purely hot and purely cold parameters, but nearly zero for mixtures. Additionally, it is clear that the two different compositions have differences in their equations of state that are statistically discernible (though not necessarily all physically relevant, see Discussion in Chapter 2), where there is little to no overlap in the 2-D 95% confidence regions for any of the parameter pairs. A final thing that must be pointed out is the ability of the stair-step plot to reveal significant parameter combination differences that are poorly visible when looking at one dimension at a time, but plain to see in two dimensions. The best example of this is shown by the correlation between K_0 and K_0', which initially appear largely consistent for the two compositions, when looking at the 1D probability distributions, but can be seen to be over inconsistent with one another at over the 3-σ level when viewed in two dimensions.

These results should then be compared in exactly the same way with previous investigations of perovskite equation of state. Unfortunately, this task is made rather difficult for two reasons. Firstly, none of the previous investigations report the full covariance matrix (or equivalent correlation matrix) for their equation of state parameters. Second, and more important, is the fact that all previous investigations fixed a number of important equation of state parameters in order to improve the apparent constraints on the parameters. Unfortunately, this results in strongly underestimating the parameter uncertainties as discussed in detail above. We are forced, therefore, to settle for the more qualitative comparisons already presented in Chapter 2.

3.4 Conclusions

In this chapter, we have used the data analysis problems that arise in powder diffraction and equation of state studies to demonstrate the power and flexibility of the Bayesian Statistics. We have demonstrated how "standard" statistical techniques, such as least-squares regression, are in fact fully consistent the Bayesian framework. We also presented a number of novel analysis methods that show how Bayesian thinking can allow for the development bespoke statistical procedures that do not suffer from the use of ad-hoc assertions, but are

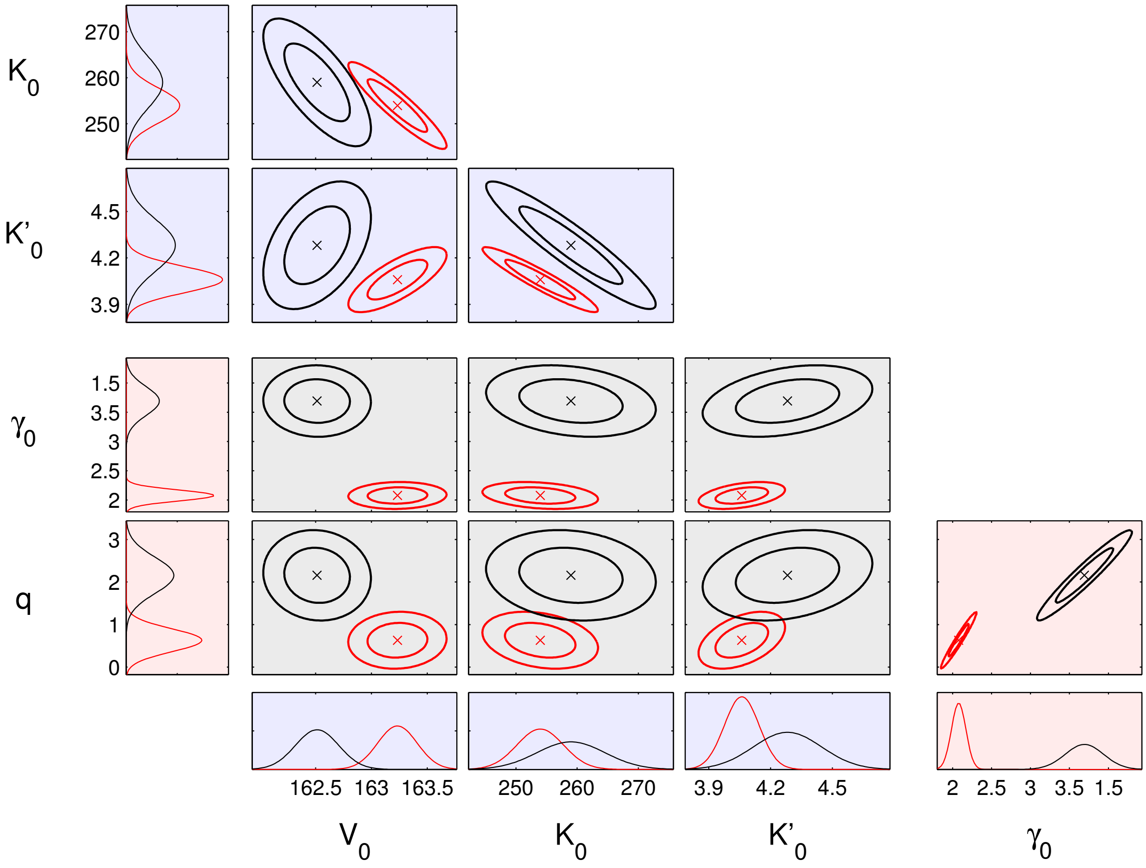

Figure 3.5: Two-dimensional posterior plots show the correlations between the different equation of state parameter uncertainties. The central plots show the nested 68% and 95% confidence regions for the 0% Fe sample in black and the 13% Fe sample in red. The marginalized one-dimensional probability distributions are shown for each parameter along the edges. Additionally, the plots are organized with the cold EOS parameters shown with a blue background, the hot parameters with a red background, and the cross-correlations with a gray background. With this scheme, it is clear that the cross-correlations between cold and hot parameters are rather small, nearly horizontal confidence ellipses, whereas the correlations between hot or cold parameters are significant, as indicated by highly angled confidence bounds.

rather directly rooted in probability theory. It is our hope that this text will help to show readers that there is no “black-magic” in Bayesian techniques, and adopting them does not require the rejection of previous statistical understandings. In short, we hope to convert others to the Bayesian way of thinking, by way of recognizing that all

Chapter 4

Coordinated HArd Sphere Model (CHASM): A simplified model for oxide liquids at mantle conditions

Abstract

We develop a new simplified model to understand and predict the behavior of oxide and silicate melts at extreme temperatures and pressures. The **C**oordinated **HA**rd **S**phere **M**odel (CHASM) is based on the extension of the hard sphere mixture model, accounting for the range of oxygen coordination states available to each cation in the liquid. By utilizing approximate analytic expressions for the hard sphere model, this method is capable of predicting complex liquid structure and thermodynamics while remaining computationally efficient, requiring only minutes of calculation time on standard desktop computers. We demonstrate how the parameters of the model can be obtained by training on equation of state data for a variety of crystal polymorphs, which discretely sample the range of available coordination states. This method is applied to the MgO system and then used to calculate the coordination evolution and equation of state of MgO liquid over a wide range of pressures and temperatures. The results produced by CHASM are finally evaluated by

comparison with predictions from published first-principles molecular dynamics calculations, indicating that CHASM is performing well and accurately capturing the dominant physics controlling the behavior of MgO liquid.

4.1 Introduction

Understanding the behavior of silicate liquids at mantle pressures is crucial to reconstructing the early evolution of Earth and other terrestrial planets. While molten silicates play a minor role in the deep mantle today, their material properties directly determined the Earth's evolutionary path as it cooled from a nearly whole-mantle magma ocean to its present-day solid state.

Recent first-principles theoretical calculations (*Stixrude et al.*, 2009) and experimental shock-wave investigations (*Mosenfelder et al.*, 2009) indicate that melting perovskite requires significantly less total energy than previously thought, supporting the idea of a deep potentially whole-mantle magma ocean early in Earth's history. The evolution of this magma ocean as it cools and crystallizes is dominated by the poorly understood behavior of silicates at extreme pressures and temperatures. Probing these conditions is difficult for both theory and experiment, and further challenges are posed by the large relevant compositional space, including at least MgO, $\mathrm{SiO_2}$, and spin-dependent FeO. The primary tool for theoretical work, first-principles molecular dynamics (FPMD), has been used to probe the behavior of a variety of silicate liquids including MgO (*Karki et al.*, 2006), $\mathrm{SiO_2}$ (*Karki et al.*, 2007), $\mathrm{Mg_2SiO_4}$ (*de Koker et al.*, 2008), $\mathrm{MgSiO_3}$ (*Stixrude and Karki*, 2005), and $\mathrm{Fe_2SiO_4}$ (*Ramo and Stixrude*, 2012). While extremely powerful, this approach has limitations including high computational cost, lower bounds on temperature due to relaxation constraints, as well as restrictions to length scales and time scales that are many orders of magnitude smaller than those relevant to the Earth or experimental methods.

As a compliment to accurate first-principles calculations, we have developed the **C**oordinated **HA**rd **S**phere **M**odel (**CHASM**), which represents oxide and silicate liquids by directly tracking the relative populations of differently coordinated cation species as a function of temperature and pressure. This model is based on extending the standard hard sphere mixture model (*Mansoori and Canfield*, 1969), which was recently applied to silicate liquids by *Jing and Karato* (2011). In order to reasonably represent compression-induced changes to internal energy, we implement a perturbation to the standard hard sphere model, in which the internal energy is represented by empirical pair potentials that are coordination dependent. Using the principle of free energy minimization, we determine the relative populations of the different coordination states. Since the hard sphere mixture model is a well-developed theory with trusted long-standing analytic approximations (*Mansoori and Canfield*, 1969), this minimization technique is computationally cheap, allowing rapid determination of the energies and coordination populations on a modern laptop computer, enabling in the near future complete mapping of the wide compositional spaces relevant to early Earth evolution. We also demonstrate how the well-known equations of state for a large set of solid structures can be used to determine appropriate values for the pair potentials, thereby raising making CHASM a fully predictive model for oxide liquids that does not require fitting liquid data.

4.2 Background

Understanding the behavior of liquids has been a challenge that has puzzled researchers for nearly a century. Since the need to understand and be able to predict liquid properties over a range of conditions is pervasive, progress has generally been distributed across a number of scientific disciplines, including thermodynamics and statistical mechanics, experimental geochemistry, and computational mineral physics. The primary research goals in each of these fields vary greatly, from gaining a mechanistic understanding of liquid properties, to providing an empirical description of compositional effects, to gaining a first-principles atomistic view of compression behavior for liquids with specific chemistries. Though each method is focused on a subset of the overall problem, together they contribute useful and complimentary insights to our understanding of liquids under a range of conditions. To provide context for the new model development described here, we give a short review of the existing approaches to modeling and measuring liquid properties from these different fields.

4.2.1 Ideal Mixing Models for Experimental Data

We start with the simplest possible model for liquid thermodynamics, which considers a set of endmember compositions that mix ideally. While this method is computationally simple, it relies on accurate and difficult to obtain experimental measurements of the properties of a range of liquids of different compositions. Such experiments include ambient pressure sound speed measurements, like those reported in *Ai and Lange* (2008), and measurements of density and internal energy in ultra high-velocity shock-wave experiments, like those of

Asimow and Ahrens (2010) and *Thomas et al.* (2012). Given a measurement of the volumes and energies of liquids for a range of compositions at a particular P-T condition, the ideal mixture model supposes that the volumes, entropies, and energies are additive and depend linearly on the fraction of each end-member composition (*Bottinga and Weill*, 1970). This assumption supposes that the excess energy and excess volume of mixing are both negligible and that the entropy of mixing is well described by simple ideal mixing of end-members. While this method employs a strong assumption of ideality, it is shown to perform fairly well under many conditions. One major difficulty with this method is that it requires extensive data sets for liquids spanning large ranges in composition in order to uniquely constrain the behavior of the endmembers. Additionally, it is fairly limited in its ability to extrapolate beyond the experimental pressure and temperature conditions, providing instead a sensible method for interpolation between data points. Even so, the ideal mixing model has found great success at describing the thermodynamics of complex many-component oxide and silicate liquids, as presented in *Ai and Lange* (2008), for instance.

Despite extensive success, there are notable shortcomings to the straightforward application of this approach, as explored by (*Guo et al.*, 2013), in which they show evidence for large changes in the partial molar volume of the iron oxide component as a function of overall composition. In particular, this change is interpreted as a strong composition dependence to the oxygen coordination number of iron, which thereby exerts large controls over the volume of the solution since 6-coordinated iron is significantly more compact than 4-coordinated iron. It is possible that the ideal mixing approach can remain useful in such cases by including separate endmembers that explore the different dominant coordination states for each cation, though the simplicity and predictive power of the method could suffer as a result. Preliminary results from shock-wave experiments on liquids in the $\mathrm{CaO{-}MgO{-}Al_2O_3{-}SiO_2{-}FeO}$ system show similar failures of the ideal mixture method (*Thomas et al., submitted* 2013), though it is not yet clear whether the practical approach of adding more end-members that sample the range of dominant coordination states is sufficient to address shortcomings of the approximation.

4.2.2 Atomic Simulation Using Molecular Dynamics

An entirely different and theoretical approach is to perform molecular dynamics simulations to directly track the motions of the atoms that make up the liquid. In this method, the trajectories of a finite number of atoms are evolved within a simulation cell. Periodic boundary conditions are imposed on the simulation, mimicking an effectively infinite tiled domain composed of side-by-side repeated images of the simulated region. Using basic Newtonian mechanics, the positions and velocities of each atom are updated in response to the imposed forces, which result from complex electronic interactions. The method of determining the forces at each time step separates this approach into two possible forms: Empirical Potential Molecular Dynamics (EPMD) versus First-Principles Molecular Dynamics (FPMD).

In the empirical approach, the interactions amongst the atoms are parameterized in terms of analytic potentials, typically using standard forms that capture the general features of atomic interactions. In the simplest and most common form, only pairwise additive interactions are considered, yielding pair potentials which are typically composed of a sum of repulsive and attractive terms. Since these equations are simple analytic expressions, the evaluation of energies and forces at each time step is comparatively rapid, and therefore EPMD is typically capable of representing between 10^4 and 10^5 atoms per simulation cell over relatively moderately long simulation times of order few nanoseconds. The major challenge with EPMD relates to whether that the parameter values for the empirical potentials are applicable outside the conditions under which they were trained and thus the quality of the simulation results can vary widely, depending on the details of the system. Improvements to accuracy can often be made by including many-body terms in the interaction energies, which can be important for determining bonding and coordination effects in many liquids *Daw et al.* (1993). Despite the challenges involved in training atomic potentials that are capable of remaining accurate over large compression ranges, there have been a few studies that apply this method to high-pressure systems, though agreement with experiment tends to break down at higher pressures (e.g., *Spera et al.*, 2009).

First-Principles Molecular Dynamics (FPMD) avoids many of these issues surrounding accuracy of the applied forces by determining them directly using first-principles quantum mechanics calculations. This strategy has gained much popularity over the last decade, as computational resources have become more powerful and affordable. In this approach, the forces imposed by the electron cloud distributions at each time step are determined self-consistently using Density Functional Theory, a practical quantum mechanical method that can provide approximate solutions to Schrödinger's equation. While accuracy is obviously less of an issue, this approach imposes a significantly higher computational burden, requiring long calculations times on parallel processor machines. Capable of handling only a handful of atoms (of order 10^2 to 10^3), it is limited to small simulation cells with length scales of roughly 10 angstroms, short evolution times of the order of picoseconds, and high temperatures above 3000K, in order to remain computationally feasible. This means that liquid behavior can only be modeled for temperatures well above the liquidus, thereby avoiding impossibly long equilibration times. In such analyses, it is implicitly presumed that picosecond evolution times for at most 1000 atoms are sufficient for accurate thermodynamic averaging. While this is likely adequate for equation of state properties of most systems, any physical effects that rely on length scales approaching the size of the simulation cell will not be captured by this approach, making the determination of dynamic material properties like diffusivity and viscosity potentially much more uncertain.

Despite these limitations, this method has been found to be extraordinarily useful, and there have been numerous recent studies on high-pressure Mg-silicates using FPMD, covering a finite set of compositions across the $\mathrm{MgO}-\mathrm{SiO}_2$ join (see, for instance, *de Koker et al.*, 2013). By focusing on a particular fixed composition and calculating the time-averaged free energies for a small set of volumes and temperatures, investigators are able to fit parametric equation of state forms to describe that liquid over a continuous pressure-temperature region. Due to computational limitations, the FPMD approach is therefore focused on simulating the detailed high-temperature behavior of a small set of particularly interesting liquid compositions.

4.2.3 Basic Hard Sphere Model

The search for a simple mechanistic model to describe liquids traces all the way back to the birth of our understanding of the molecular nature of matter, to the work of Johannes Diderik van der Waals in 1873. In his investigations of the behavior of fluids at very high temperatures, he sought an explanation for the disappearance of the vapor to liquid transition upon compression (for temperatures above the critical point). This led him to develop an improved equation of state, the familiar van der Waals equation:

$$P = \frac{RT}{V-b} - \frac{a}{V^2} \tag{4.1}$$

In proposing this expression, van der Waals hypothesized that a fluid is composed of equivalent discrete particles of matter (molecules). Due to their finite size and rigidity, the molecules that make up the liquid are forced into non-overlapping configurations, which therefore reduces the free volume, which describes the fraction of the volume available for the molecules to occupy. In the first term of his equation, van der Waals replaced the volume in the ideal gas equation with an effective free volume, which is just the total volume V less the excluded volume b due to the molecules, $V_{\text{free}} = V - b$. The second term represents weak attractive forces—now referred to as van der Waals forces—that help bind the fluid together and can be represented to leading order as proportional to $1/V^2$. With this equation, van der Waals was able to represent the critical behavior of fluids, but more importantly, he laid the groundwork for understanding liquid thermodynamics in terms of effects resulting from the finite size of fluid particles.

Nearly a century later, the idea of representing a fluid as a collection of hard spheres experienced a sudden renaissance with the advent of large computing resources. As hinted by the early success of the van der Waals equation, the behavior of liquids is dominated by strong short-range repulsive forces that prevent liquid particles from occupying the same space. Therefore, the behavior of liquids can be represented using a simple hard-wall potential function:

$$\Phi_{ij}\left(r\right) = \begin{cases} \infty & ; \quad r < R_i + R_j \\ 0 & ; \quad r \geq R_i + R_j \end{cases} \tag{4.2}$$

where the internal energy Φ_{ij} of particles i and j is zero for all non-overlapping configurations. Particle overlap is expressly forbidden by a jump to infinite energy when their separation distance is less than the sum of their radii $r < R_i + R_j$. This almost trivial equation is the basis of the hard sphere model, which has been the source of a wealth of information about the general properties of liquids. The initial investigations of hard sphere fluids represented some of the first computer simulations ever performed. Later efforts by many investigators provided analytic expressions that approximate the results of these computational simulations to high degrees of accuracy. It is these analytic expressions which make the hard sphere model so attractive, since they enable rapid calculation of liquid properties many orders of magnitude faster than molecular dynamics simulations and on standard personal computers rather than large super-clusters.

