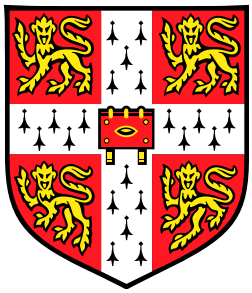
University of Cambridge

Aspects of structural and electronic disorder in network materials: approaches to simulation



Andrew D. Walkingshaw

Jesus College

February 6, 2007

A dissertation submitted to the University of Cambridge in partial fulfilment of the requirements for the degree of Doctor of Philosophy

Contents

1 Introduction **18**

2 Methods of materials simulation **23**

2.1 Introduction 23

2.2 Atomistic simulation 24

2.2.1 Common features of our atomistic simulations 24

2.3 Classical force-field simulation 28

2.3.1 Other interactions 29

2.4 Effective medium theory 30

2.5 Density functional theory 31

2.5.1 Born-Oppenheimer, the Hohenberg-Kohn theorem, and the Kohn-Sham equations 31

2.5.2 Pseudopotentials 36

2.5.3 The SIESTA method and code 37

2.6 Dynamic languages and XML; rapid algorithm design 39

2.6.1 Introduction 39

2.6.2 Python and high-level programming languages in numerical computation 39

2.6.3 Chemical Markup Language 43

3 Charge doping in WO_3 46

3.1 Introduction 46

3.2 Methodology 48

3.3 Calculated phase diagram 50

3.4 Discussion 55

3.5 Conclusions 63

4 Constrained Linear Maximization 64

4.1 Introduction 64

4.2 Constrained Linear Maximization 65

4.3 Constrained Linear Maximization: Further Details 78

4.4 Efficiency of Constrained Linear Maximisation 85

4.4.1 Step size 85

4.4.2 Bracketing tolerance 87

4.4.3 Reorientation algorithm 89

4.5 Tests on 2D potentials 90

4.6 Platinum (001) surface 98

5 Diffusion of Li^+ in crystalline silicates **100**

5.1 Introduction 100

5.2 Method 102

5.2.1 Supercell and potential 102

5.2.2 Event discovery 106

5.3 Quartz 108

5.3.1 Prior results 108

5.3.2 Calculated activated pathways 109

5.3.3 Stable states 111

5.3.4 Variations in channel width throughout the diffusion process . 119

5.3.5 Li–Al interactions 126

5.4 Cristobalite and Tridymite 134

5.4.1 Cristobalite 135

5.4.2 Tridymite 139

5.5 Conclusions 139

6 Li^+ diffusion in SiO_2 glass **144**

6.1 Introduction 144

6.2 Method 145

6.3 Calculated transition states 146

6.3.1 Stable states; distribution and comparison with crystalline phases 149

6.3.2 Transition state events; nature, and correlation with stable states 152

6.3.3 Role of network distortion in transition-state events 158

7 Conclusions **167**

List of Tables

3.1 Cutoff radii, $r(\zeta)$, for the basis sets corresponding to W and O species; as the basis set is double–ζ polarized, there are two basis functions (and therefore radii) per hypothetical atomic orbital, as specified by quantum numbers n and l. All radii are given in bohr. The starred entries are polarization orbitals. 49

3.2 Confinement radii for each of the s, p, d, and f channels of the W and O pseudopotentials. All radii are given in bohr. The W pseudopotential used also included a partial-core correction, with a radius of 1.30 bohr, according to the scheme of Louie et al [1]. 49

3.3 Lattice parameters, experimental, from my simulations, and from the work of de Wijs [2] for (hypothetical) cubic, low temperature monoclinic, room temperature monoclinic, high-temperature tetragonal and room temperature triclinic phases of WO_3. All distances are given in Å. Note that de Wijs' calculations were performed under the local-density approximation for monoclinic phases, and under the generalized gradient approximation for triclinic and tetragonal phases. 51

4.1 Parameters for the Gaussian functions added to the modified LEPS potential, after Henkelman and Jónsson [3]. 92

4.2 Results, stepsize=0.05; Note that transition states 1 and 2 are inaccessible from basin 3. 94

4.3 Results, stepsize=0.1; Note that transition states 1 and 2 are inaccessible from basin 3. 94

5.1 Parameters for the force-field used in these simulations. 103

5.2 Parameters for the α-quartz supercell, compared against experiment. As can be seen, c is 0.6% larger, and a 0.84% larger, in this doped cell than in the experimental undoped values at 20 K measured by Tucker et al; this is regarded as acceptable, certainly within the spirit of empirical-potential calculations. 104

5.3 Parameters for the tridymite supercell; experimental data from Kihara et al. [4] 105

5.4 Parameters for the α-cristobalite supercell; experimental data from Pluth et al. [5] 106

List of Figures

3.1 (a) Lattice parameters (for normalized cubic supercell with respect to the ϵ phase) of WO_3. Phase boundaries are denoted by solid lines, and labelled with their position (in electrons per centre); (b) Volumes of $2\times2\times2$ supercell of WO_3; (c) Bond lengths within WO_6 octahedra, all versus electron dopant concentration. 52

3.2 Offcentring in WO_6 octahedra; (a) along [001], (b) along both [001] and [110]-type directions 54

3.3 Order parameters for the doping-induced phase transitions in WO_3. (a) Pc – $P2_1/c$, in Å, where xy_1 and xy_2 are the inequivalent W-O bond lengths in the the $x-y$ plane; (b) $P2_1/c$ – $Pnma$, in degrees; (c) $Pnma$ – $P4/nmm$ (in Å); (e) $P4/nmm$ – $Pm\bar{3}m$ (in Å), where z_1 and z_2 refer to the two W–O bonds aligned predominantly with the z axis. Dashed vertical lines denote predicted phase boundaries, and are labelled with their positions in electrons per centre. 56

3.4 Polyhedral tilts in WO_3, as varying with dopant charge. The charges in each case are, in electrons per W atom, are; (a) 3/16, (b) 5/16, (c) 7/16, (d) 1/2, (e) 5/8, (f) 1; these correspond to, respectively, the monoclinic Pc phase; the $P2_1/c$ phase, approaching the $Pnma$ boundary; the middle of the orthorhombic phase; the $Pnma$ – $P4/nmm$ boundary; within the stability range of the tetragonal phase; and the aristotype. It is clear that the distortions between some phases are extremely small. This poses difficulties in ascertaining exactly where they are located in composition space. 57

3.5 Increase in energy per formula unit as a function of fractional change in W-O bondlength, for both W off-centring and Jahn-Teller distortion along a cubic axis. 60

3.6 Mean occupancy per W atom of $5d$ and $6s$ orbitals versus dopant charge. 61

4.1 Converged CLM runs on a 2D harmonic potential. 74

4.2 System trajectories for a quadratic/quartic 2D potential. 76

4.3 System energy profiles for a quadratic/quartic 2D potential, in one successful (the lower plot) and one unsuccessful (the upper) case. . . 77

4.4 Dimensional scaling of the interpolation strategy for a potential of the form $E(x_1, x_2 \ldots x_n) = (1 - x_1^2)^2 + x_2^2 + \ldots + x_n^2$ 79

4.5 Trajectories through the system for two different step sizes; dashed lines represent minimization in the tangent space. 86

4.6 As can be seen, the calculated path energy only falls at step 4; by this point, one has already overshot the transition state. 88

4.7 2D LEPS surface 91

4.8 2D LEPS surface with added Gaussians 93

4.9 Histogram of number-of-step distribution for runs from each basin; stepsize = 0.05 95

4.10 Histograms of (Number of force evaluations / 10) versus count for each of three bracketing tolerances; note the general shift of features of the histogram to the left. 95

4.11 Comparative positions of our calculated transition states on the modified 2DLEPS potential for bracketing tolerances of 0.01 and 0.04 units. 96

4.12 Transition states on the Pt(001) surface; from the top, a crowdion, a direct hop and a dimer exchange. 99

5.1 Connectivity of minima and transition states found in singly Li^+-doped α-quartz; diffusion up the c-axis is shown in the labelled vertices M1–M5 and T1–T5. Ellipses represent stable states, diamonds transition states; all energies are presented in eV. 110