We can intuitively understand some of the basic properties of the hard sphere model and liquids in general by examining snapshots of hard sphere fluids at different packing fractions. At ambient conditions, the internal energy of a liquid plays only a very minor role in determining liquid thermodynamics—instead, the dominant term in the free energy of the system comes from the configurational entropy resulting from the freedom of liquids to adopt nearly any structure that avoids particle overlap. In particular, quantities like the expected particle separation distance and pressure for a given volume are primarily dependent on the number of available microstates or configurations that exist for a particular set of macroscopic variables (e.g., volume and temperature). This structural freedom is cast in terms of the free volume available for each particle to move through. At the relatively low densities near vapor equilibrium considered by van der Waals, the free volume is shared by all liquid particles and is thus simply the total volume less the excluded volume, as reflected in the van der Waals' equation. It is important to note that the excluded volume is larger than the volume occupied by liquid particles. For low density configurations of hard spheres, the closest two particle centers can get to one another is the sum of their radii. This results in an excluded region for a pair of particles that has a radius of twice the sphere radius, and thus eight times the volume. On a per particle basis, the excluded volume is therefore four times the particle volume for disperse particles, as was noted by van der Waals. This approximation breaks down, however, for dense packings, where each liquid particle can have multiple nearby neighbors. In such tightly packed configurations, the excluded volume regions of neighboring particles overlap one another, resulting in a total excluded volume that is only a fraction of van der Waals' upper limit.

To visualize the effect of packing density on the excluded volume, Figure 4.1 shows a snapshot of a 2D hard sphere fluid at a range of liquid densities. The relative liquid density

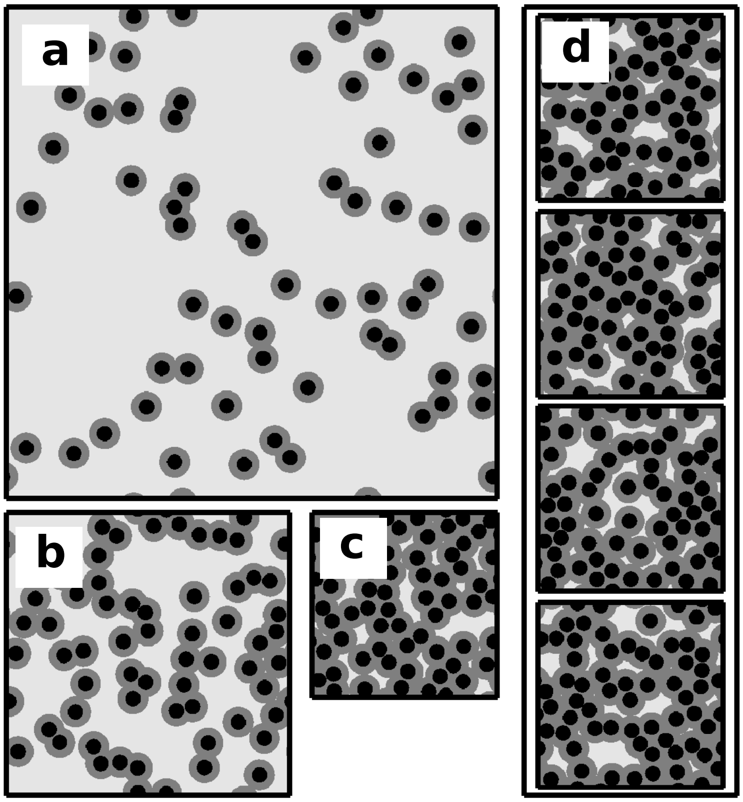

Figure 4.1: Snapshots of the simple hard sphere model for a range of packing fractions. Each hard sphere particle is show in black and the gray halo around it represents the excluded volume region that other particle centers are prevented from occupying. Panels a–c span the range of packing fractions from nearly ideal non-interacting fluids to extremely dense liquids. Panel d shows a variety particle configurations at fixed density.

is expressed in terms of the packing fraction

$$\eta \equiv \frac{\langle V_{\mathrm{HS}} \rangle}{V} = \frac{1}{V}\frac{4\pi}{3N}\sum_{i}^{N} R_i^3 \tag{4.3}$$

which is just the average hard sphere volume V_{HS} scaled by the volume of the liquid per particle, V. In the single component hard sphere fluid, the expression simplifies since all particles have the same fixed size. [1] Each sphere is shown in black with a gray halo around it representing the potential excluded volume region. In the disperse (gas-phase) limit, shown in panel a of Figure 4.1, particles generally lie far from one another interacting through collisions only rarely. In this case, the finite particle volume is negligible, and therefore particles can be treated as point-like, resulting in ideal gas behavior. For moderate packing fractions (low density liquid phase), particles regularly interact with one another revealing the importance of the excluded volume. As shown in Figure 4.1 panel b, particles at these moderate densities tend to have only one or two immediate neighbors, and therefore the excluded volume is approximately additive. For high packing fractions (high-density liquid phase), particles are forced to significantly crowd together, causing their excluded volume regions to strongly overlap, as seen in panel c of Figure 4.1.

A crucial attribute of liquids that relates directly to the excluded volume effect is the dependence of excluded volume on configuration. To demonstrate this, a number of random snapshots at constant packing fraction are shown in panel d. From these images, it can be seen that the evolving random liquid structure leads to fluctuations in the total ex-

---

[1] It should also be noted that all of the figures depicting snapshots of hard sphere configurations are shown for explanatory purposes, created from calculations of 2D hard sphere liquids. All of the equation of state calculations are carried out properly in three dimensions using the analytic approximate methods reviewed in the Appendix.

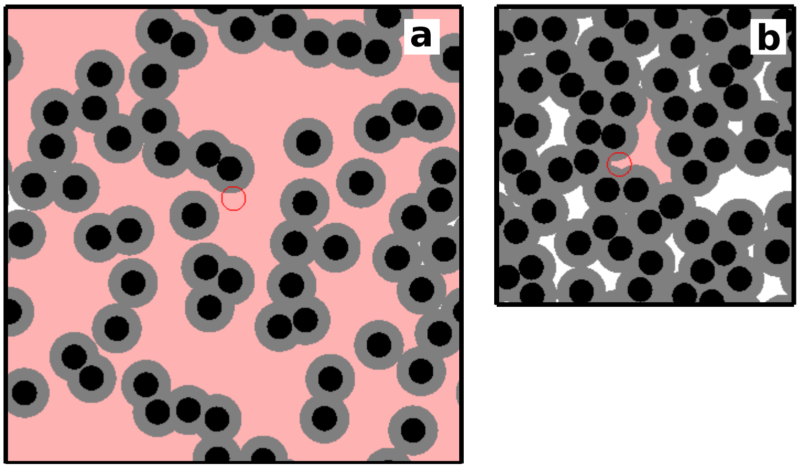

Figure 4.2: Snapshots of a hard sphere fluid at moderate (a) and high (b) packing fractions. These configurations are identical to panels (b) and (c) of Figure 4.1, but a single particle, outlined in red, has been removed. The light red shaded region shows the free volume available to the reference particle. At low packing fractions (a), particles all share the same free volume, whereas at high packing fractions (b), each particle is trapped by its neighbors.

cluded volume as well as the locally available free volume, shown in white. As the particles bounce around in the fluid, the degree of overlap between their excluded volume regions evolves as a function of configuration, resulting in complex sub-additive behavior for the total excluded volume of the liquid. Another consequence of tightly packed configurations is that particles no longer share the same free volume. In Figure 4.2, we repeat two of the same configurations from the previous figure, but in this case, one of the particles has been removed to display the local instantaneous free volume available for it to explore. The location of the removed particle is shown with a red outline near the center of the simulation cell, and its local free volume region is shaded in light red. In the low density case, the particle can access nearly the entire free volume of the liquid. In the high density case, however, the particle is confined within an instantaneous "cage" formed by its neighbors. This leads to equations of state and average liquid structures that have complex dependencies on the packing fraction. It is quite remarkable that such complex behaviors such as the sub-additivity of excluded volumes and the evolution of particle-specific free volumes can be well represented by the simple hard sphere fluid.

4.2.4 Previous Application of Hard Spheres to Silicate Liquids

The hard sphere model has been investigated through both computer simulation and analytic theoretical development over the last half-century. This has resulted in a rich and mature body of literature that is readily available for applications to any particular problem of interest. For the sake of brevity and readability, we will wait to present these classic results for hard spheres until it becomes relevant to our modeling efforts. We do want to draw attention, however, to the wealth of analytic approximations that are now available to describe many different aspects of the hard sphere model. These include fundamental thermodynamic relations like the equation of state, which relates the volume and the temperature to the pressure. The time-averaged structure of hard sphere fluids is also known analytically. Unlike crystalline solids that have fixed and periodic structures, liquids present a time evolving continuum of relative atomic positions. In order to handle this ambiguity, the average liquid structure is usually described in terms of the radial distribution function (RDF), which expresses the average local density of particles at a given separation distance relative to the density of the liquid as a whole. Most important for real-world applications is the fact that these properties are also known for random mixtures of hard spheres of varying sizes.

Given that most of the primary thermodynamic results for hard spheres have been in place for almost 40 years, it is rather surprising that hard sphere theory is rather new to the study of silicate melts. The first notable application of hard spheres to silicate liquids was presented by *Guillot and Sarda* (2006), in which they modeled the compression of particular fixed composition silicate liquids in order to investigate the solubility behavior of noble gases. In this first application of hard sphere theory to geologically relevant materials, *Guillot and Sarda* (2006) obtained the compression behavior for MORB, silica, and olivine composition liquids up to 10 GPa. This was accomplished by determining a single effective hard sphere size for each liquid that best matched a measurement of the 1 bar density at a reference temperature. It is important to note that this method yields only a very approximate description since it relies on representing each liquid as a collection of identical hard spheres, where it is unclear which structural aspects of a multicomponent liquid this physically represents. Even more important is the fact that this method is limited only to describing liquids at particular fixed compositions for which data already exists, and is therefore incapable of producing a predictive theory for liquids with unmeasured compositions. Nevertheless, this approach was reasonably successful at fitting low degree compression data with only a single free parameter.

More recently, *Jing and Karato* (2011) made considerable progress in the hard sphere description of silicate melts by fitting a set of hard sphere diameters to a collection of liquid density measurements in the $\mathrm{CaO{-}MgO{-}Al_2O_3{-}SiO_2{-}FeO}$ system. To accomplish this, they utilized an analytic approximate result for describing mixtures of hard spheres of different sizes (*Lebowitz*, 1964). (This classic work derived a generalized expression for the Percus-Yevick equation, which describes the radial distribution function for hard spheres, to allow for a random mixture of differently sized spheres.) Unlike the previous geological application, it is somewhat easier to draw a tie between the hard spheres of *Jing and Karato* (2011) and the physical aspect of the liquid they represent. They obtained a set of best-fit sphere sizes, where each sphere stands for a particular oxide component, like $\mathrm{SiO_2}$ or MgO. Since the spheres are meant to represent the relatively incompressible local units of the liquid, and they are specific to each oxide component, we can think of them as the region surrounding a cation and including a fraction of the shared oxygens that are directly bonded to that cation.

In order to obtain the oxide sphere sizes, *Jing and Karato* (2011) fit a large data catalog of liquid densities and bulk moduli (from sound speed measurements) for a range of temperatures and compositions at ambient pressure. Initially, they found the fixed hard sphere diameters that best matched the liquid data in the least-squares sense. Later in the paper, they also added in a physically motivated correction for temperature, which has often been applied in the hard sphere literature with moderate success, and an entirely empirical correction for pressure, to attempt to better match high-pressure shock-wave measurements. By including the temperature correction, they were able to demonstrate quite compellingly that they were able to reproduce the generally observed trends in density and compressibility for a large chemical system at ambient pressures. Using the empirical correction to allow the spheres to be slightly deformable with increasing pressure, they were also able to match low-pressure data up to 25 GPa from sink/float experiments. By adding a further empirical term to the pressure-dependent deformability of the spheres, they could better describe a small set of shock-wave measurements up to 135 GPa, but as all of these correction factors are no physically motivated, they can only provide modest improvements without adding significant insight. As the authors remark themselves, it would be better to be able to directly represent the compression-induced structural changes that are physically occurring within the liquid in order to have a theory with greater predictive and explanatory power.

4.2.5 Local Liquid Structure

As discussed in the previous section, efforts to adapt hard sphere theory to silicate melts up till now have focused on empirical determination of hard sphere parameters. Using this method, *Guillot and Sarda* (2006) and *Jing and Karato* (2011) were able to lean on the formalism of hard spheres to act as a physically motivated "averaging kernel", allowing them to effectively interpolate within the data set upon which the hard sphere sizes were trained. Extrapolation outside the training set, however, was fairly limited, as the hard spheres were not strongly connected with physically meaningful aspects of the liquid structure. It is for this reason that neither model was capable of accurately representing large compression ranges, with pressures over 30 GPa being generally inaccessible. Furthermore, any compositionally induced structure changes in liquids would also fall outside the capabilities of this approach. In the process of developing a more physically rooted application of the hard sphere model, we therefore have strong motivation to consider the structural attributes of dense silicate melts with a focus on how they are affected by changes in pressure and composition.

While studying atomic structure in melts is more difficult than in crystalline solids, there exist a number of methods that can be brought to bear in this effort. We start by describing the method that is the most directly sensitive to atomic structure, and is thus the primary tool for probing crystalline solids: x-ray or neutron diffraction. Each atom in the material acts as a scattering center for incoming high energy particles (typically photons or neutrons). Due to its wave-like nature, the phase of the incoming radiation is retained upon scattering, allowing the radiation from the collection of atoms in the material to add constructively or destructively. The general condition for constructive interference is that the path length traveled by a set of scattered rays is equal to an integer number of wavelengths. When considering a periodic crystalline structure, this is typically stated in terms of the spacing between adjacent crystal planes, d, in terms of Bragg's Law:

$$n\lambda = 2d\sin\theta \tag{4.4}$$

where λ is the wavelength of radiation used to probe the structure, n is the order of the reflection (typically $n = 1$ for most diffraction data since higher order reflections have negligible intensity), and 2θ is the scattering angle, defined as the angle between incoming and diffracted rays. From this equation, it is clear that radiation is scattered at high intensity into only a finite set of preferred directions which are characteristic of the crystal structure.

Liquids, on the other hand, are significantly more complex than crystalline solids since they lack periodicity. They, therefore, do not have large sets of atoms at precise locations from which coherent scattering can lead to sharp and intense diffraction peaks. Rather, they contain a collection of disordered atoms which lie at random absolute locations in space, reflecting a lack of long-range order. Nevertheless, the atoms occupy characteristic locations relative to one another, reflecting their considerable short-range order. As briefly mentioned above, the probabilistic nature of the structure of liquids is best-described in terms of the Radial Distribution Function (RDF), $g(r)$, which is proportional to the probability, $P(r)dr$, of finding two atoms a given distance apart:

$$g(r) \equiv \frac{\rho(r)}{\langle\rho\rangle} \qquad \text{and} \qquad P(r)dr = 4\pi r^2 \langle\rho\rangle g(r) dr \tag{4.5}$$

The RDF therefore gives the typical particle density a distance r away from a reference particle, normalized by the average density $\langle\rho\rangle$. The RDF is easily calculated from atomic simulations by counting the number of particles inside a set of thin spherical shells of varying sizes and averaging over the all particles and time-steps. Figure 4.3 shows a typical RDF for a simple monatomic liquid depicting the two primary features of average liquid structure: a sharp inner edge to the particle density and decaying density oscillations with increasing distance. The inner edge is roughly equal to the sum of particle radii and simply results from finite particle size and the corresponding tendency of particles to avoid overlapping one another. The decaying density oscillations about the average value reflect the limited short-range order in the liquid, where each subsequent maximum corresponds to the next-nearest-neighbor shell of particles, and at sufficient distance, all ordering tendency is lost yielding the average density value.

While the radial distribution function is not a direct observable for real-world liquids like it is for atomic simulations, it is nevertheless closely related to the data produced by x-ray or neutron diffraction. By replacing periodic lattices with a tendency toward local ordering, diffraction experiments for liquids result in a characteristic broad, low intensity pattern referred to as “diffuse scattering”, which displays azimuthal symmetry and slow os-

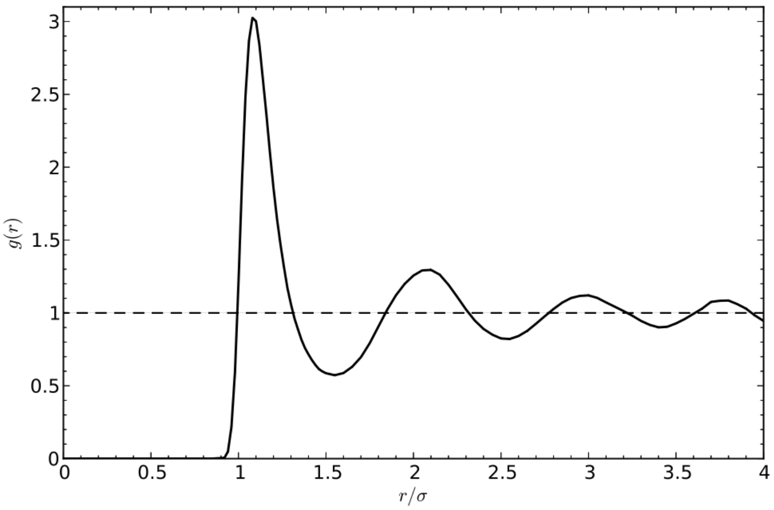

Figure 4.3: Depiction of the radial distribution function (RDF) of a simple model liquid (in this case the Lennard-Jones fluid). The primary features of the RDF are its primary peak, indicating the average nearest-neighbor particle distance, and its decaying oscillations, reflecting successive shells of neighboring atoms which become less distinct with distance as long-range disorder comes to dominate. [image credit: wikipedia.org]

cillations in intensities as a function of the scattering angle. The integrated 1D diffraction pattern can be straightforwardly transformed into the structure factor, which is a normalization of the scattered intensity giving the absolute scattering per atom as a function of the scattering vector $\vec{q} = \vec{k}_{scat} - \vec{k}_{inc}$ (the difference between the scattered and incident wavevectors). The structure factor for a liquid bears a strong resemblance to its radial distribution function, reflecting the fact that they are Fourier transforms of one another. The basic structural depiction of liquids can therefore be obtained with minimal processing of diffraction data.

Challenges arise from the fact that, unlike its crystalline counterpart, the diffuse scattering pattern is inherently broad and low intensity, making it easily confused with the similarly broad yet unconstrained background pattern. It is therefore only with great difficulty, taking care to minimize background signal using large samples and careful experimental designs, that reliable liquid structure factors are obtained. It is with even greater difficulty that structure factors can be measured for high-temperature, high-pressure liquids like silicate melts (*Funamori et al.*, 2004). It is for this very reason that much of our knowledge of liquid structure relies on insights gained from quenched glasses.