5.2 Transition state T1 (0.750 eV), connected to M1 (the "wall" site, Figure 5.3). This is the activation barrier for diffusion through channel walls, and hence in [110]-type directions in this structure. 111

5.3 Minimum M1: the lowest minimum found on the c-axis mechanisml this is what has been dubbed a "wall" site, residing (as it does) in a [100]-channel wall. Connected to T1 (Figure 5.2) and T2 (Figure 5.4, the twofold site using Sartbaeva [6]'s notation.) 112

5.4 T2 (0.053 eV): transition state connecting M1 (the wall site, Figure 5.3) and M2, the twofold site (Figure 5.6.) 113

5.5 T5 (0.101 eV): this transition state, showing twofold linear coordination of Li^+ by Si^{4+}, connects M2 (Figure 5.6 with the equivalent twofold site directly below it (M5, not shown). 113

5.6 M2 (0.028 eV): connected to M1, the wall site, by T2 (Figure 5.4). This is Sartbaeva and Wells' "fourfold" site, as can be seen on the left. 114

5.7 T3 (0.222 eV): connects M2 (Figure 5.6, the fourfold site, with M3, the twofold site (Figure 5.8). This is the rate-determining step for [001] diffusion, with an activation barrier of 0.194 eV. 114

5.8 M3 (0.175 eV); this is Sartbaeva and Wells' "twofold" site, with the Li^+ ion directly between two oxygens, and is connected to the upper twofold site (M4, Figure 5.10) and the lower twofold site (M2, Figure 5.6) through T4 (Figure 5.9) and T3 respectively. 115

5.9 T4 (0.211 eV): this is the transition between twofold M3 (Figure 5.8) and and upper fourfold M4 (Figure 5.10). 116

5.10 M4 (0.053 eV): this is a fourfold site like M2 (Figure 5.6), connected to twofold M3 by T4 (Figure 5.9). 116

5.11 Approximate equilibrium constant between the two most stable Li^+–Si sites in α-quartz 118

5.12 Three equivalent O–O pairs in three channels of α-quartz - primary (containing Li), adjacent, and tertiary (two channels removed from Li). The two oxygens selected are the primary oxygens in the "twofold" coordination site described above. 120

5.13 Distance between three equivalent O–O pairs in three channels of α-quartz, for each of the six sites in our diffusion mechanism above; in order, sites T5, M2, T3, M3, T4 and M4. 121

5.14 Three equivalent Si–Si pairs in three channels of α-quartz - primary (containing Li^+), adjacent, and tertiary (two channels removed from Li^+). 122

5.15 Distance between three equivalent Si–Si pairs in three channels of α-quartz, for each of the six sites in our diffusion mechanism above; in order, sites T5, M2, T3, M3, T4 and M4. 123

5.16 Distance between O–O pairs across our six sites in the diffusion process, numbered 1-4, bottom-up (with the first stable state, M2 (Figure 5.6), lying between Pair 1.) 125

5.17 The four Si–Si atom pairs relevant to our diffusion process; these are numbered 1–4 from the bottom upwards. 127

5.18 Distance between Si–Si pairs across our six sites in the diffusion process, numbered as in Figure 5.17. 128

5.19 Histograms of number-of-states versus activation energy for our sample quartz configurations, with and without defect charge delocalization. 129

5.20 Absolute energy versus activation energy for our sample quartz configurations, with and without defect charge delocalization. 130

5.21 Connectivity of minima and transition states found in singly Li^+-doped, Al^{3+}-defect-delocalised α-quartz; diffusion up the c-axis is shown in the labelled vertices M1–M5 and T1–T5. Ellipses represent stable states, diamonds transition states; all energies are presented in eV. 131

5.22 Absolute energy versus Li–Al distance; the curve is a functions of the form $a + c/x$. Our fitted value for a is 0.886 eV; for c, -5.44 eV/Å. . . 132

5.23 Electrostatic-energy corrected activation energy versus uncorrected transition energy; note the lack of deviation from x=1, showing that (apart from three outliers) electrostatic Al–Li interactions have a negligible effect on the activation energies observed. 133