Glasses are strongly preferred to melts for most analytic techniques since they can be measured at ambient conditions on large stable samples. Glasses form as a result of rapid cooling of the liquid, where the cooling timescale is shorter than the kinetic relaxation timescale, forcing the liquid to fall out of equilibrium and thereby freezing its structure. By imposing a range of cooling rates, experimentalists can obtain a set of snapshots of the liquid structure at different temperatures, where the effective temperature at which the liquid structure is frozen into a glass is called the “fictive temperature”. These different cooling rates are achieved using a variety of cooling methods, which include (in order of increasing cooling rate): slow controlled cooling by reducing furnace power over time, air cooling where the melt is removed from the furnace, water cooling where it is submersed in a water-bath, and splat quenching where melt is squirted at high speed onto freezing cold metal. By varying the cooling rate over ten orders of magnitude through such techniques, the fictive temperature can be adjusted by typically of order a few hundred degrees for silicate glasses. Some of these techniques can also be adapted for use in a high-pressure, high-temperature apparatus like a multi-anvil furnace to generate glasses from high-pressure silicate melts. These glasses are then available for conventional measurements, like the diffraction determined Radial Distribution Functions described above.

There is also a broad set of spectroscopic techniques that can been used to probe the local environment of particular atoms in liquids and quenched glasses. These methods rely on a variety of complex interactions between photons of different wavelengths and the nuclei or electron distributions inside the material. Without going into the details, we briefly mention some of the available techniques and the general insights they provide about high-pressure silicate melts. Nuclear magnetic resonance spectroscopy is one of the most successful techniques for probing the local environment around certain target atoms within a material. This term is actually used to describe a wide set of methods that exploit the properties of certain nuclei and their tendency to couple with and react to externally imposed strong magnetic fields. For particular set of atoms, it is possible to observe emitted photons that are characteristic of different properties of the local environment about the nucleus, using a variety of specially designed magnetic fields that may vary in space, time and orientation. For Earth scientists, measurements of silicate materials using ^{17}O-, ^{27}Al-, and ^{25}Mg-NMR have proved particularly useful in studying the evolution of coordination state (number of bonded nearest neighbors) as a function of both pressure and composition. There have been numerous studies over the last two decades that have examined the evolution of oxygen coordination of the Al cation increasing smoothly with increasing pressure (see, e.g., *Lee*, 2011). Typically these studies rely on measurements of quenched glasses rather than in situ high-temperature melts, but it is quite clear that the smooth shift in coordination populations to higher values with increasing pressure is the dominant compression mechanism of most oxide liquids over mantle pressure ranges. In particular, ^{27}Al-NMR studies of Al-silicate glasses are particularly useful since aluminum often substitutes in for silicon within the polymerized network of the silicate structure. By

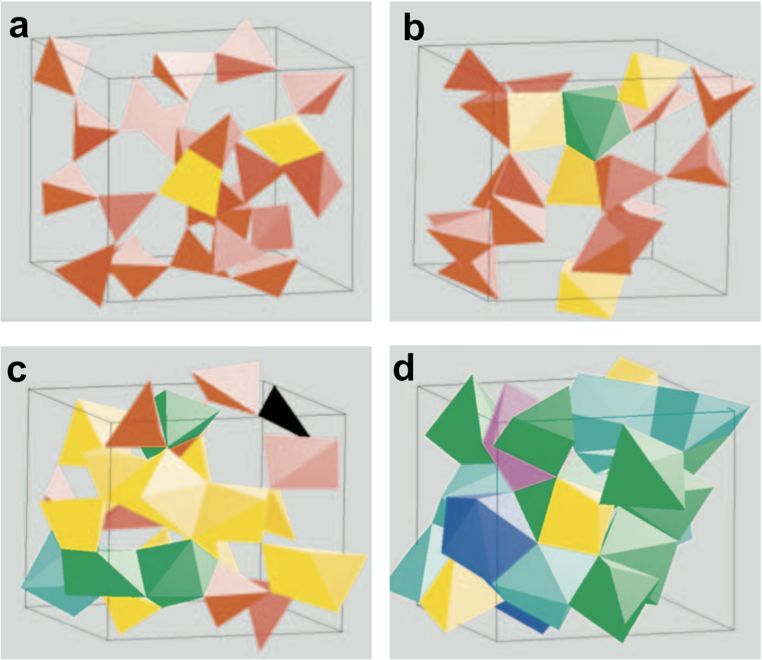

Figure 4.4: Evolution of Si−O coordination polyhedra with increasing compression. Snapshots are shown for a range of volumes, relative to the initial reference volume, (1.0, 0.8, 0.6, 0.4) and a range of temperatures (3000K, 4000K, 5000K, 6000K). Dominantly 4-coordinated Si gives way gradually to 6-coordinated Si with compression. [image credit: *Karki et al.* (2007)]

examining the shift from dominantly 4-coordinated aluminum at ambient conditions up to 5- and 6-coordinated Al, we can gain insights about the general shifts occurring in the liquid's silica framework.

The last primary way in which local liquid structure is probed is through molecular dynamics simulations like those discussed in Section 4.2.2. By turning to FPMD techniques, researchers can obtain reasonable predictions of both liquid structure and thermodynamics at extreme temperatures (above ~ 3000 K) and high pressures that are unattainable in most experimental setups. Such high temperatures pose great difficulties for in situ measurements and quenched glasses are restricted to fictive temperatures with a theoretical maximum at the liquidus ranging down to a few hundred degrees below. Additionally, rapid quenching is made more difficult inside high-pressure apparatuses like multi-anvil presses or diamond anvil cells. Atomic simulation, with its direct knowledge of the exact position of each atom in the fluid, therefore, has a large role to play in the evaluation of liquid structure. The general results from FPMD studies of terrestrial liquids at mantle conditions support the same general conclusions of the diffraction experiments and spectroscopy studies. For visualization purposes, we reproduce a figure from *Karki et al.* (2007) shown in Figure 4.4, which shows the evolution of the different $\mathrm{Si-O}$ coordination polyhedra with increasing compression. As can be seen in the first panel, at a volume V_X corresponding to ambient pressure at 3000K, the Si cations lie predominantly in a tetrahedral 4-coordinated state, and as pressure is increased they gradually shift toward higher coordination numbers, occupying primarily octahedral 6-coordination state at 6000K and $V = 0.4V_X$. It should also be noted that the variety of simultaneously occupied coordination states increases with increasing temperature, as expected from simple Boltzmann-like statistical considerations. In fact, not only is coordination evolution the primary compression mechanism for pure silica, it is found to actually fight against bond length changes, where the nearest-neighbor silicon-oxygen bond length is seen paradoxically to initially *increase* on compression, indicating just how strong the energetic benefit to increasing coordination actually is. Similar coordination behavior has been seen across many different compositions relevant to rock-forming liquids including MgO (*Karki et al.*, 2006), all along the MgO—$\mathrm{SiO_2}$ binary (*de Koker et al.*, 2013), and $\mathrm{Fe_2SiO_4}$ (*Ramo and Stixrude*, 2012).

4.3 Coordinated HArd Sphere Model (CHASM)

Given the combined evidence from diffraction studies, NMR spectroscopy, and FPMD simulations, it appears that oxygen coordination evolution is a common compression mechanism across all high-pressure oxide and silicate melts. Despite this fact, there does not currently exist a simplified theory for oxide melts which accounts for the coordination effects that dominate compression throughout the Earth's mantle. We therefore develop a simplified model to represent high-pressure oxide liquids that explicitly accounts for the range of available coordination states for each cation, and thereby predicts the evolution of coordination with changing pressure, temperature, and composition. The **C**oordinated **HA**rd **S**phere **M**odel (CHASM) provides a general framework for representing the compression behavior of silicate and oxide liquids in terms of evolving the relative populations of different oxygen coordination states for each cation in the liquid, using a perturbation to the hard sphere mixture model. Furthermore, by exploiting the analytic expressions for mixtures of hard spheres from the physics and statistical mechanics literature, we construct a model that captures the basic physical mechanisms while remaining computationally efficient, being easily run on a standard personal computer in minutes, rather than taking months on multi-node supercomputers. Following the classic approach of training empirical potentials, this is a model that relies on parameters that are determined from the comparatively well-understood behavior of solid polymorphs, and can thus be fully predictive for silicate liquids.

4.3.1 A Qualitative Picture of Liquid Compression

To introduce the basic concepts important to constructing such a model, we first examine the essential differences between compression in solids and liquids and how this relates to their respective equations of state. In Figure 4.5, we show a two-dimensional cartoon representation of the compression of materials along a continuum from solid-like to liquid-like behavior. (It should be noted that while this diagram is only meant to illustrate the important concepts, it has been constructed properly using a Monte Carlo hard sphere model to represent liquid particles of different sizes.) Each panel of the figure show two snapshots of the particle structure in the material at particle densities relevant to low- and high-pressure states in light and dark gray, respectively.

Up to this point, we have used a variety of terms, including atom, molecule, or “particle”, to refer to the matter making up the liquid. Generally speaking, a liquid particle should be thought of as the smallest structural unit that can be used to describe the liquid. Oftentimes, these are simply atoms or molecules. In the case of ionic liquids, like the silicates and oxides that are the focus of this study, the particles should be thought of as cation-anion groupings. In order to minimize the Coulombic energy, the large oxygen anions remain grouped around each central cation even in the liquid state, where the number of nearest-neighbor oxygens about each cation is expressed through the coordination number of that cation. As time progresses, the cations and oxygens both move through space, causing both the coordination number and the identity of the coordinated oxygens to evolve with time. Nevertheless, the energy penalties for segregating the oxygens and cations are so high that each cation is always surrounded by an ever-changing local shell of oxygens

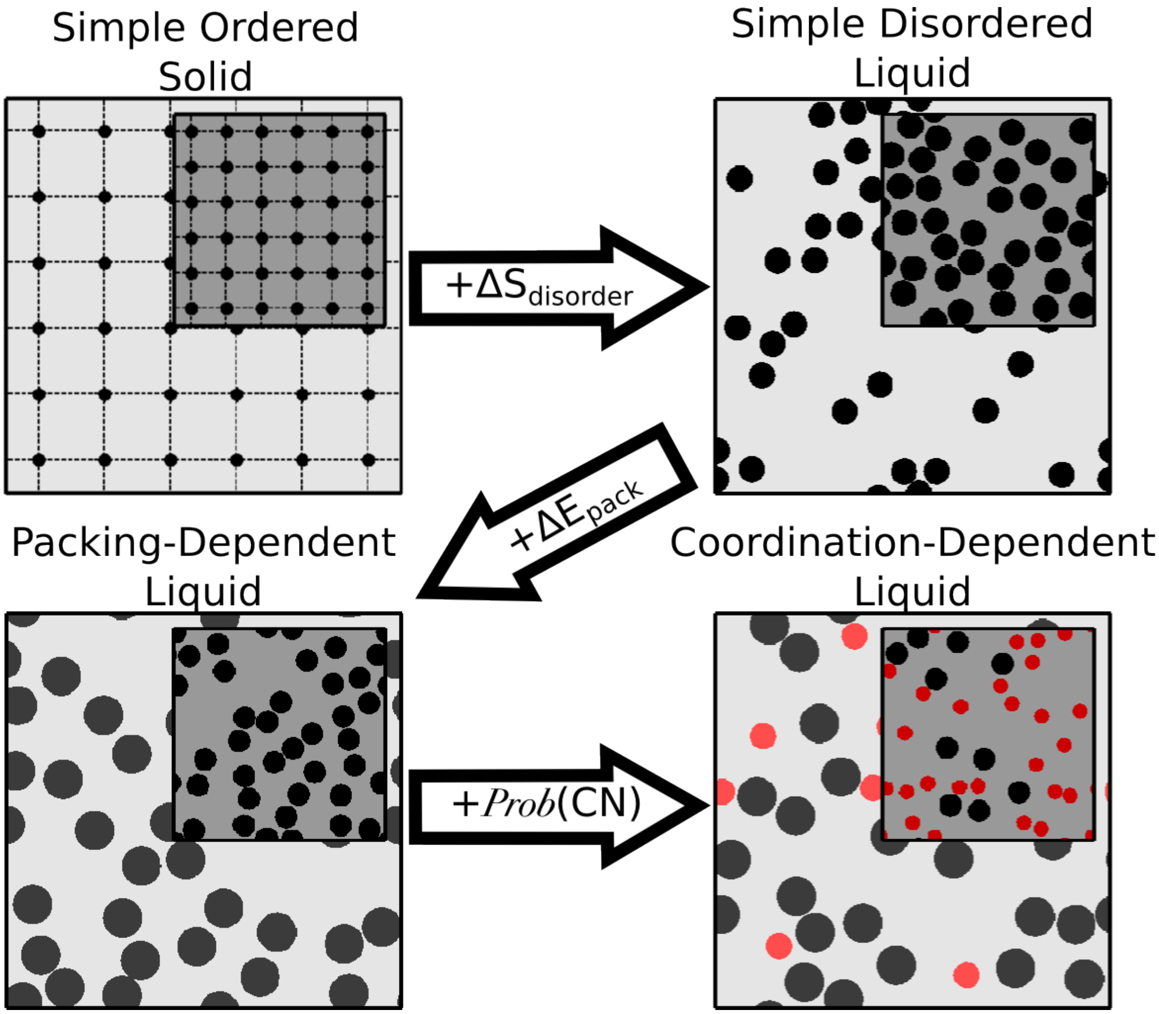

Figure 4.5: Compression mechanism cartoon for oxide liquids showing the different conceptual steps involved in getting from solid-like compression to liquid-like compression. We start with simple ordered solid compression, where all volume changes are taken up by changing bond lengths. The addition of considerable entropy upon melting leads to a simple disordered liquid that can be described with the hard sphere model. Including dependence of the internal energy on packing fraction results in a perturbed hard sphere model, where the sphere sizes adjust with compression. Finally, allowing for varying coordination states with populations that depend on packing fraction yields a reasonable picture of a coordination depend liquid. In the final panel, high and low coordination are shown in red and black, respectively, and the liquid adjust the coordination populations toward higher coordination states with increased compression.

and can be usefully described as evolving through a series of coordination states. (In reality, the oxygens also undergo some coordination evolution, increasingly being shared by 3 instead of only 2 cations, but this has a less major impact on the overall liquid thermodynamics.) In this way we can adapt the hard sphere model to ionic liquids by thinking of the spheres as representative of coordinated ionic complexes, where each is centered on a single cation. The interactions of each sphere with its neighbors is therefore governed primarily by its coordination state, which affects both the size of the complex and the way in which it is connected to its neighbors through the background field of oxygens that are implicitly present and surrounding each cation.

Keeping this understanding of liquid particles in mind, we can look to Figure 4.5 to gain an intuitive grasp of the mechanisms that operate in liquids to accommodate compression and how they differ from compression in solids. In the upper left of the figure, we start with the orderly compression of solids, which is dominated by changes in bond length. With atoms fixed in their relative positions within the unit cell, compression in solids is marked by only small changes in disorder associated with the decreasing volume available for thermal atomic vibrations. Upon melting, materials gain a significant bump in entropy, associated with the loss of long-range order. As discussed above, liquids still maintain short-range order, which can be approximately represented using the basic hard sphere model. This simple disordered liquid is depicted in the second panel of Figure 4.5, where the changes in the liquid structure are controlled by the packing fraction. It is important to note that in this case, the hard sphere radius is assumed to be fixed and independent of both temperature and compression. A simple hard sphere model such as this is the type utilized by *Jing and Karato* (2011) to describe low-pressure silicate liquids.

A further refinement to the physical picture can be made by including the feedback effect of compression on the excluded volume. In the third panel, we show how a more realistic liquid picture allows the sizes of the hard spheres particles change as a function of pressure (or temperature). This can be achieved in one of two major ways: most simplistic is the application of empirical rules for how the hard sphere sizes compress with increasing pressure or temperature. This class of method was extremely popular for many years due to its computational simplicity and resulted in many applications of the hard sphere model to simple liquids, and is also the approach taken by *Jing and Karato* (2011) to improve the accuracy of extensions to elevated pressures and temperatures. While simple to implement, this empirical method suffers from its lack of strong physical grounding, where the rules governing the compressibility of the hard spheres are fairly arbitrary and chosen only to match the generally observed trends in liquid behavior. They therefore rely on fitting liquid data and have only limited capability for extrapolation and interpolation in pressure–temperature–composition space.

It is strongly preferable to use a physically based perturbation method, which uses hard spheres as the reference system to which higher-order corrections are made. The physical phenomenon that slightly compressible hard spheres seek to represent is the gradual decrease of the minimum particle spacing experienced by liquid particles with compression. This behavior reflects the limits of the hard sphere model, which approximates the steep Fermi repulsive energy barrier for particles at small separation distances with a step-function, or infinitely sloped hard wall. Of course, such a rough approximation was only ever appropriate over a limited range of conditions. As temperature increases, the liquid particles gain thermal energy allowing them to “climb” further up the repulsive slope, shrinking the effective hard sphere size. Similarly, increasing pressure confines liquid particles into ever smaller regions, balancing the energy penalty associated with climbing the repulsive barrier as particle sizes shrink against the energy benefit resulting from the increased free volume available to each particle. For this reason, it is energetically advantageous to steadily reduce the hard sphere size of each particle with compression, as depicted by the shrinking hard spheres in the third panel of Figure 4.5. By altering particle size, the liquid is able to maintain large configurational entropies in spite of decreasing volume, which helps to offset its increasing internal energy. This trade-off between the internal energy (associated with the packing of liquid particles) and the free energy (associated with the dependence of configurational entropy on free volume) is at the core of the “variational method” for hard spheres, which seeks an optimum hard sphere size through total free energy minimization. We will discuss the details of this method below, but for now, it is important to recognize how it relates to the evolving structure and energetics of the liquids upon compression and heating.

As indicated in the previous section, the final and most important aspect of oxide melt structure for the purposes of understanding compression is the availability of a variety of different oxygen coordination states for each cation. In the last panel of Figure 4.5, we show a snapshot of a simple liquid that has two different coordination states available to it, a large low coordination state in black and a compact high coordination state in red. The liquid is free to populate these states in whatever manner minimizes the total free energy.

This results in low coordinations dominating at low pressures and high coordinations at high pressure, as depicted in the figure by the increased presence of red particles in the high-pressure state. Both the populations of coordination states, shown as red and black in the figure, and the effective particle sizes for each state respond to changes in pressure and temperature, seeking to minimize the total free energy.