5.24 Transition-state electrostatic-energy correction (in eV) versus stable-state Li–Al distance (in Å). Note the three outliers. 134

5.25 CM1 (0 eV): this is the lowest cristobalite minimum found. 135

5.26 CT1 (0.665 eV): this transition state is connected to minimum CM1 in Figure 5.25 with an activation barrier of 0.665 eV. 136

5.27 CT2 (0.692 eV): this transition state is, like CT1 (Figure 5.26), connected to minimum CM1 (in Figure 5.25). The activation barrier for this process would be 0.692 eV. 136

5.28 CM2 (0.0271 eV): The second-lowest minimum found by my search in cristobalite. 137

5.29 CT3 (0.584 eV): This transition state is connected to minimum CM2 in Figure 5.28 by an activation barrier of 0.557 eV. 137

5.30 CM3 (0.794 eV): Note that in this minimum, and minimum CM4 (Figure 5.32, there is a reoriented chain of tetrahedra. 138

5.31 CT4 (1.360 eV): transition state connected to minimum CM3 in Figure 5.30. This move is similar to the CM1–CT1 (Figures 5.25 and 5.26) move above, apart from the reoriented chain; the activation energy is also similar, at 0.566 eV. 138

5.32 CM4 (1.024 eV): as in CM3 (Figure 5.30), a chain of tetrahedra is reoriented with respect to the minimum-energy structure. 139

5.33 CT5 (1.245 eV): transition state connected to minimum CM4 in Figure 5.32. This move, with a low activation energy, is particularly interesting - as it causes the reoriented chain of polyhedra to return to their normal position with an activation barrier of only 0.221 eV. . 140

5.34 TM1 (0 eV): Lowest tridymite minimum found. 140

5.35 TT1 (0.295 eV): transition state connected to TM1 (Figure 5.34), with an activation energy of 0.295 eV. It is difficult to see the lithium ion in this picture; it can be found just to the left of the centre of the image. 141

5.36 TM2 (0.206 eV): Of the two tridymite minima found, the higher in energy. 141

5.37 TT2 (1.283 eV): Transition state connected to TM2 (Figure 5.36), with an activation energy of 1.077 eV. 142

5.38 TT3 (0.628 eV): like TT2, connected to TM3 (Figure 5.36), but with a much lower activation energy of 0.657 eV. 142

6.1 Extended event network in glass (i). 147

6.2 Extended event network in glass (ii). 148

6.3 Histogram of stable states identified in our Li^+-doped glass sample. . 149

6.4 Spectrum of stable states and associated activation energies in our glass sample. 150

6.5 Distribution of activation barriers in SiO_2 glass, quartz, tridymite and cristobalite. 151

6.6 Coordination of Li by O (as defined in the text) in each of our stable states. 153

6.7 Coordination of Li by O in the stable state (as defined in the text) versus activation energy for each process 154

6.8 Relative Li-O and Li-Si coordination in stable states in our glass . . . 155

6.9 Mean Li-O(1) . . . Li-O(10) distances for three energy ranges 156

6.10 Change in coordination of Li by O between the stable and transition state (as defined in the text) versus activation energy, in eV 157

6.11 Normalized change in density of Si and O around Li between stable and transition states in glass 158

6.12 Distance between stable and transition state, excluding Li, for events in our sample, in Å, versus activation energy in eV 160

6.13 Magnitude of Si–O network distortion in the transition state (in Ang) versus the energy of the stable site (in eV) to which the transition state is connected, and from which the distortion is measured. 161

6.14 Number of atoms displaced by more than 0.1 Å in a transition-state event in our sample, versus activation energy in eV 162

6.15 Number of atoms displaced by more than 0.1 Å in a transition-state event in our sample, versus magnitude of Si–O lattice distortion (in Ang) 164

6.16 Number of atoms displaced by more than cutoff x (in Angstroms) versus cutoff; the curve shown has the equation 165

"Honour thy error as a hidden intention."

– Brian Eno, from the Oblique Strategies

To the many people who have supported me throughout my education thus far, to all of whom I owe a debt I can never repay; but firstly, and foremostly, to my parents.