At this point, we must briefly digress to discuss the issue of coordination number. Cations in crystalline solids exist in a small finite number of coordination environments depending on the crystal structure. Typically, these coordination states are thought of as being restricted to integer values, corresponding to the number of oxygens bonded to the cation. This is a sensible notion for "ideal" structures, in which all nearest-neighbor oxygens lie at equal distance from the central cation. Many crystal structures, however, contain cation sites that are highly distorted, where the nearest-neighbor oxygens lie at a discrete set of unequal distances, making it more difficult to define which oxygens are directly bonded to the cation and which are shared unequally with neighboring cations. This complication reveals the need to define the "effective coordination number", which gives higher coordination weight to nearer neighbors and less to more distant ones. While there are many possible definitions of effective coordination number, we use that put forward by *van Meel et al.* (2012), since it is straightforward, using a sensible solid-angle-based weighting scheme, and does a good job reflecting the differences between different distorted crystal structures. [2]

---

[2] A more commonly used definition was introduced by *Hoppe* (1979), who used an exponential averaging scheme to give greater weight to closer neighbors. *Hoppe* (1979) fully admitted that the form of the weighting scheme is fairly arbitrary, and can only be judged on its ability to recover the correct values for ideal structures while providing useful insight into more distorted ones. Accordingly, we have tested this scheme against the more recent one by *van Meel et al.* (2012), and have found *van Meel et al.* (2012) to return more sensible values. Additionally, this new method results in coordination weights that are constrained to values less than

Like distorted solids, liquids also place cations inside geometrically non-ideal sites. While we depicted only two coordination states in Figure 4.5 for the purposes of clarity, liquids actually contains a spectrum of local coordination environments. Released from the constraints of periodicity, the distances and angles between cations and oxygens in liquids are dynamically evolving, causing the nearest-neighbor oxygens to sample a *continuous* distribution of bond distances, further emphasizing the need to consider fractional coordination states. While the concept of fractional coordination may be unfamiliar to many, it is crucial to understanding the structure and resulting thermodynamics of silicate and oxide melts.

Using this qualitative picture of liquid structure and thermodynamics as our guide, we are now prepared to consider the development of a quantitative model that captures these important behaviors.

4.3.2 Perturbing the Hard Sphere Mixture Model

Unlike dynamical models, such as first-principles molecular dynamics, which seek to directly simulate the motions of particles in a system, statistical mechanical models only attempt to represent the behaviors of a system probabilistically. The advantages to this alternative approach are that it produces models that are *many* orders of magnitude faster—the difference between desktop computers and super-computers, minutes and months of calculation time—and it also leads to models that are inherently scale independent, and so they do not suffer from simulation size or evolution time effects.

---

or equal to one for each neighboring particle, whereas the method of *Hoppe* (1979) frequently gives weights considerably greater than one for close-by neighbors. We find this to be fairly unphysical, since an atom cannot be said to have a bonding weight of more than 100%.

At the root of statistical mechanics is the recognition that each macroscopic state of the system, described by the state variables of volume, temperature, and composition, actually corresponds to a huge number of combined microscopic states. A macroscopic system (with of order at least Avogadro's number $\sim 10^{24}$ of particles) can only be completely described by its microscopic configuration, a list of the position and velocity vectors for every particle at every time step. Clearly, a "macroscopic dynamical model" would require an impossible amount of computing power. Instead, statistical mechanical models consider the range of available microstates, and assign relative populations to these states according to the rules of thermodynamics. For a closed system at constant temperature T, volume V, and composition $\vec{X}$, the second law of thermodynamics states that the system will achieve equilibrium when its Helmholtz free energy F is minimized, which is given by the sum of independent contributions $F = \sum F_i = \sum (E_i - TS_i)$, where E is the internal energy and S is the entropy. Therefore, the microstates of the system are populated in such a way as to minimize the total free energy. [3] This provides us with a simple recipe for constructing a statistical mechanical model, where we need only write down an expression for the free energy and minimize it.

Using the hard sphere mixture model as a guide, we can think of silicate and oxide melts as a collection of coordination-specific hard spheres, where each sphere represents a particular oxygen bonding environment about a cation lying at its center. By explicitly considering the range of available coordination states, we are able to capture how the contribution of each cation to the overall liquid structure depends on the oxygen coordination

---

[3] The simplest example of this is the familiar Boltzmann distribution, which describes the population of a set of independent states with different energy levels.

number, which controls the local connectivity between the cations. Given this physical picture, we have developed the **C**oordinated **HA**rd **S**phere **M**odel (CHASM), which applies a perturbation to the standard hard sphere mixture by both accounting for multiple coordination states for each oxide component and estimating how the average liquid structure affects the internal energy of the liquid. The total free energy of the CHASM liquid can thus be divided into its major components:

$$F_{\mathrm{Tot}} \approx F_{\mathrm{ideal}} + \Delta F_{\mathrm{HS}} + E_{\mathrm{struc}} + F_{\mathrm{internal}} \tag{4.6}$$

where F_{ideal} is the ideal gas contribution, ΔF_{HS} is the configurational free energy contribution as calculated by the hard sphere mixture model, E_{struc} is the dependence of the internal energy on the average liquid structure, and F_{internal} captures the additional free energy contribution caused by the internal degrees of freedom available to each hard sphere particle, primarily associated with the freedom of movement of the bonded oxygens. Given the dependencies of each term, the total free energy can be written:

$$F_{\mathrm{Tot}} = F_{\mathrm{Tot}}(V, T, \vec{X}, \mathbf{p_{cn}}, \mathbf{d_{cn}}) \tag{4.7}$$

where the free energy is a function of the thermodynamic state variables $(V, T, \vec{X})$, as well as the model parameters describing the hard sphere attributes $(\mathbf{d_{cn}}, \mathbf{p_{cn}})$, which are the diameters and populations of the hard spheres as a function of coordination number for each oxide component of the liquid.[4] We now give a brief description of each term in

---

[4] The coordination populations are defined separately for each component such that $\mathbf{p_{cn}}_{ij}$ gives the fraction of cations of type i that are in coordination state j and therefore $\sum_j \mathbf{p_{cn}}_{ij} = 1$ for each component.

the total free energy expression. While, for the purpose of clarity, we save most of the details for the Appendix, it is important to recognize that all terms can be evaluated using a combination of straightforward analytic expressions and simple 1-D numerical integrals.

The first term in the free energy captures the contribution associated with a mixture of ideal gases. This captures the momentum carried by each particle, given by the standard expression for a single component ideal gas $F_{\text{id-gas}}$, combined with the ideal entropy of mixing for completely randomized particles $S_{\text{id-mix}}$:

$$\begin{aligned} F_{\text{ideal}} &= \sum_i X_i F_{\text{id-gas}}(V, T, m_i) - T S_{\text{id-mix}} \\ S_{\text{id-mix}} &= -k_B \sum_{i,j} x_{ij} \log x_{ij} \qquad (\text{where} \quad x_{ij} \equiv X_i \mathbf{p_{cn}}_{ij}) \end{aligned} \tag{4.8}$$

Note that these contributions are evaluated for the hard sphere particles rather than the individual atoms, and thus m_i and X_i are the mass and mole fraction of each oxide component (in a per cation basis). The mole fraction of oxide component i in the j^{th} coordination state is given by x_{ij}, which is equal to the relative population of that coordination state times the absolute abundance of that component. The occupation of multiple local coordination states leads to a significant increase in the entropy of mixing for liquids, reflecting the increase in the number of distinguishable randomized microstates. This fact is well known, and was found to adequately explain the high entropy of mixing observed across the MgO-SiO_2 binary in FPMD calculations (*de Koker et al.*, 2013).

The next term in the free energy is the adjustment resulting from a mixture of hard spheres. This analytic equation was derived using an approximate solution for the radial distribution function, referred to as the Percus-Yevick equation, and was first extended to hard sphere mixtures by *Lebowitz* (1964). Subsequent work improved the approximation (*Mansoori et al.*, 1971), yielding an expression that can be thought of as a standard black box for calculating the free energy of hard sphere mixtures. While this expression is shown in its entirety in Appendix 4.A, we note here that it takes the form:

$$\Delta F_{\text{HS}} = kT \times f_{\text{HS}}(V, \vec{X}, \mathbf{p_{cn}}, \mathbf{d_{cn}}) \tag{4.9}$$

where the free energy difference is a purely geometric function of the volume, hard sphere sizes and populations. The dependence on temperature is only as a multiplicative prefactor, in effect setting the overall energy scale. Together, the ideal and hard sphere terms represent the pieces of the standard hard sphere mixture model.

While generally useful, the standard hard sphere model lacks any contributions to the internal energy, which are crucial for describing large degrees of compression. As presented in Figure 4.5 and in the subsequent discussion, there exists a trade-off between the configurational free energy (captured by the hard sphere model) and the internal energy cost associated with the average liquid structure. This balance enables the particles to adjust their effective sizes to respond to changes in volume or temperature in order to maintain the minimum free energy. We therefore need an expression for the energy cost associated with a particular liquid structure. The simplest possible approach is to approximate the total structural energy as resulting from the sum over pairwise interactions—through an effective "pair potential" E_{pair}—of every particle with every other particle. This is achieved by considering interactions between a central reference particle of type c and the surrounding particles of type s:

$$E^{cs}_{\text{struc}} = \frac{1}{2} x_c \int_0^\infty E^{cs}_{\text{pair}}(r) \operatorname{Pr}^{cs}(r) dr \qquad \text{and} \qquad E_{\text{struc}} = \sum_{c,s} E^{cs}_{\text{struc}} \tag{4.10}$$

where the total energy is given by integrating the probability-weighted energy cost ($\text{Pr} \times E_{\text{pair}}$), which depends on the separation distance r, over all possible separation distances. Additionally, x_c, the mole fraction of c particles, gives the probability of choosing a c particle at random, and the factor of $\frac{1}{2}$ exists to prevent double counting (since each particle is counted once as the central particle, c, and once as the surrounding particle, s). This expression is then evaluated for every possible pairing of distinguishable particles and summed over, cycling through each coordination state for every component.

The probability of finding particles at each distance is tied directly to the structure of the liquid through the radial distribution function $g(r)$, and thus we can rewrite the above equation:

$$\begin{aligned} E^{cs}_{\text{struc}} =& \frac{1}{2} V \rho_c \int_0^\infty E^{cs}_{\text{pair}}(r) 4\pi r^2 \rho_s g_{cs}(r) dr \\ =& 2\pi V \rho_c \rho_s \int_0^\infty r^2 E^{cs}_{\text{pair}}(r) g_{cs}(r) dr \end{aligned} \tag{4.11}$$

where the number densities of the central and surrounding particles are given by ρ_c and ρ_s. The key point is that the approximate RDF for the liquid is a direct output of the hard sphere mixture model, and therefore we need only impose a reasonable pair potential in order to calculate the structural perturbation. Determining the empirical pair potentials for the coordinated hard spheres is therefore at the heart of applying the CHASM method and requires significant attention. We defer that discussion until Section 4.3.3, where we detail how to determine liquid-appropriate pair potentials solely from data on the solid structures. For the meantime, we assume that such a determination has been made and that Equations (4.10) and (4.11) can be evaluated.

Finally we come to the last term in the free energy expression, which captures the effects of the internal degrees of freedom of each hard sphere particle. Thus far, every term has relied entirely on thinking of the hard spheres as indivisible particles—at this point we consider corrections that account for the limitations of this description. Each particle is composed of a central cation surrounded by a set of coordinated oxygens that are shared between neighboring particles. The oxide particle picture reflects how much cations and oxygens resist separating from one another, but it ignores the freedom of independent motion that the oxygens retain. We thus split the internal free energy contribution into its two major parts:

$$F_{\text{internal}} \approx F_{\text{bond ang}} + F_{\text{bond len}} = F_{\text{internal}}(V, \vec{X}, \mathbf{p_{cn}}) \tag{4.12}$$

where $F_{\text{bond ang}}$ expresses the free energy resulting from the positional freedom that is coupled with bond-angle distortion, and $F_{\text{bond len}}$ expresses the contribution due to the variability in bond length about the average value.

We have derived an expression for the first term in Equation (4.12), using elementary quantum mechanics, but to avoid distraction, we relegate that derivation to Appendix 4.D. The results show that the bond-angle freedom term relates to the cation-oxygen bond length, which is unsurprising since the oxygens can be thought of as lying on the surface of a "bonding sphere" with radius equal to the bond length. Most of the extra dynamics relate to motions of the oxygens along the surface of this sphere. There must, of course, be some motion of the oxygens along the radial direction causing bond length distortions, as represented by the second term in Equation 4.12, though we currently neglect any contributions from motion of this type. As we show in the next section, we can empirically relate oxygen bond length for a particular cation to the overall volume by examining a variety of solid structures. Additionally, the coordination number directly controls how many oxygens lie on the surface of the bonding sphere, determining the typical bond angles, and thus the coordination populations are also important to calculating the contribution from internal degrees of freedom. Notice, however, that there is no direct dependence upon either hard sphere sizes or temperature, as indicated by the right-hand side of Equation (4.12).

By combining the equations above with those in Appendix 4.A-4.D, we now have the ability to evaluate the total CHASM free energy. This relies specifically on providing the coordinated hard sphere pair potentials, which is discussed in detail in the next section. As alluded to previously, the final step is to apply the principle of energy minimization:

$$F_{\mathrm{liq}}(V,T,\vec{X}) = \min_{\mathbf{p_{cn}},\mathbf{d_{cn}}} \{F_{\mathrm{Tot}}\} \tag{4.13}$$

where the state of the system is that which minimizes the total free energy. Thus, for a particular set of state variables V, T, and $\vec{X}$, we can obtain self-consistent thermodynamically derived values by minimizing the total free energy with respect to the hard sphere parameters. This generic method of determining the model-dependent parameters is also called the "variational method" for hard spheres, due to its connections to the calculus of variations[5], and was developed by *Mansoori and Canfield* (1969). In that and subsequent works, the authors developed improvements to the description of compositional hard sphere mixtures, but remained limited to any considering the added effect of liquid structure. We have thus extended the same technique to directly tracking the evolution of coordination state. To simplify and speed up calculations, the dependence of the hard sphere sizes and populations on coordination number can be parameterized with simple smooth functions, thereby enforcing the expected systematic behavior while also reducing the number of parameters that must be varied (details are presented in Section 4.3.4). Because all of the expressions involved in evaluating the total free energy are either analytic or simple 1D numerical integrals, evaluating the free energy is a trivially fast operation, taking less than a second on a modern personal computer. We are therefore able to use a standard numerical minimization routine in MATLAB, allowing the solution for each state point to be calculated in only a few minutes, in contrast to the weeks or months necessary on multiprocessor machines in dynamical simulations.

4.3.3 Training Pair Potentials on Solid Structures

Now that we have presented the general framework of the CHASM approach, we will discuss the details of how to obtain reasonable values for the pair potentials by training on the equations of state of solid polymorphs. This step is critical, since the pair potentials represent the primary input to this model, determining the structural energy perturbation

---

[5] The liquid free energy in equilibrium can be thought of as a functional of the total free energy, which is subject to a minimization constraint. We imagine using a series of trial expressions for the free energy, and vary the hard sphere parameters until the minimum energy state is achieved.

that is at the heart of the variational method. Additionally, we must also determine the relationship between oxygen bond length and volume, which is provided by the empirical trends apparent in the solid crystal structure data. While bond lengths are not an explicit part of the general CHASM framework—since oxygens are only represented in terms of coordination number—they are required to estimate the bond-angle correction to the free energy.

As discussed at length in the previous section, liquids populate a continuous spectrum of local atomic structures, while crystalline solids are restricted to the small subset that are both periodic and minimal energy configurations. Though the range of available options is much narrower, differently structured crystals of constant composition (called crystal polymorphs) can be thought of as discretely sampling the same continuum that is available to the less highly-constrained liquids. This implies that we can gain considerable insight into the energetics of particular atomic arrangements by studying the solid structures. Ideally, we would have a computational tool that could rapidly determine the internal energy of any arbitrary ionic arrangement, but such a tool would need to solve Schrödinger's equation and therefore require first-principles methods like density functional theory, which are inherently both iterative and extremely computationally intensive. Instead, we look to the well-studied energetics of the solid structures to develop a set of empirical pair potentials that can approximately capture the dependence of internal energy on atomic configuration. When applied to the atoms themselves, this strategy represents a fairly standard approach for determining potential functions that can be used in empirical molecular dynamics simulations. Though we are not performing MD simulations, we do benefit from the general knowledge such studies provide about the features and limitations of pair potentials.

Empirical potentials represent the internal energy as a sum over either pairwise or many-body particle interactions using a suitably flexible and simple functional form. The most well-known example is the Lennard-Jones potential, which is appropriate when weak van der Waals attractive forces dominate at large separation distances: $E_{LJ}(r) = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$, where ϵ and σ are the energy and distance scales. A more flexible potential form is given by the generalized Morse potential, which represents the repulsive and attractive parts using exponential decay functions:

$$E_{\text{exp-exp}}(r) \approx \epsilon_R \exp\left(-\frac{r}{\sigma_R}\right) - \epsilon_A \exp\left(-\frac{r}{\sigma_A}\right) \tag{4.14}$$

where ϵ_R & ϵ_A are the energy scales and σ_R & σ_A are the length scales for repulsion and attraction, respectively. In order to remain physical, with an asymmetric energy well that has a steep repulsive barrier at short distances, the parameters must take values with $\epsilon_R >> \epsilon_A$ and $\sigma_R < \sigma_A$. When applied to ionic particles, it is also necessary to add a Coulombic term to the interaction energy. While functions of this type are loosely motivated by physical considerations, they can only be expected to capture the observed energetic variation over a limited range of conditions. Accordingly, it is common to develop pair potentials that are appropriate for small deviations from a particular configurational environment, such as atomic vibrations within a particular crystal structure.

In order to extend the suitability of the potentials, it is usually necessary to add more and more corrective terms, which add angular dependence for dipole fields, impose bond-angle restrictions, or account for 3-body interactions, significantly reducing the simplicity of the approach and increasing the care required for obtaining reasonable parameter values. The most common downside to using pairwise-only atomic potentials, like Equation (4.14), arises from the fact that such expressions are forced to represent total bonding energies as strictly proportional to coordination number, since each bond contributes the same energetic weight. This tends to impose an unrealistic preference for high coordination numbers, since the total energy can be lowered by surrounding each cation with as many oxygens as possible. In reality, there is an increasing cost to placing more particles in a confined area, caused by the increase in local electron density (*Daw et al.*, 1993). Each subsequent particle actually sees a reduced energy benefit for bonding, compelling the ions to adopt less extreme coordination numbers.[6]

It is important to recognize that our goal is quite different from MD studies—we are not seeking to directly model the individual positions of cations and oxygens. Instead, we need only capture the average relative placement of effective hard sphere particles, which represent the basic structural units from which each liquid or crystal configuration is built. This simplification affords significant advantages over the standard pair potential approach by sidestepping the issue of suitability. Rather than developing complicated atomic potentials that are capable of representing vastly different local configurations, we instead restrict ourselves to only representing the interactions of hard sphere particles that are inherently coordination specific. By imposing the constraint that the pair potential values must be smooth and slowly varying functions of coordination number, we ensure that the

---

[6] The embedded atom method is a very successful many-body potential theory which accounts for this effect when directional bonding remains unimportant (*Daw and Baskes*, 1984; *Daw et al.*, 1993), though there is a significant added cost in terms of both complexity and the number of additional parameters whose values must be determined in order to apply the technique.