Except as acknowledged above and throughout the thesis, the present work is my own, and except where specifically acknowledged in the text was not performed in collaboration; furthermore, it is not substantially the same as work I have submitted for any other degree or diploma at any University.

Summary

Disorder in solids is one of the critical problems facing exponents of simulation-based studies in the mineral sciences. In this thesis, I present some novel techniques which can be used to probe the structure and dynamics of such systems under simulation, and use the techniques thus developed to gain insights into some well-known and scientifically-important classes of materials.

In Chapter 2, I present an overview of current practice, considering both classical and quantum-mechanical methods. I also introduce some recent developments in computer science and software engineering which have begun to be taken up in scientific circles.

In Chapter 3, I present a study of charge doping in the electrochromic defect perovskite WO_3. By using a novel method to represent the doping process, I am able to neglect dopant-size effects and chemical disorder; this leads to the gaining of new insights into the sequence of dopant-induced phase transitions in this material.

In Chapter 4, an algorithm for discovery of transition states (named Constrained Linear Maximization) is introduced. This is a development of an older approach, called the Intrinsic Reaction Coordinate method; we present the key differences between this approach and other competing techniques, and test the new method on three energy surfaces; two test surfaces (one being a modified LEPS2D potential), and the (001) surface of platinum simulated using Effective Medium Theory.

In Chapter 5, Constrained Linear Maximization is used to undertake studies of lithium dopant motion in α-quartz, cristobalite and tridymite; in Chapter 6, it is used to study lithium motion in a fully-coordinated SiO_2 glass. In both of these chapters, the interplay between relaxation of the SiO_2 framework and the motion of lithium is considered extensively.

Chapter 7 summarises and draws conclusions from the work in the thesis, and presents possible directions for further research in the future.

Acknowledgements

I wish to acknowledge funding from the Engineering and Physical Sciences Research Council. I would like to thank Prof. Emilio Artacho and Prof. Martin Dove, my research supervisors, for their patience and forbearance; without their guidance none of this work would have taken shape. My colleagues past and present have all shaped this thesis for the better in many ways, whether they were aware of it or not! I owe particular thanks to Dr. William Lee for discussions on transition-state modelling, Dr. Asel Sartbaeva and Dr. Stephen Wells for sharing their thoughts on lithium dynamics, Profs. Jim Scott and Nicola Spaldin for their thoughts on perovskites, and Dr. Marivi Fernandez-Serra for her SIESTA expertise. Dr. Kostya Trachenko generated the glass structures used in the latter part of this thesis, and I thank him for his help. The work on transition-state search was begun during a visit to the Centre for Atomic Scale Materials Physics (CAMP) of the Technical University of Denmark, and I wish to thank Dr. Asbjørn Christensen and Prof. Karsten Jakobsen for their support and advice. The last part of this thesis has been written whilst working at the Unilever Centre for Molecular Informatics; some thought-provoking insights from my colleagues there, especially Prof. Peter Murray-Rust, have been invaluable.

Any mistakes remaining in this thesis are mine entirely, but there would be many more without my proofreaders; the aforementioned Dr. Lee, Dr. Peter Corbett, Dr. Toby White, Jim Downing and Richard Moore. Finally, I would like to thank my friends for their support and encouragement over the last four years; without them I would have never got this far.

Chapter 1

Introduction

Over the course of the last sixty years, the simulation of materials at equilibrium has progressed from proof-of-concept to fine art. There are many reasons for this; foremost is the staggering increase in available calculating power. The first stored-program computer, EDSAC, was switched on in 1949; just over fifty years later, in 2001, using a bespoke computer a molecular dynamics simulation on NaCl thirty-three million atoms in size [7] was performed. It is not hyperbole to say that a whole new approach to the physical sciences has been opened up by this new resource.