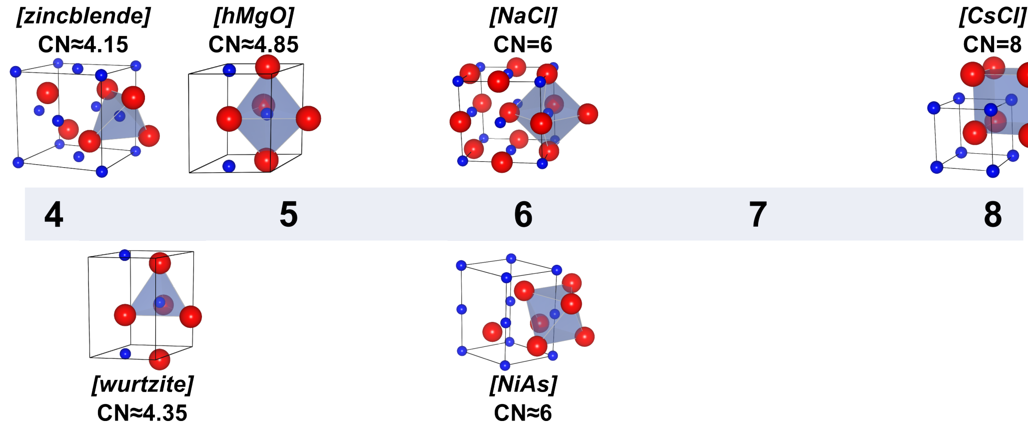

Figure 4.6: Representation of the unit cell structures and oxygen coordination polyhedra for the six theoretically realizable polymorphs of MgO. Each structure is displayed along an axis representing the effective coordination number. In reality, only the NaCl structure (periclase) is stable under terrestrial conditions. Nevertheless, the other structures are theoretically possible and can provide great insight into the energetics of oxygen-coordinated Mg.

resulting potentials are well behaved and improve the interpolation behavior of the derived potentials. This is particularly important since the solid polymorphs only sample a small set of coordination states and thus reasonable interpolation properties are an absolute requirement of potentials that are used for liquid calculations. Most importantly, as the local coordination is a prescribed attribute of each particle rather than a derived quantity based on model output, we no longer suffer from the common issue of improperly favoring high oxygen coordination states. Additionally, since each particle is neutrally charged, there is no need for additional Coulomb terms. By specifying pair potentials for the effective particles, rather than the atoms themselves, we are able to retain a simple functional form, as in Equation (4.14), without sacrificing the flexibility necessary to represent the changing nature of atomic interactions that results from evolving coordination states.

Having motivated and developed a general strategy for representing hard sphere potentials, we now demonstrate how to train the potential parameter values using equation of state data for the MgO polymorphs. MgO is chosen because it represents the simplest possible oxide and yet is extremely important to our understanding of the Earth. In the pressure range experienced in the silicate Earth, pure MgO exists only as the mineral periclase, which has the B1 or NaCl "rock-salt" structure, with an oxygen coordination number of 6. Depending on the packing efficiency of the unit cell, MgO could potentially adopt a range of structures, which are shown in Figure 4.6 ordered along an effective oxygen coordination line. This figure indicates the range of periodic structural options available MgO and other "AB" crystals, including CsCl, ZnS, ZnO, CaO, and BN. Whether a particular chemical system will actually adopt a given structure depends on the energetics of that configuration as dictated by quantum mechanics. So stable is the NaCl structure for MgO that density functional theory calculations have placed the pressure-induced phase transition to the denser CsCl (or B2) structure at somewhere between 400 and 500 GPa (*Karki*, 1997; *Jaffe et al.*, 2000; *Joshi et al.*, 2012). Looking to other analog AB crystal systems, we know that MgO should be able to theoretically adopt the low coordination zinc-blende (B3) and wurtzite (B4) structures, though at unrealizable "negative pressures". Though no structures other than NaCl have ever been observed for MgO in the lab, there are numerous studies that have examined both the high and low (negative) pressure polymorphs using density functional theory (see, for example, *Karki et al.* (2007); *Joshi et al.* (2012); *Jaffe et al.* (2000); *Oganov and Dorogokupets* (2003)). These studies provide energy-volume curves, which when coupled with known atomic positions, can form the basis of a training dataset for simple coordinated pair potentials.

For training our MgO pair potentials, we rely on the dataset produced by *Joshi et al.*

(2012), as it examined all six of the available structures shown in Figure 4.6. The fact that all energy curves were calculated self-consistently within a single study is important, since each study necessarily introduces an arbitrary energy offset parameter By using a self-consistent training set, we need only allow for a single global offset term, rather than one for each study which could amplify degeneracies in the fitted parameters. The calculated equations of state from *Joshi et al.* (2012) for the six theoretical structures of MgO are shown in Figure 4.8. In addition to the four previously mentioned structures, these also include two non-ideal distorted structures: the NiAs structure is six coordinated like periclase, but has a highly inefficient packing scheme for the cations resulting in large distortions of the nearest-neighbor cation distances, while the hMgO structure is a buckled boro-nitride variant that provides nearly 5-coordination for the cations. Together, these six structures with their computed effective coordination numbers and relative cation positions provide sufficient information to train the coordination-dependent pair potentials.

In order to obtain a simple parameterization that describes the total internal energy of each atomic configuration in terms of the pairwise interactions of cations alone, we must first select a form for the pair potential. The attributes of the pair potential are assumed to vary systematically with coordination number, thereby providing a simple way to capture how local bonding environment in turn affects the energy. For both its familiarity and flexibility, we choose the generalized Morse potential, which can be described most simply as the difference of two exponential decay functions (see Equation 4.14). While this parameterization has the simplest mathematical form, it is not very useful for fitting. The potential

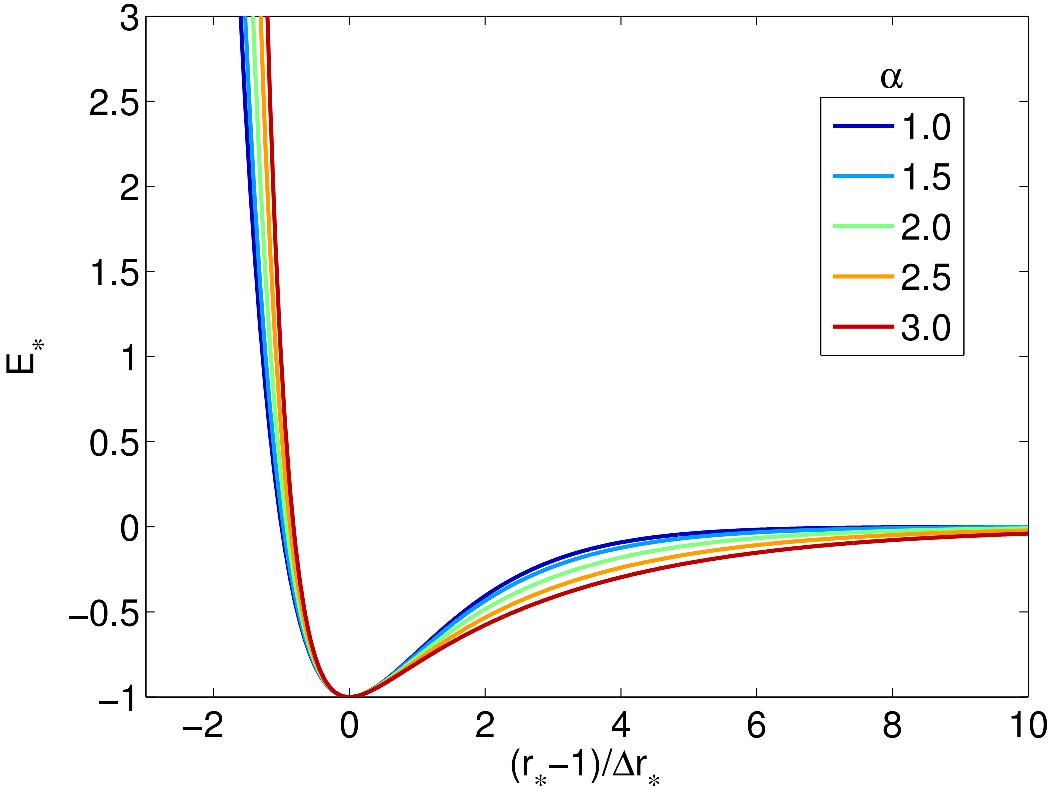

Figure 4.7: Normalized plot of the Generalized Morse Potential. The axes are chosen to force all wells to overplot one another. The only remaining degree of freedom is the shape parameter α. We choose to fix its value to $\alpha = \sqrt{2}$, which corresponds to the standard Morse Potential form, in order to prevent over-fitting.

is thus recast into the following "normalized Morse potential" form, depicted in Figure 4.7:

$$\begin{aligned} E^*_{\text{pair}}(r^*) =& \frac{1}{\alpha^2 - 1} \left\{ \exp\left[-\alpha \frac{(r^* - 1)}{\Delta r^*} \right] - \alpha^2 \exp\left[-\frac{1}{\alpha} \frac{(r^* - 1)}{\Delta r^*} \right] \right\} \\ &\text{with,} \quad E^* = E/E_D \quad \text{and} \quad r^* = r/r_e \end{aligned} \tag{4.15}$$

where the energy and length scales defining the depth and position of the energy well are given by E_D and r_e, and the relative width of the energy well is defined by Δr^*. Together, these 3 parameters scale the position and dimensions of the energy well, as depicted in Figure 4.7. Lastly, the overall shape of the well is given by the value of α, which primarily controls the curvature of the distant half of the well as shown in Figure 4.7. This parameterization, equivalent to Equation 4.14 with 4 free parameters, can be reduced to the standard Morse potential by fixing $\alpha = \sqrt{2}$. While fixing α removes flexibility from the function, we employ this strategy to avoid overfitting since the six compression curves do not sample a sufficient variety of cation distances to independently constrain the precise well shape.

The fitting procedure itself is fairly straightforward. First the atomic positions for the periodic crystals are calculated up to some large cutoff radius, which is many times larger than the nearest-neighbor cation distance. Then the total pairwise interaction energy is determined by evaluating Equation 4.15 at each separation distance and summing over all cation pairs. Finally, we also include a "tail correction", which roughly accounts for the effect of truncating the sum at a finite cutoff radius. The form of this correction term is exactly the same as Equation 4.11, where beyond the cutoff radius, we neglect the effect of small fluctuations in particle density setting $g_{cs}(r > r_{\text{cut}}) = 1$. This results in a trivial exponential integral for the Morse potential that is performed analytically.[7]

In addition to the direct interaction energy reflecting the positions of each cation, we must introduce a correction to account for the strong distortions present in some of the structures like NiAs and hMgO. This correction counteracts the excessive repulsive energies predicted by the pair potential in the case of distorted closely-spaced cations. The pair potentials are designed to capture the general energetic dependence of coordinated cations, and can therefore be thought of as describing the energy cost of "ideal" atomic arrangements. Distorted structures, however, have the freedom to rotate oxygen positions to move them out of the way and make room for the closely spaced cations, thereby reducing the effective energy cost of such an arrangement. This effect depends almost exclusively on the separation distances of the nearest-neighbor cation pairs, which induce the largest repulsive energies. To account for this difference, we apply a simple additive corrective term based on the concept of the effective coordination number:

$$\Delta E_{\text{corr}} = -\sum_i w_i (E_{\text{pair}}(d_i) - E_{\text{pair}}(d_{\text{eff}})) \tag{4.16}$$

where w_i is the coordination weight of the i^{th} nearest-neighbor cation and d_{eff} is the effective cation separation distance considering all neighboring cations, both calculated using the method of *van Meel et al.* (2012). Applying this expression to each cation, we can imagine that we are determining the energy impact of moving all of the nearest-neighbor cations from their actual positions to their average effective separation distance, which are more reflective of the energetic cost of the atomic arrangement. This corrective term is only

---

[7] The inclusion of this tail correction causes the total pairwise energy to converge much more rapidly, enabling smaller values of the cutoff radius.

important for distorted structures like NiAs, where it is large and negative. For ideal atomic arrangements, the actual and effective distances are equal, and thus the correction term is zero.

Using the calculation outlined above, we can calculate a model energy compression curve for any chosen set of pair potential parameters r_e, E_D, and Δr^* (where $\alpha = \sqrt{2}$ is fixed in this case to avoid overfitting). Since each parameter represents a scale factor that could potentially vary over orders of magnitude and is constrained to be positive, we use $\log r_e$, $\log E_D$ and $\log \Delta r^*$ as our fitting parameters. Each compression curve is first fit individually. Since we seek a model that merely captures the general behavior of coordinated cations, it is not to important assign error bars to each measurement as we know that inadequacies of the model, not errors in the DFT data, will dominate the residuals. Therefore we apply the standard method of unweighted least-squares regression. In this case, the goodness of fit is described in terms of the explained fraction of the variance or $R^2 = 1 - \mathrm{Var}(E_{\text{resid}})/\mathrm{Var}(E_{\text{data}})$, with a value of 1 corresponding to a "perfect" fit. After fitting each curve individually, we construct a global model that enforces smooth dependence on coordination by asserting that each potential parameters is described by a low-order polynomial in coordination number. To get an initial guess to the final global model, we fit the individual best-fit parameters with a quadratic, after which the entire collection of energy-volume curves are fit jointly. The final results of this global fit are shown in Figures 4.8 and 4.9. As seen in Figure 4.8, the general quality of the fit reasonably good with $R^2 = 0.986$. The largest residuals belong to the two highly distorted structures, NiAs and hMgO, where the corrective term given by Equation 4.16 is able to account for

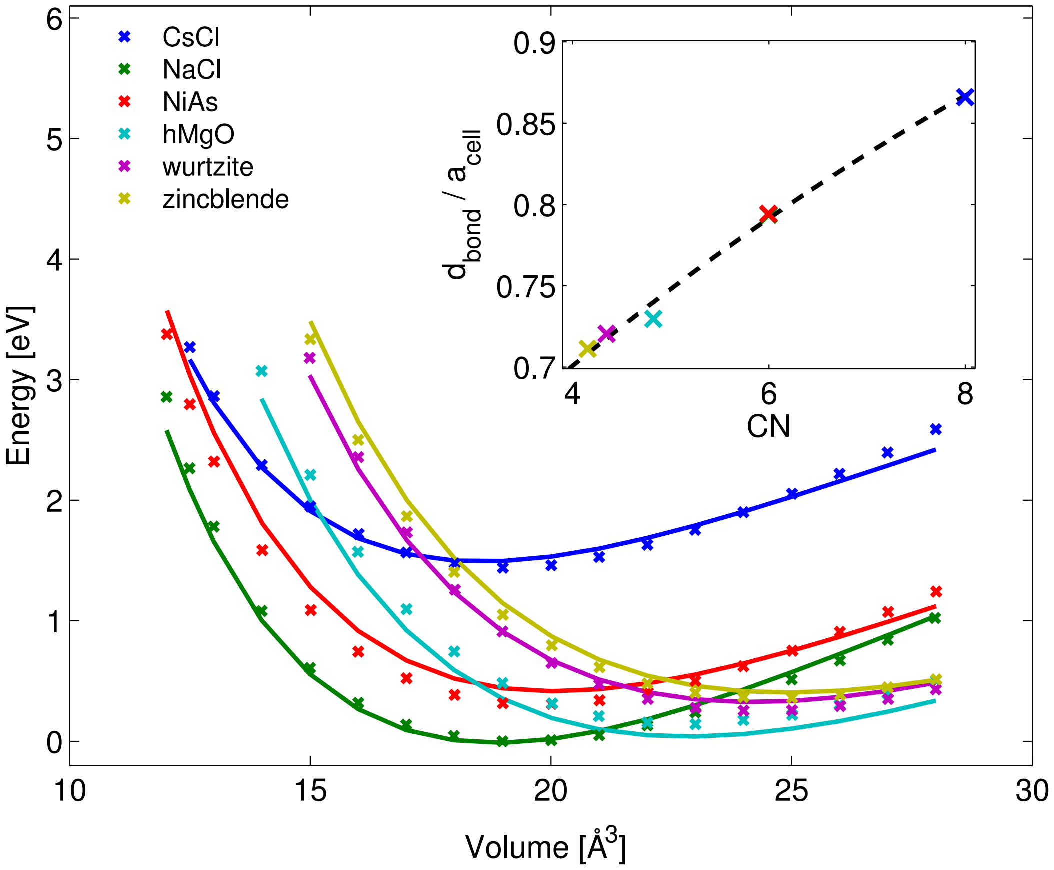

Figure 4.8: The first-principles energy-volume curves for the six polymorphs of MgO (*Joshi et al.*, 2012), are plotted together with the results of the global fit for the coordinated pair potentials. This fit assumed quadratic dependence of the three Morse potential scale parameters on coordination number, and kept the shape parameter fixed to $\alpha = \sqrt{2}$, the value corresponding to the standard Morse potential. Thus, using 9 free parameters, we obtain a good fit to the data with a goodness of fit (explained fraction of the variance) of $R^2 = 0.986$. Residuals are higher for more distorted structures, like NiAs and hMgO, as expected. The inset shows the smooth dependence of the oxygen bond–length scale ratio on coordination number, and is fit with a quadratic. This dependence is only important as a higher order correction to the free energy, in terms of the bond-angle contribution.

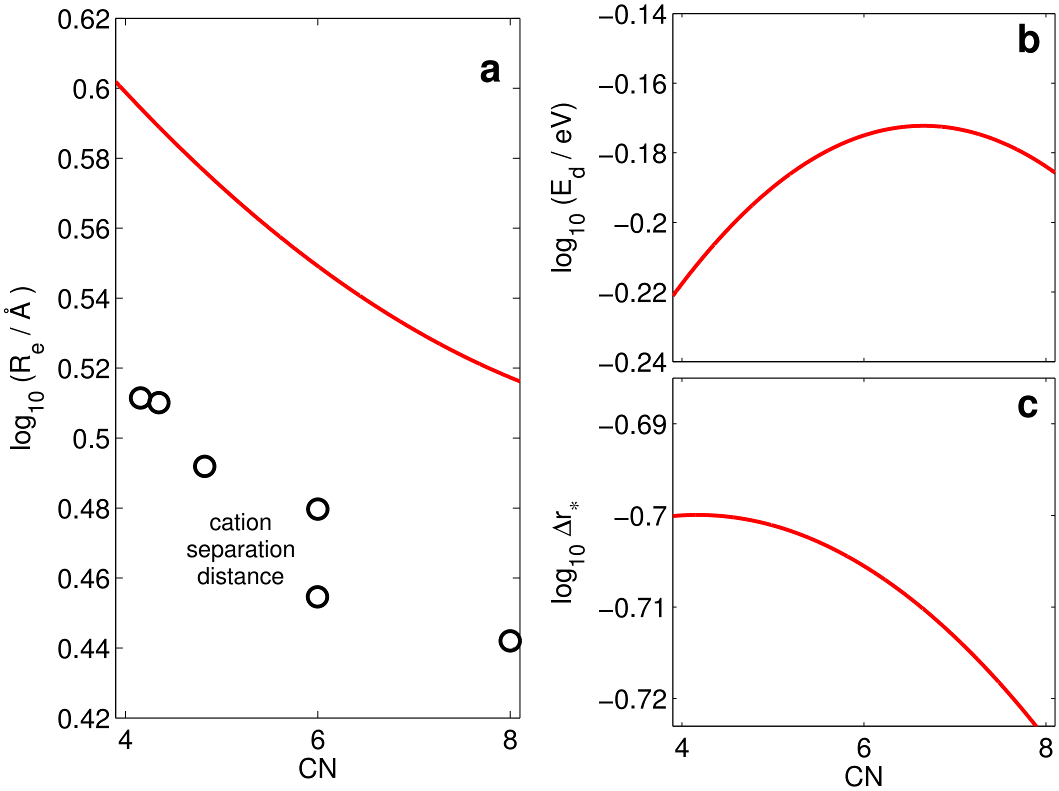

Figure 4.9: The best-fit trends for the pair potential scale parameters are plotted against coordination number. These trends correspond to the global best fit, shown in Figure 4.8. Panel (a) shows the variation of the equilibrium radius, (b) shows the energy well depth, and (c) shows the width of the energy well. Also plotted in panel (a) is the cation separation distances for each crystal structure at ambient pressure. The relationship between them and the equilibrium radius can plainly be seen, where the equilibrium radii are approximately given by scaling up the cation separation distance.

much, but not all, of the distortion effect. Also shown in the inset of Figure 4.8, is the systematic relationship between coordination number and the oxygen bond length in terms of the ratio of bond length to unit cell parameter, where $a_{\text{cell}} = V^{1/3}$. This nearly linear trend reflects the ability to more efficiently pack space with higher coordination numbers. Approximate values for the oxygen bond length, given by the quadratic fit shown in Figure 4.8, are needed for evaluating the internal degrees of freedom term to the free energy, which reflects the added internal states of each hard sphere caused by the motions of the oxygens.

The polynomials corresponding to the global best fit are shown in Figure 4.9. In panel a, the systematic decrease in the equilibrium radius with increasing coordination is plainly seen. Also plotted are the effective nearest-neighbor cation distances for each structure at ambient pressure. It is unsurprising that the equilibrium distance is related to the actually observed separation distances, though they are larger by roughly 20%, reflecting the constraints that periodicity places on the atomic positions. Panels b and c show the variation of the energy well depth and width with coordination number. To visualize what these polynomials mean, Figure 4.10 shows the pair potential energy wells calculated for a range of coordination numbers. The systematic changes in the pair potential parameters has the effect of shifting the energy wells toward smaller radii with increasing coordination. The depths of the wells also vary, causing six-coordinated pairs to have the lowest equilibrium energy, related to the stability of the 6-coordinated periclase structure. From this figure, we can see that the lowest energy curve at each separation distance shifts from low to high coordination as the separation distance decreases. It is primarily this trend that induces the gradual evolution of coordination number in oxide liquids with increasing compression.

4.3.4 Determining Liquid Properties using CHASM

With the coordination-dependent pair potentials in hand, we are now ready to apply the CHASM to determine the average liquid structure and thermodynamics. In this section, we demonstrate how to use the free energy minimization framework, described in Section 4.3.2, to calculate the thermodynamic properties of oxide liquids, specifically applying them to the MgO system.

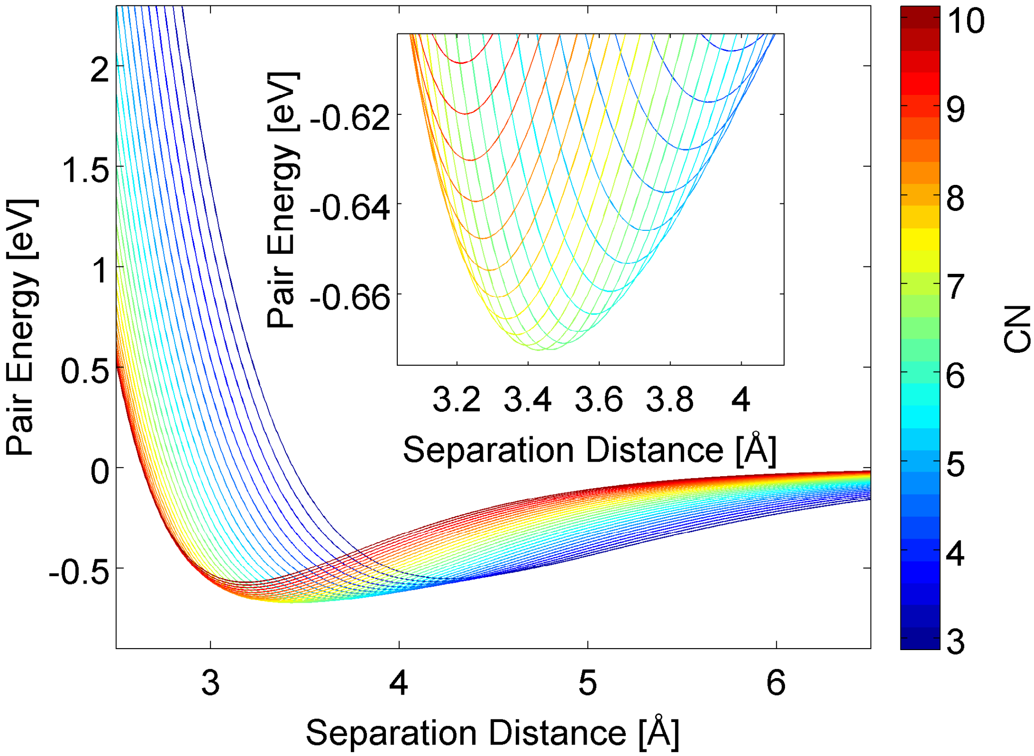

Figure 4.10: Results from the pair potential global fit, showing the energy well corresponding to each fractional coordination number along a continuum between 3 and 10. The inset figure shows a zoomed view of the minimum, which occurs for a coordination number near 6.

We begin by thinking of a complex coordinated liquid as being composed of an ideal mixture of liquids different oxygen coordination states. While this model neglects the interactions between different coordination states, which can play an important role in both the energy and entropy of the liquid, it represents a first order approximation to the liquid's properties. The variational method for hard spheres dictates that for a simple single component liquid, the correct value for the hard sphere size is simply the one that minimizes the free energy, represented by the sum of terms in Equation 4.6. The basic meaning of each term is described in Section 4.3.2 and the details of the calculation are given in the Appendix. The most important term to recall at this point is the structural energy perturbation term, which approximates the configurational internal energy using pair potentials. In the previous section, we developed a simple polynomial description of the pair potential terms as a function of coordination number, as shown in Figure 4.8. Using that model description, we obtain the pair potential parameters describing a range of coordination states, as shown in Fig. 4.10, for use in calculating the total free energy. It is important that coordination spectrum is sampled with sufficiently high resolution to adequately represent the continuous nature of coordination number for liquids. A sampling frequency of 0.25 coordination number is used in this study and deemed accurate based on the energetic differences between adjacent states as compared to kT.

At this point, we treat the total free energy function as a black box, which due to its primary reliance on analytic expressions, provides rapid function evaluations (each lasting only a fraction of a second). To determine the hard sphere size for each fixed coordination liquid, it is simply a matter of using standard function minimization routines[8]. At a range of

---

[8] The code developed in this investigation is implemented in MATLAB, and thus uses the library functions

sample volumes, the appropriate hard sphere size is thus determined for each coordination state along with the associated liquid free energy. The resulting set of free energy curves for the sampled coordination states of MgO are shown in Figure 4.11. Focusing on individual curves, each color-coded by their coordination state, we can see the overall trend that increasing coordination state simultaneously increases the zero-pressure free energy (shifting it upward), decreases the zero-pressure volume (shifting it left), and decreases the typical bulk modulus (reducing the curvature). This results in the expected behavior that the lowest energy coordination state changes gradually from low to high coordination as the liquid is compressed. Also shown in the inset of the same figure is the variation of the determined hard sphere diameters with compression. While this is an internal model variable that does not provide direct physical insight, we show it here to note that over much of the high compression region, the hard sphere sizes are nearly linear in log-volume. This supports the empirical deformability correction imposed by *Jing and Karato* (2011) to roughly account for compression effects. Unfortunately, there is non-negligible curvature at low pressures, and the shift of coordination states with compression will also induce added curvature. Nevertheless, this indicates that such a correction is not unreasonable for extending the applicability of the standard hard sphere model to slightly higher pressures.

After determining the compression behavior of independent coordination states, we obtain a rough first-order approximation of the liquid by treating each coordination as a quasi-independent state. Thinking of them as a set of non-interacting states that range in energy, Boltzmann's constants can be used to estimate the population distribution that minimizes the free energy: $p_i \propto \exp(-F_i)$, where F_i is the free energy and p_i is the pop-

---

fminunc and fminbnd, though any multidimensional nonlinear minimization routine should work.

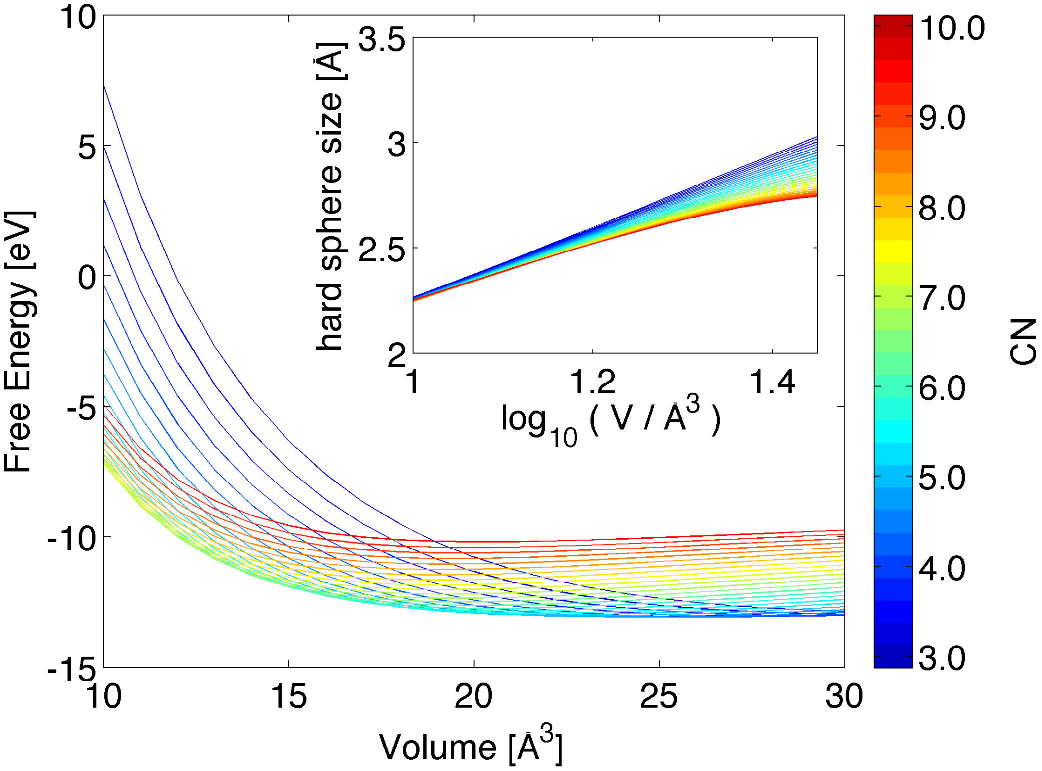

Figure 4.11: The free energy compression curves for a set of hypothetical MgO liquids each composed of a single fixed coordination number. An approximate set of coordination populations can be determined using Boltzmann factors at each volume. Thus, we can see that coordination state with the minimum free energy at a particular volume should lie near the peak of the coordination population distribution. The inset shows the variation of hard sphere sizes with increasing log-volume. The near linearity of these curves provides support for the empirical hard sphere deformability correction used in *Jing and Karato* (2011).

ulation of the i^{th} coordination state. This results in a population distribution that is a good initial guess for the true distribution, which accounts for interactions between the differently coordinated hard spheres. To enable the efficient calculation of the global free energy minimum, we parameterize the variations in population and hard sphere size, relying on low-order polynomials to capture smooth trends with a simple functional form. The hard sphere diameters are relatively straightforward, with a polynomial dependence of the diameter on log-coordination number capable of representing the observed variations with only a few degrees of freedom (3 to 5). The population distributions are somewhat trickier, adopting typical Gaussian-like peaks. To retain the flexibility afforded by polynomials, we model the log-population values as having polynomial dependence on coordination number; the choice of a second-order polynomial would restrict the model to a standard normal distribution (or Gaussian peak), which is not adequate to describe the observed shapes. We therefore choose a fourth-order polynomial, which is sufficient to capture the shapes of the initial population distributions over a range of volumes.[9]

Together, these parameterizations vastly reduce the size of phase space that must be searched to minimize the free energy, reducing the number of variables from over 50 to between 5 to 10, depending on the chosen polynomial degree. Given this simple parameterization, it is only a matter of using a standard minimization algorithm to locate the set of hard sphere size and population parameters that yields the lowest total free energies. We found that the best general procedure is to start by evaluating the hard sphere sizes (and free

---

[9] It should be mentioned that we restrict allowable solutions to those that result in only a single peak in the population distribution over the modeled coordination number range.. Therefore, only parameter combinations that result in monotonic increasing and decreasing behavior to the left and right of the peak are allowed. This condition is crucial to prevent unphysical minor peaks from appearing at either very high or very low coordination number, which occur as an artifact of trying to match the shape of the primary peak.

energies) individually for each coordination state. From there, the Boltzmann distribution is used to obtain a population estimate, which is represented using the polynomial formalism described above. Alternating between the two groups of parameters, fitting one set while the other remains fixed, appears to yield the most rapid convergence. In practice, we find that only a single additional iteration step is generally necessary to reach convergence for the free energy values. While the total free energy convergence fairly rapidly, after only tens to hundreds of function evaluations, the details of the populations distributions themselves can take more time, as seen by the fluctuations in the population curves in Figure 4.12.

Using the procedure described above, we calculate a set of isothermal compression curves over a range of temperatures and volumes. As stressed throughout the development of CHASM, obtaining the stable minimum free energy configuration is quite rapid requiring only a few minutes of time on a modern processor. Furthermore, each state point is entirely independent of the others, and thus the code can be run simultaneously on multiple processors if fine sampling of large compositional spaces is required. From the output of a set of equilibrium calculations, we can immediately examine the evolution of coordination number without any further processing. Figure 4.12 shows the evolution of the coordination state populations as a function of compression for MgO liquid at 3000K. In the main panel of the figure, the calculated populations of each coordination state are colored according to their relative probability values. From this figure, we can see the smooth evolution of the average coordination number with compression from values near 4 at 30 $\mathring{A}^3$ up to roughly 7.5 at 10 $\mathring{A}^3$, which is in strong agreement with the findings of the variation of average coordination number with compression presented by *Karki et al.* (2006). This plot also indicates the width of the distribution over coordinations states at a given volume. That width responds directly to temperature in a Boltzmann-like fashion, populating a wider range of more energetically costly states at higher temperatures. The inset to the same figure displays the same information in a different way, now showing the evolution of the occupation of each coordination state as a function of compression. For viewing purposes, we have binned the fractional coordination states together into integer width bins. When viewed in this way, we can highlight the contribution of each coordination state to the overall changing structure of the liquid. While *Karki et al.* (2006) does not show any plots like this tracking the distribution of coordination states, a later paper on silica (*Karki et al.*, 2007) does show the evolving populations of each coordination number, which shows the same qualitative behaviors as seen here, and are likely to be rather similar since the average coordination number are in agreement.