There have been huge advances in technique as well (and as a result); advanced classical forcefields, and quantum-mechanical methods like Density Functional Theory [8, 9], enable the accurate (and frequently inexpensive) prediction of the mechanical, electrical and spectroscopic properties of the vast majority of crystalline solids. However, not every class of problem is equally accessible to commonly-used simulation techniques. If the system is ordered – and therefore, symmetry can be exploited – or if the property one is interested in is time-independent, and does not require calculation of the motion of atoms within the system, then on the whole it is likely that it will be relatively straightforward to calculate by some means. Even if one needs to resort to a very high level of theory – quantum Monte Carlo [10], say

– it can be done, although one may need to find a very fast computer or be very patient!

Some problems, unfortunately, cannot be attacked so easily, and a lot of work in recent years has focussed on developing techniques to attack these more-challenging systems. In the case of disordered materials, the periodic symmetry of the crystal lattice is disturbed or destroyed; therefore, to simulate these one must either find some way of placing the disorder in a periodic context or give up and simulate a radically larger system, with a concomitant increase in computing power needed. Often this can make the difference between a given system being tractable to simulation or not – particularly when using quantum mechanical methods, as one must when interested in the electronic behaviour of a system, where typically the CPU time taken is of the order of the cube of the number of free electrons or worse.

This problem, serious as it is, still pales somewhat when compared to one of the most significant problems facing computational materials science; that of simulating the time-dependent dynamic properties of materials. Over the temperature range we are typically interested in, the nuclei behave classically and we can therefore use Newtonian mechanics to simulate their motion (in other words, use molecular dynamics). We can only use this, though, if we're likely to see the event; molecular dynamics only gives us a probe out to microsecond timescales, and often the processes we are interested in – diffusion, creep, and so on – take seconds to years. Diffusion in particular is a challenge for simulation; the characteristic rate for diffusion scales as $exp(-\Delta E/kT)$, where T is temperature and ΔE the activation barrier; thus, diffusion problems can easily vary from being on the picosecond-to-microsecond timescale, microsecond lengths being right at the limit of what can currently be simulated, to years. If one takes the natural timescale of an atomic motion as being one femtosecond, and the natural timescale of human experience as a second, three microseconds for an atom corresponds to ninety-five years on the human timescale - an entire lifetime!

There is a very, very large gap between the timescales we can simulate directly and

the timescales of the events we wish to probe; simple, unbiased molecular dynamics are clearly not the complete answer. In attacking this problem we, therefore, need to be a bit more ingenious; we need to consider what we can calculate, rather than what we can't. In the case of activated processes, harmonic transition state theory is very well-established; as long as we can map out the minima and connecting passes between them, we can determine the kinetics of the system. We don't need to observe events – just to know the prerequisites for them to occur.

In the present work, aside from an overview of the state of the art, I introduce novel approaches to the analysis of characteristic problems of both these kinds. Firstly, I consider chemical and electronic disorder, in the form of the effect of doping on the electrochromic ceramic tungsten trioxide. Perovskites are enormously important materials commercially, quite apart from the dominant role of $CaTiO_3$ in the mineralogy of the deep Earth; tungsten trioxide is used for its electrochromic [11, 12] properties, as mentioned – its optical properties vary with applied charge. It also undergoes a metal-insulator transition, and undergoes many structural phase transitions under temperature/pressure, including phases which exhibit antiferroelectricity.

Indeed, many perovskites exhibit remarkable behaviour, ranging from subtle elastic behaviour such as the inverse plateau effect in $[Sr,Ca]TiO_3$ [13] to their high-tech applications – for instance, the ferroelectric switching behaviour of thin films, which is used in the design of non-volatile memory devices and has been examined intensively [14]. Tungsten trioxide is a good prototype for doping behaviour in perovskite oxides; regarding it as ABO_3 with a vacancy on the B site, it is a very good example of a system displaying dopant-dependent Glazer [15] tilting transitions. Even more remarkably, at around 91 K, sodium-doped WO_3 displays islands of high-T_c superconductivity [16]; in its own right, it is a fascinating material.

By finding a way of representing doping as a continuous process without destroying the overall symmetry of the supercell and, thus, increasing the size of the calculation beyond reasonable limits, I map out the series of phase transitions which this material undergoes when doped with electron donors. These simulations shed new

light on experimental analyses of the structure of Na_xWO_3; this provides a potent example of the importance of simulation and experimental analysis proceeding in parallel. Furthermore, the chapter presents a general technique which can be used to simulate the dopant behaviour of a large part of this class of technologically-critical materials.