Finally, we can obtain the equation of state curves for a set of isotherms calculated using CHASM and compare with first-principles results. In Figure 4.13, we show the derived pressure-volume curves for a set of MgO liquid isotherms. The direct output of the CHASM are a set of free energy curves, which are generally of less interest to most people. In order to convert each curve into a pressure volume curve, we use the same strategy as that used by the FPMD practitioners (see, e.g., *de Koker and Stixrude*, 2009), which is simply to fit a finite strain equation of state to each isothermal curve. The finite strain model is also known as the Birch-Murnaghan equation of state. We find that a standard third-order model, with 4 parameters, V_0, F_0, K_0, and K_0' is perfectly adequate to represent the free energy

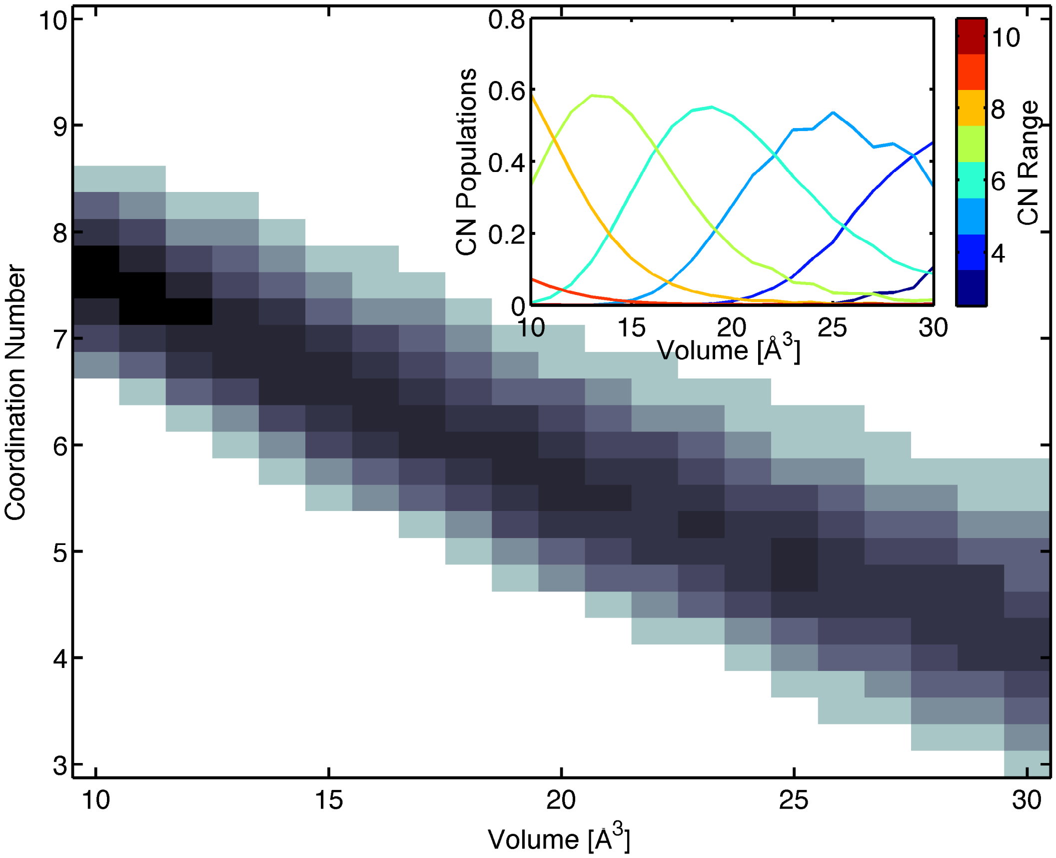

Figure 4.12: Evolution of the coordination populations for MgO liquid at 3000K. The main figure shows probability distribution, indicated by the darkness of the colormap, amongst the coordination states at each volume. The general trend of increasing coordination number with compression can be seen, as can a slight reduction in the width of the distribution with increasing compression. The inset figure shows binned coordination populations as a function of volume, showing dominantly 4-coordinated Mg transforming smoothly to dominantly 7 and 8-coordinated Mg at the highest pressures.

curves. Using this fit to each isotherm, it is then simple to calculate the pressure for a range of volumes, as shown in Figure 4.13. Also shown in the inset of the same figure is plot of the free energy curves against a normalized volume, V_X, which is taken to be the zero-pressure volume at 3000 K. We find a 3000K volume scale of $V_X = 25.4\AA^3$, which falls short of the first-principles result of $V_X = 27.4\AA^3$ (*Karki et al.*, 2006). In all likelihood, the fact that we are underestimating the zero-pressure volume is tied to some inadequacies in the internal degrees of freedom correction, probably related to the as-yet unaccounted for freedom of varying oxygen bond length. Nevertheless, if we plot the scaled free energy curves, we can see that we are doing quite an excellent job of matching the behavior in the first-principles calculations. This is our strongest indicator that the CHASM formalism is capable of accurately capturing the behavior of oxide melts.

4.4 Conclusions

We have developed a new simplified model to model the structure and thermodynamics of oxide and silicate melts. The **C**oordinated **HA**rd **S**phere **M**odel (CHASM) is based on an extension to the standard hard sphere mixture model which accounts for the range of local oxygen coordination state about each cation, which appears to be the key to compression mechanisms in oxide melts. In this formalism, a coordinated hard sphere is taken to represent a single cation surrounded by a bonded shell of coordinated oxygens, where the coordination number represents the local connectivity and bonding relations between each cation and its neighbors. The CHASM is based on the principle of free energy minimization, using it to determine sizes and populations of the hard sphere oxides. In this

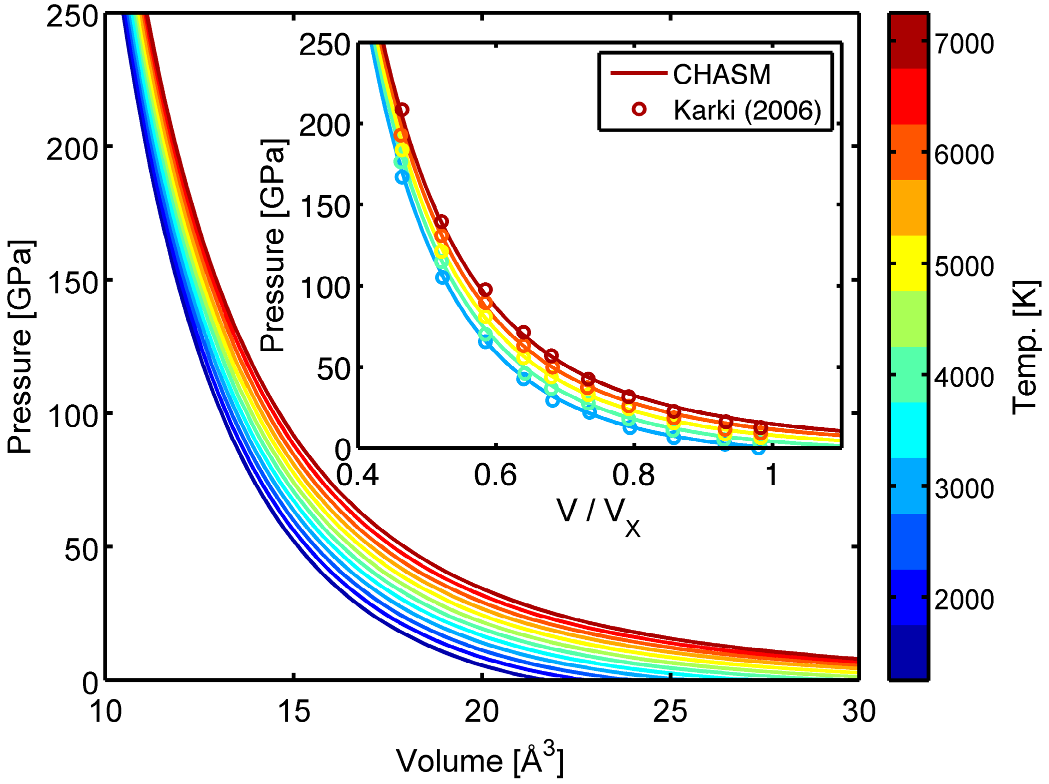

Figure 4.13: Resulting equation of state curves for MgO liquid over a range of temperatures. The curves are obtained by fitting third-order Birch-Murnaghan equations to the free energy curves determined by CHASM, from which the pressure curves are readily determined. The inset figure shows the same curves replotted against a relative volume, where the volume scale is given by the 3000K zero-pressure volume. Also plotted are the first-principles results from *Karki et al.* (2006), which our calculation is in strong agreement with, except for the value of our volume scale, which is 25.4 $\AA^3$ as opposed to 27.4 $\AA^3$ from *Karki et al.* (2006).

way the evolution of coordination populations is a direct prediction of the model, where the coordination states are chosen so as to minimize the free energy.

In addition to representing coordination states, the CHASM also account for internal energy variations that depend on the average structure of the liquid. This is incorporated through the use of coordinated cation pair potentials. The pair potentials actually represent nearly the only input to the model that is unique to the liquid being calculated—all remaining differences are derived from these all important potentials. We also demonstrate for the first-ever CHASM represented liquid, MgO melt, how the pair potentials can be obtained by training their values on the known energy-volume curves for a variety of solid structures, which discretely sample periodic examples from generally disordered structural continuum that is available to the liquid. By obtaining all of the input variables for the CHASM from solid structures, we have therefore generated a model that is fully predictive for liquid structure and thermodynamics, not requiring any input from liquid measurements (though it will be possible, but not necessary, to incorporate information from liquid experiments to adjust the pair potentials and improve the model).

Finally we demonstrate the use of the CHASM to calculate equations of state using MgO as an example. By design, the CHASM relies entirely on analytic function evaluations and simple 1D integrals, and thus computation remains rapid. The CHASM can obtain a converged calculation for a single state point (with a particular volume, temperature, and composition) in roughly one to a few minutes on a standard desktop processor. We show how the resulting prediction of coordination state evolution compares favorably with the orders of magnitude more time-consuming first-principles calculations. We also produce a set of predicted equation of state isotherms which match the first-principles results nearly perfectly after accounting for a shift in the zero-pressure volume scale.

In future work, we plan to apply this framework to the silica system, and then use it to explore thermodynamic properties along the $\mathrm{MgO–SiO_2}$ binary. Given that we have a method for describing the free energy surface of the liquid, we should also be able to look into the possibilities of making melting curve calculations, when combined with models for the solid free energies. Once multiple oxide components are introduced to the calculation, it may be important to improve the method of training pair potentials to allow for directly fitting the intermediate composition structures, rather than relying solely on the endmember description. Additional work should be put in to improving the internal degrees of freedom term in the free energy. There are also gains that can be made in implementation, potentially decreasing calculation times by a further factor of 3 to 10. Finally, we expect in the near future to begin the challenging process of incorporating FeO as a component, as it is much more complex than the other oxides, but crucial to our understanding of early Earth evolution. With its multiple internal spin state that can drastically effect local volumes, the Fe cations will have an added dimension to the available states, populating a range of oxygen coordination numbers while simultaneously populating a few discrete spin states.

— Appendices —

4.A Generalized Hard Sphere Mixture

The complete set of expressions for the equation of state of a hard sphere mixture is presented in *Mansoori et al.* (1971). For completeness, we include these expressions here, which represent the configurational free energy contribution ΔF_{HS} to the total free energy. We start with the expression for the packing fraction:

$$\eta_i \equiv \frac{V_{\text{HS},i}}{V} = x_i \frac{4\pi}{3V} R^3_{\text{HS},i} \quad \text{where} \quad \eta = \sum_i \eta_i \tag{4.17}$$

The compressibility Z is expressed as a function of the packing fraction:

$$Z = \frac{PV}{kT} \approx [1 + \eta + \eta^2 - 3\eta(y_1 + y_2\eta) - \eta^3 y_3](1-\eta)^{-3} \tag{4.18}$$

where the needed additional relations are:

$$\begin{aligned} y_1 &= \sum_{i=1} \sum_{j>i} \Delta_{ij}(R_i + R_j)(R_i R_j)^{-1/2} \\ y_2 &= \sum_{i=1} \sum_{j>i} \Delta_{ij} \sum_{k=1}^{N} \left(\frac{\eta_k}{\eta}\right) \frac{(R_i R_j)^{1/2}}{R_k} \\ y_3 &= \left[\sum_{i=1} \left(\frac{\eta_i}{\eta}\right)^{2/3} x_i^{1/3}\right]^3 \\ \Delta_{ij} &= \left[\frac{(\eta_i \eta_j)^{\frac{1}{2}}}{\eta}\right] \left[\frac{(R_i - R_j)^2}{R_i R_j}\right] (x_i x_j)^{\frac{1}{2}} \end{aligned} \tag{4.19}$$

These are all combined to express the contribution to the free energy and entropy of the liquid:

$$\begin{aligned}\frac{\Delta F_{HS}}{kT} = \frac{F_{id}}{kT} - \frac{3}{2}(1 - y_1 + y_2 + y_3) + (3y_2 + 2y_3)(1-\eta)^{-1} + \\ \frac{3}{2}(1 - y_1 - y_2 - \frac{1}{3}y_3)(1-\eta)^{-2} + (y_3 - 1)\log(1-\eta)\end{aligned} \tag{4.20}$$

$$\Delta S_{HS} = k \log Z - \frac{(F_{HS} - F_{id})}{T} \tag{4.21}$$

4.B Hard Spheres with Structural Perturbation

The variational approach to hard spheres allows the incorporation of a structural perturbation into the free energy expression, which captures the dependence of the internal energy on the structure of the liquid. This perturbation method was developed by *Mansoori and Canfield* (1969), and the resulting expressions are shown here.

We begin by approximating the structural energy contribution using the radial distribution function $g(r)$, together with pair potentials $u(r)$:

$$\begin{aligned}F \approx F_{HS} + F_{\text{struc}} &= F_{HS} + \int_0^\infty P(r)u(r)dr \\ &\approx F_{HS} + 4\pi\rho \int_{R_{min}}^\infty r^2 u(r) g(r) dr\end{aligned} \tag{4.22}$$

Mansoori and Canfield (1969) showed that the above integral for the structural perturbation energy E_{struc} can be transformed using a Laplace transformation, yielding:

$$F \approx F_{HS} + 4\pi\rho R_{HS}^3 \int_0^\infty U(s)G(s)ds \tag{4.23}$$

It turns out that while expressing the radial distribution function directly is quite challenging, relatively simple analytic approximations exist for its Laplace transform (which is an exponential integral transform similar to a Fourier transform, but one where the exponential weighting function is real valued and thus imposes a decaying rather than oscillatory envelope). We can understand this on an intuitive level, since the radial distribution function is a function that inherently decays to an asymptotic constant value of one. Therefore, by expressing the RDF in terms of its Laplace transform, it is ensured that the RDF will necessarily possess the required asymptotic behavior for large separation distances.

$G(s)$ is the Laplace transform of the hard sphere radial distribution function and $U(s)$ is the inverse Laplace transform of the pairwise potential energy function, defined by:

$$G(s) = \int_1^\infty \exp(-sx) g_0(x) x dx \tag{4.24}$$

$$xu(x) = \int_0^\infty \exp(-sx) U(s) ds \tag{4.25}$$

These expression for $G(s)$ has a simple functional form given by Percus-Yevick theory, and can be found in any standard text on the hard sphere model. (One easily readable account is given by *Yuste et al.* (1998). The inverse Laplace transform is a standard integral, and thus we can obtain the expression for an exponential, appropriate for the generalized Morse potential, using simple integral tables or Mathematica.

4.C Representing Variable Coordination

In order to carry out the structure perturbation calculation, we must have an expression for the pair potential between every possible pair of particles. This is not generally practical because the solid structures generally sample only a single coordination state, and thus do not constrain the energetic cost of two different coordination states interacting. To handle this challenge, the empirical molecular dynamics community long ago developed the practice of obtaining rules for "mixing" the pair potentials of two different particles to estimate an effective cross-pair-potential. The most accurate of the available mixing rules are given by *Kong* (1973), and we summarize the results of the mixing rules below, considering interactions between particles of type 1 and 2:

$$\begin{aligned}
&\text{attractive mixing:} && \frac{1}{\sigma_A^{12}} = \frac{1}{\sigma_A^{11}} + \frac{1}{\sigma_A^{22}} \quad , \quad \epsilon_A^{12} = \sqrt{\epsilon_A^{11}\epsilon_A^{22}} \\
&\text{repulsive mixing:} && \sigma_R^{12} = \frac{1}{2}\left(\sigma_R^{11} + \sigma_R^{22}\right) \quad , \quad \left(\frac{\epsilon_R^{12}}{\sigma_R^{12}}\right)^{\sigma_R^{12}} = \sqrt{\left(\frac{\epsilon_R^{11}}{\sigma_R^{11}}\right)^{\sigma_R^{11}} \left(\frac{\epsilon_R^{22}}{\sigma_R^{22}}\right)^{\sigma_R^{22}}} \\
&\text{simple mixing terms:} && \sigma_A^{-1}, \;\; \log(\epsilon_A), \;\; \sigma_R, \;\; \sigma_R \log(\epsilon_R/\sigma_R)
\end{aligned} \tag{4.26}$$

the simple mixing terms therefore represent the quantities that can be mixed through simple arithmetic means.

4.D Free Energy Contribution of Bond-Angle Deviations

The free energy contribution due to the freedom of bond-angle is approximately determined using elementary quantum mechanics. We think of the oxygens as being constrained to lie on the surface of a sphere whose radius is the bond length. The oxygens are not allowed to freely wander along the full surface of the sphere, but are rather confined to a portion of it defined by the average locations of the neighboring bonded oxygens. Thus we can imagine each oxygen as being confined to portion of the bonding sphere whose area is just the total area divided by the coordination number:

$$A_{\text{oxy}} = \frac{4\pi l^2}{N_{cn}} = 2\pi l^2(1 - \cos\phi_{\text{oxy}}) \tag{4.27}$$

Neglecting the detailed shape of the confinement area, we can approximate this area on the sphere as a spherical cap, whose opening angle is ϕ_{oxy}, given above.

When considering the true confinement area, we must recognize that the neighboring bonded oxygens will avoid overlapping one another which effectively reduces the size of the confinement area. The expression for this is given by:

$$\phi_{\text{oxy}} = \cos^{-1}(1 - 2/N_{cn}) \quad \text{and} \quad \Delta\phi_{\text{NN}} = \sin^{-1}(r_{\text{oxy}}/l) \tag{4.28}$$

Thus the adjusted confinement area for each bonded oxygen is just:

$$A_{\text{conf}} = 2\pi l^2(1 - \cos\phi_{\text{conf}}) \quad \text{and} \quad \phi_{\text{conf}} = \phi_{\text{oxy}} - \Delta\phi_{\text{NN}} \tag{4.29}$$

Now that we have an estimate of the confinement region, we can calculate the free energy contribution by approximating this as a particle in a two-dimensional box. Neglecting the effect of the curved surface, the 2D single particle in a box expression is:

$$F^{1}_{\text{bond Ang}} = -kT \log\left(\frac{mkTA_{\text{conf}}}{2\pi\hbar^{2}}\right) \tag{4.30}$$

where the value for the confinement area is given above. The total free energy is given by the sum over each particle (coordinated oxygen), but we must also account for the fact that the oxygens are shared between neighboring cations. Therefore, to avoid over-counting, we also multiply by the factor $f = n_{\text{oxy}}/N_{cn}$, yielding the final expression for the bond-angle freedom:

$$F_{\text{bond Ang}} = fN_{cn}F^{1}_{\text{bond Ang}} = n_{\text{oxy}}F^{1}_{\text{bond Ang}} \tag{4.31}$$

Bibliography

Ai, Y., and R. A. Lange (2008), New acoustic velocity measurements on CaO-MgO-Al2O3-SiO2 liquids: Reevaluation of the volume and compressibility of CaMgSi2O6-CaAl2Si2O8 liquids to 25 GPa, *J. Geophys. Res.*, *113*(B4), B04,203–.

Anderson, O. L. (1998), Thermoelastic properties of MgSiO 3 perovskite using the Debye approach, *American Mineralogist*, *83*(1-2), 23–35.

Andrault, D., N. Bolfan-Casanova, M. Bouhifd, N. Guignot, and T. Kawamoto (2007), The role of Al-defects on the equation of state of Al-(Mg,Fe)SiO3 perovskite, *Earth and Planetary Science Letters*, *263*(3-4), 167–179.

Angel, R. J. (2000), Equations of State, *Reviews in Mineralogy and Geochemistry*, *41*(1), 35–59, doi: 10.2138/rmg.2000.41.2.

Asimow, P. D., and T. J. Ahrens (2010), Shock compression of liquid silicates to 125 GPa: The anorthite-diopside join, *J. Geophys. Res.*, *115*(B10), B10,209–.