Secondly, the majority of the thesis is devoted to the prediction of diffusion behaviour in network minerals. A modified, improved derivative of one of the oldest transition-state search methods, the Intrinsic Reaction Coordinate method, is presented; we derive the fundamental mathematical behaviour of this method, the method of Constrained Linear Maximization, in great detail, and further test the method by briefly examining the complex problem of crowdion formation on the surface of platinum. After that, using this novel method I investigate the diffusion of lithium dopant ions through framework silicates; firstly in ordered polymorphs of quartz (quartz itself, tridymite, cristobalite), and culminating with a study of diffusion dynamics in silica glass.

The quartz–β-eucryptite system has been studied extensively by experiment and approximate simulation methods over the last couple of years [6, 17]; I improve on that work by using my exact (up to the limits of the potential) transition state search method. The results thus obtained provide insights into both the mechanism of transition in these materials and into the length-scale of strain associated with diffusion in network silicates of these types; given that β-eucryptite, which shares the β-quartz framework, is a one-dimensional fast-ion conductor [18], once again these materials are of technological value above and beyond their considerable theoretical interest. By studying the different polymorphs of quartz, it is possible to consider the effect of their differing structures on the activation barrier for lithium diffusion.

Using those results, we consider diffusion of Li^+ in silica glass. Although there have been many studies of more highly doped glasses using molecular dynamics, such as the lithium metasilicate glass Li_2SiO_3 [19], there have been very few of the low-doping limit; I analyse the behaviour found in the glass in comparison with the

behaviour seen in the regular silicate phases.

Therefore, the thesis builds towards this simulation of long-time-scale processes in a highly disordered system; that, in itself, shows how far the field has come. There is much further to go in the future, and that is addressed in the conclusion to this work, but I would like to invite you to survey this thesis and, hopefully, enjoy the view back over the land it covers.

Chapter 2

Methods for materials simulation; an overview

2.1 Introduction

In the present study, two major categories of methods for materials simulation have been used; first-principles electronic structure calculation through density functional theory [9, 8], as embodied in the SIESTA [20] code, and classical forcefield simulation as found in GULP [21]. In addition, some minor sections have been performed with effective-medium theory (through the EMT [22] forcefield included in the CAMPOS suite); I therefore give a brief overview of the common points of each of these methods, particularly as they relate to the transition-state discovery algorithm proposeed later in the thesis, of the key theoretical concepts underlying each of them, and of any important peculiarities of each of the methods embodied by the programs.

In the final section of this chapter, however, I digress into software engineering; some aspects of the work done in this thesis have touched on new approaches to the process of implementing novel algorithms and approaches to simulation, and it is worth spending a little time to introduce those.

2.2 Atomistic simulation

All of the simulations performed in this thesis can broadly speaking be considered as an operation on an explicit set of atoms with well-defined positions in a 3D vector space. Models (and codes) which implement this kind of simulation are generally referred to as *atomistic*.

2.2.1 Common features of our atomistic simulations

In essence, one can think of each model as specifying a Hamiltonian, taking the simulation cell (and the positions of all atoms contained therein, denoted $\mathbf{x}$ and consisting of a concatenation of the Cartesian positions of atoms 1 through n in sequence - $x_1, y_1, z_1, x_2 \ldots z_{n-1} x_n y_n z_n$) as a variable and returning an energy; calling this $H(\mathbf{x})$, the corresponding forces on the system is given by;

$$F = -\frac{\delta H(\mathbf{x})}{\delta \mathbf{x}} \tag{2.1}$$

The forces on any given atom are therefore just the relevant components of this vector. The transition state search algorithm developed later in this thesis depends solely on being able to evaluate the energy of a given configuration of atoms and the forces acting on that configuration; therefore any energy-evaluation function which fulfills these needs can be used with our methods, and indeed all of the atomistic models used in this thesis provide well-defined forces on atoms in this manner. This was, as might be expected, a primary criterion for their selection.