Bass, J. D., R. C. Liebermann, D. J. Weidner, and S. J. Finch (1981), Elastic properties from acoustic and volume compression experiments, *Physics of the Earth and Planetary Interiors*, *25*(2), 140–158.

Bottinga, Y., and D. F. Weill (1970), Densities of liquid silicate systems calculated from partial molar volumes of oxide components, *American Journal of Science*, *269*(2), 169–182.

Burbidge, J. B., L. Magee, and A. L. Robb (1988), Alternative Transformations to Handle Extreme Values of the Dependent Variable, *Journal of the American Statistical Association*, *83*(401), 123–127.

Catalli, K., S.-H. Shim, P. Dera, V. B. Prakapenka, J. Zhao, W. Sturhahn, P. Chow, Y. Xiao, H. Cynn, and W. J. Evans (2011), Effects of the Fe3+ spin transition on the properties of aluminous perovskiteNew insights for lower-mantle seismic heterogeneities, *Earth and Planetary Science Letters*, *310*(3-4), 293–302, doi: 10.1016/j.epsl.2011.08.018.

Chambers, J. E. (2004), Planetary accretion in the inner Solar System, *Earth and Planetary Science Letters*, *223*(3-4), 241–252.

Cohen, R. E., O. Glseren, and R. J. Hemley (2000), Accuracy of equation-of-state formulations, *American Mineralogist*, *85*(2), 338–344.

David, W. I. F., and D. S. Sivia (2001), Background estimation using a robust Bayesian analysis, *Journal of Applied Crystallography*, *34*(3), 318–324, doi: doi:10.1107/S0021889801004332.

Davies, D. R., S. Goes, J. Davies, B. Schuberth, H.-P. Bunge, and J. Ritsema (2012), Reconciling dynamic and seismic models of Earth's lower mantle: The dominant role of thermal heterogeneity, *Earth and Planetary Science Letters*, *353-354*(0), 253–269.

Daw, M. S., and M. I. Baskes (1984), Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals, *Phys. Rev. B*, *29*(12), 6443–6453.

Daw, M. S., S. M. Foiles, and M. I. Baskes (1993), The embedded-atom method: a review of theory and applications, *Materials Science Reports*, *9*(7-8), 251–310.

de Koker, N., and L. Stixrude (2009), Self-consistent thermodynamic description of silicate liquids, with application to shock melting of MgO periclase and MgSiO3 perovskite, *Geophysical Journal International*, *178*(1), 162–179.

de Koker, N., B. B. Karki, and L. Stixrude (2013), Thermodynamics of the MgO-SiO2 liquid system in Earth's lowermost mantle from first principles, *Earth and Planetary Science Letters*, *361*(0), 58–63.

de Koker, N. P., L. Stixrude, and B. B. Karki (2008), Thermodynamics, structure, dynamics, and freezing of Mg2SiO4 liquid at high pressure, *Geochimica et Cosmochimica Acta*, *72*(5), 1427–1441.

Dewaele, A., F. Datchi, P. Loubeyre, and M. Mezouar (2008), High pressure–high temperature equations of state of neon and diamond, *Phys. Rev. B*, *77*(9), 094,106–9.

Dorfman, S. M., V. B. Prakapenka, Y. Meng, and T. S. Duffy (2012), Intercomparison of pressure standards (Au, Pt, Mo, MgO, NaCl and Ne) to 2.5 Mbar, *J. Geophys. Res.*, *117*(B8), n/a–n/a.

Dorfman, S. M., Y. Meng, V. B. Prakapenka, and T. S. Duffy (2013), Effects of Feenrichment on the equation of state and stability of (Mg,Fe)SiO3 perovskite, *Earth and Planetary Science Letters*, *361*(0), 249–257.

Dziewonski, A. M., and D. L. Anderson (1981), Preliminary reference Earth model, *Physics of the earth and planetary interiors*, *25*(4), 297–356.

Fiquet, G., D. Andrault, A. Dewaele, T. Charpin, M. Kunz, and D. Hasermann (1998), P-V-T equation of state of MgSiO3 perovskite, *Physics of the Earth and Planetary Interiors*, *105*(1-2), 21–31.

Fiquet, G., A. Dewaele, D. Andrault, M. Kunz, and T. Le Bihan (2000), Thermoelastic Properties and Crystal Structure of MgSiO3 Perovskite at Lower Mantle Pressure and Temperature Conditions, *Geophys. Res. Lett.*, *27*, –.

Funamori, N., S. Yamamoto, T. Yagi, and T. Kikegawa (2004), Exploratory studies of silicate melt structure at high pressures and temperatures by in situ X-ray diffraction, *J. Geophys. Res.*, *109*(B3), B03,203–.

Guillot, B., and P. Sarda (2006), The effect of compression on noble gas solubility in silicate melts and consequences for degassing at mid-ocean ridges, *Geochimica et Cosmochimica Acta*, *70*(5), 1215–1230, doi: 10.1016/j.gca.2005.11.007.

Guo, X., R. A. Lange, and Y. Ai (2013), The density and compressibility of CaO-FeO-SiO2 liquids: evidence for four-coordinated Fe2+ in the CaFeO2 component, *Geochimica et Cosmochimica Acta*, *(submitted)*.

Hammersley, A. P., S. O. Svensson, M. Hanfland, A. N. Fitch, and D. Hausermann (1996), Two-dimensional detector software: From real detector to idealised image or two-theta scan, *High Pressure Research*, *14*(4), 235–248.

Heinz, D. L., and R. Jeanloz (1987), Temperature measurements in the laser-heated diamond cell, in *Geophys. Monogr. Ser.*, vol. 39, pp. 113–127, AGU, Washington, DC.

Hill, R. (1952), The Elastic Behaviour of a Crystalline Aggregate, *Proceedings of the Physical Society. Section A*, *65*(5), 349–.

Hinrichsen, B. (2006), Advances in data reduction of high-pressure x-ray powder diffraction data from two-dimensional detectors: a case study of schafarzikite (FeSb2O4).

Hinrichsen, B., R. E. Dinnebier, and M. Jansen (2008), Chapter 14 Two-dimensional Diffraction Using Area Detectors, in *Powder Diffraction: Theory and Practice*, pp. 414–438, The Royal Society of Chemistry.

Hoppe, R. (1979), Effective coordination numbers (ECoN) and mean Active fictive ionic radii (MEFIR)[1,2]*, *Zeitschrift fr Kristallographie*, *150*(1-4), 23–52, doi: 10.1524/zkri.1979.150.1-4.23.

Iizuka, R., H. Kagi, and K. Komatsu (2010), Comparing ruby fluorescence spectra at high pressure in between methanol-ethanol pressure transmitting medium and its deuteride, *Journal of Physics: Conference Series*, *215*(1), 012,177–.

Ishii, M., and J. Tromp (1999), Normal-Mode and Free-Air Gravity Constraints on Lateral Variations in Velocity and Density of Earth's Mantle, *Science*, *285*(5431), 1231–1236, doi: 10.1126/science.285.5431.1231.

Jackson, J. M., W. Sturhahn, O. Tschauner, M. Lerche, and Y. Fei (2009), Behavior of iron in (Mg,Fe)SiO3 post-perovskite assemblages at Mbar pressures, *Geophys. Res. Lett.*, *36*, –.

Jacobsen, S. D., C. M. Holl, K. A. Adams, R. A. Fischer, E. S. Martin, C. R. Bina, J.-F. Lin, V. B. Prakapenka, A. Kubo, and P. Dera (2008), Compression of single-crystal magnesium oxide to 118 GPa and a ruby pressure gauge for helium pressure media, *American Mineralogist*, *93*(11-12), 1823–1828, doi: 10.2138/am.2008.2988.

Jaffe, J. E., J. A. Snyder, Z. Lin, and A. C. Hess (2000), LDA and GGA calculations for high-pressure phase transitions in ZnO and MgO, *Phys. Rev. B*, *62*(3), 1660–1665.

Jing, Z., and S.-i. Karato (2011), A new approach to the equation of state of silicate melts: An application of the theory of hard sphere mixtures, *Geochimica et Cosmochimica Acta*, *75*(22), 6780–6802, doi: 10.1016/j.gca.2011.09.004.

Joshi, K., B. Sharma, U. Paliwal, and B. Barbiellini (2012), Pressure-dependent electronic properties of MgO polymorphs: a first-principles study of Compton profiles and autocorrelation functions, *Journal of Materials Science*, pp. 1–9.

Karki, B. (1997), Structure and elasticity of MgO at high pressure, *American Mineralogist*, *82*(1), 51–60.

Karki, B. B., D. Bhattarai, and L. Stixrude (2006), First-principles calculations of the structural, dynamical, and electronic properties of liquid MgO, *Phys. Rev. B*, *73*(17), 174,208–.

Karki, B. B., D. Bhattarai, and L. Stixrude (2007), First-principles simulations of liquid silica: Structural and dynamical behavior at high pressure, *Phys. Rev. B*, *76*(10), 104,205–.

Kesson, S. E., J. D. Fitz Gerald, and J. M. Shelley (1998), Mineralogy and dynamics of a pyrolite lower mantle, *Nature*, *393*(6682), 252–255.

Kong, C. L. (1973), Combining rules for intermolecular potential parameters. II. Rules for the Lennard-Jones (12–6) potential and the Morse potential, *J. Chem. Phys.*, *59*(5), 2464–2467.

Kudoh, Y., C. T. Prewitt, L. W. Finger, A. Darovskikh, and E. Ito (1990), Effect of iron on the crystal structure of (Mg,Fe)SiO3 perovskite, *Geophys. Res. Lett.*, *17*(10), 1481–1484.

Labrosse, S., J. W. Hernlund, and N. Coltice (2007), A crystallizing dense magma ocean at the base of the Earth/'s mantle, *Nature*, *450*(7171), 866–869.

Lebowitz, J. L. (1964), Exact solution of generalized Percus-Yevick equation for a mixture of hard spheres, pp. –.

Lee, S. K. (2011), Simplicity in melt densification in multicomponent magmatic reservoirs in Earths interior revealed by multinuclear magnetic resonance, *Proceedings of the National Academy of Sciences*, *108*(17), 6847–6852, doi: 10.1073/pnas.1019634108.

Lundin, S., K. Catalli, J. Santilln, S.-H. Shim, V. Prakapenka, M. Kunz, and Y. Meng (2008), Effect of Fe on the equation of state of mantle silicate perovskite over 1Mbar, *Physics of the Earth and Planetary Interiors*, *168*, 97–102.

Mansoori, G. A., and F. B. Canfield (1969), Variational Approach to the Equilibrium Thermodynamic Properties of Simple Liquids. I, *J. Chem. Phys.*, *51*(11), 4958–4967.

Mansoori, G. A., N. F. Carnahan, K. E. Starling, and T. W. Leland, Jr. (1971), Equilibrium Thermodynamic Properties of the Mixture of Hard Spheres, *J. Chem. Phys.*, *54*(4), 1523–1525.

Mao, H. K., J. Xu, and P. M. Bell (1986), Calibration of the Ruby Pressure Gauge to 800 kbar Under Quasi-Hydrostatic Conditions, *J. Geophys. Res.*, *91*, –.

Mattern, E., J. Matas, Y. Ricard, and J. Bass (2005), Lower mantle composition and temperature from mineral physics and thermodynamic modelling, *Geophysical Journal International*, *160*(3), 973–990.

Mosenfelder, J. L., P. D. Asimow, and T. J. Ahrens (2007), Thermodynamic properties of Mg2SiO4 liquid at ultra-high pressures from shock measurements to 200 GPa on forsterite and wadsleyite, *J. Geophys. Res.*, *112*(B6), B06,208–.

Mosenfelder, J. L., P. D. Asimow, D. J. Frost, D. C. Rubie, and T. J. Ahrens (2009), The MgSiO3 system at high pressure: Thermodynamic properties of perovskite, postperovskite, and melt from global inversion of shock and static compression data, *J. Geophys. Res.*, *114*(B1), B01,203–.

Mukhopadhyay, S. (2012), Early differentiation and volatile accretion recorded in deep-mantle neon and xenon, *Nature*, *486*(7401), 101–104.

Mukhopadhyay, S., R. Parai, M. K. Peto, and J. Tucker (2012), The recycling efficiency of water and noble gases to the mantle, in *AGU Meeting*.

Nakajima, M., and D. Stevenson (2012), The Initial State of the Moon Forming Disk and the Earth's Mantle Based on SPH Simulations, in *Lunar and Planetary Institute Science Conference Abstracts*, vol. 43, p. 2627.

Oganov, A. R., and P. I. Dorogokupets (2003), All-electron and pseudopotential study of MgO: Equation of state, anharmonicity, and stability, *Phys. Rev. B*, *67*(22), 224,110–.

O'Keeffe, M., B. Hyde, and J.-O. Bovin (1979), Contribution to the crystal chemistry of orthorhombic perovskites: MgSiO3 and NaMgF3, *4*(4), 299–305–.

Orear, J. (1982), Least squares when both variables have uncertainties, *American journal of physics*, *50*, 912.

Press, W., S. Teukolsky, W. Vetterling, and B. Flannery (2007), *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, – pp., Cambridge University Press.

Ramo, D. M., and L. Stixrude (2012), Spin crossover in Fe2SiO4 liquid at high pressure.

Rivers, M., V. B. Prakapenka, A. Kubo, C. Pullins, C. M. Holl, and S. D. Jacobsen (2008), The COMPRES/GSECARS gas-loading system for diamond anvil cells at the Advanced Photon Source, *High Pressure Research*, *28*(3), 273–292, doi: 10.1080/08957950802333593.

Shen, G., M. L. Rivers, Y. Wang, and S. Sutton (2001), Laser heated diamond cell system at the Advanced Photon Source for in situ x-ray measurements at high pressure and temperature, *Review of Scientific Instruments*, *72*(2), 1273–1282.

Shimizu, H., H. Imaeda, T. Kume, and S. Sasaki (2005), High-pressure elastic properties of liquid and solid neon to 7GPa, *Phys. Rev. B*, *71*(1), 014,108–.

Silvera, I. F., A. D. Chijioke, W. J. Nellis, A. Soldatov, and J. Tempere (2007), Calibration of the ruby pressure scale to 150 GPa, *physica status solidi (b)*, *244*(1), 460–467.

Singh, A. K. (1993), The lattice strains in a specimen (cubic system) compressed nonhydrostatically in an opposed anvil device, *J. Appl. Phys.*, *73*(9), 4278–4286.

Singh, A. K. (2009), Analysis of nonhydrostatic high-pressure diffraction data (cubic system): Assessment of various assumptions in the theory, *J. Appl. Phys.*, *106*(4), 043,514–8.

Sivia, D., and J. Skilling (2006), Data analysis: a Bayesian tutorial.

Solomatov, V. S., and D. J. Stevenson (1993), Suspension in Convective Layers and Style of Differentiation of a Terrestrial Magma Ocean, *J. Geophys. Res.*, *98*, –.

Spera, F. J., D. Nevins, M. Ghiorso, and I. Cutler (2009), Structure, thermodynamic and transport properties of CaAl2Si2O8 liquid. Part I: Molecular dynamics simulations, *Geochimica et Cosmochimica Acta*, *73*(22), 6918–6936.

Stevenson, D. (2007), 9.01 - Earth Formation and Evolution, in *Treatise on Geophysics*, edited by E. in Chief:Gerald Schubert, pp. 1–11, Elsevier, Amsterdam.

Stixrude, L., and B. Karki (2005), Structure and Freezing of MgSiO3 Liquid in Earth's Lower Mantle, *Science*, *310*(5746), 297–299.

Stixrude, L., N. de Koker, N. Sun, M. Mookherjee, and B. B. Karki (2009), Thermodynamics of silicate liquids in the deep Earth, *Earth and Planetary Science Letters*, *278*(3-4), 226–232.

Sugahara, M., A. Yoshiasa, Y. Komatsu, T. Yamanaka, N. Bolfan-Casanova, A. Nakatsuka, S. Sasaki, and M. Tanaka (2006), Reinvestigation of the MgSiO3 perovskite structure at high pressure, *American Mineralogist*, *91*(4), 533–536, doi: 10.2138/am.2006.1980.

Tackley, P. J. (2011), Dynamics and evolution of the deep mantle resulting from thermal, chemical, phase and melting effects, *Earth-Science Reviews*, (0), –, doi: 10.1016/j.earscirev.2011.10.001.

Takemura, K. (2007), Pressure scales and hydrostaticity, *High Pressure Research: An International Journal*, *27*(4), 465–472.

Takemura, K., and A. Dewaele (2008), Isothermal equation of state for gold with a He-pressure medium, *Phys. Rev. B*, *78*(10), 104,119–.

Tan, E., and M. Gurnis (2007), Compressible thermochemical convection and application to lower mantle structures, *J. Geophys. Res.*, *112*, –.

Thomas, C. W., Q. Liu, C. B. Agee, P. D. Asimow, and R. A. Lange (2012), Multi-technique equation of state for Fe2SiO4melt and the density of Fe-bearing silicate melts from 0 to 161 GPa, *J. Geophys. Res.*, *117*(B10), B10,206–.

Toby, B. (2001), EXPGUI, a graphical user interface for GSAS, *Journal of Applied Crystallography*, *34*(2), 210–213, doi: 10.1107/S0021889801002242.

Tsuchiya, T., and K. Kawamura (2002), First-principles study of systematics of high-pressure elasticity in rare gas solids, Ne, Ar, Kr, and Xe, *J. Chem. Phys.*, *117*(12), 5859–5865.

van Meel, J. A., L. Filion, C. Valeriani, and D. Frenkel (2012), A parameter-free, solid-angle based, nearest-neighbor algorithm, *J. Chem. Phys.*, *136*(23), 234,107–12.

Vinet, P., J. Rose, J. Ferrante, and J. Smith (1989), Universal features of the equation of state of solids, *Journal of Physics: Condensed Matter*, *1*, 1941–1963.

Xie, S., and P. J. Tackley (2004), Evolution of UPb and SmNd systems in numerical models of mantle convection and plate tectonics, *Journal of Geophysical Research: Solid Earth (19782012)*, *109*(B11).

You Shu-Jie, C. L.-C., and J. Chang-Qing (2009), Hydrostaticity of Pressure Media in Diamond Anvil Cells, *Chinese Physics Letters*, *26*(9), 096,202–.

Yuste, S. B., A. Santos, and M. L. de Haro (1998), Structure of multi-component hard-sphere mixtures, *J. Chem. Phys.*, *108*(9), 3683–3693.

Zhang, D., J. M. Jackson, W. Sturhahn, and Y. Xiao (2011), Local structure variations observed in orthoenstatite at high pressures, *American Mineralogist*, *96*(10), 1585–1592.


CALIFORNIA INSTITUTE OF TECHNOLOGY
GEOLOGICAL AND PLANETARY SCIENCES
Labor Bonus et Multus