Periodic boundary conditions and supercells

The vector space in which all our simulations are performed must be finite, else it will not be computable. However, in a study of the bulk properties of periodic – and (in principle) infinitely large, in the case of a perfect crystal – materials, artificially truncating our simulations at some surface would either be unacceptably inaccurate or require too big, and therefore computationally expensive, a simulation. Instead, throughout this thesis the approximation of periodic boundary conditions (PBCs) has been used.

In short this approximation can be expressed as follows:

$$\mathbf{x} + u\mathbf{a} + v\mathbf{b} + w\mathbf{c} \equiv \mathbf{x} \tag{2.2}$$

where $\mathbf{x}$ is a position vector in our system, u, v and w are integers, and $\mathbf{a}, \mathbf{b}$ and $\mathbf{c}$ are three vectors defining the edges of our simulation box (termed the *supercell*). This imposes periodicity on our system; the supercell used may be as small as a single unit cell of the system under study, or as large as is required and is computationally feasible, but it must be borne in mind that this approximation will impose artificial long-range order on a system. If the supercell is sufficently small, interactions of atoms with their own periodic images – self-interaction – will occur, and may seriously impair the accuracy of the simulation.

Electrostatic interactions

Every realistic atomistic simulation must consider electrostatic interactions – between ions, between positively-charged nuclei and negatively-charged electrons, between electrons and electrons, or between nuclei and nuclei, depending on the level of detail embodied by the model. However, in all these cases it is necessary to sum

the electrostatic interaction over multiple species over some distance, and this is not easy; one must therefore find a way of achieving this.

The electrostatic interaction between two point charges, at positions $\mathbf{r_1}$ and $\mathbf{r_2}$ with charges q_1 and q_2, has the functional form:

$$E(\mathbf{r_1}, \mathbf{r_2}) = \frac{q_1 q_2}{(\mathbf{r_2} - \mathbf{r_1})} \tag{2.3}$$

$$F(\mathbf{r_1}, \mathbf{r_2}) = \frac{q_1 q_2}{|\mathbf{r_2} - \mathbf{r_1}|^3}\mathbf{r_2} - \mathbf{r_1} \tag{2.4}$$

where E is the energy of the interaction, and F the force exerted. This is just Coulomb's law. However, this needs to be summed over, in principle, every pair of atoms in the system, and further to this, under periodic boundary conditions these interactions will extend over more than one supercell; neglect of these long-range interactions can lead to significant errors in the calculated properties - including the energy and forces of the system - and as such, it is critically important that an accurate method is used for the summation of long-range interactions.

The Ewald sum

For this purpose, the Ewald summation [23] is ideal. We decompose the Coulomb interaction into two parts - a strictly short-ranged interaction, and a long-range interaction which is therefore short-ranged in reciprocal space. Both of these therefore have convergent sums in the basis where they are short-ranged; one can then set a cutoff in real and reciprocal space which allows us to calculate the Coulombic interaction up to acceptable accuracy.

The full derivation can be found in standard textbooks [24], but there are certain points of interest which should be expanded on here. In principle, the derivation proceeds by noting that:

$$\frac{1}{r} = \frac{2}{\sqrt{\pi}} \int_0^\infty \exp(-r^2\rho^2) d\rho \tag{2.5}$$

and, by substituting this lemma into the Coulomb interaction, separating the Coulomb interaction into two integrals on different length scales, as mentioned above. The short-range interaction falls to zero rapidly; the long-range interaction can be converted, by Fourier transformation, into a short-ranged interaction in reciprocal space (here, E_{recip}). The short-ranged interaction, E_{real}, can be summed over directly.

$$E_{recip} = \left(\frac{1}{2}\right) \frac{4\pi}{V} \sum_{\mathbf{G}} \frac{\exp\left(-\mathbf{G}^2/4\eta\right)}{\mathbf{G}^2} \times \sum_i \sum_j q_i q_j \exp(-i\mathbf{G}r_{ij}) \tag{2.6}$$

$$E_{real} = \sum_{l=0}^{\infty} \sum_{i,j} \frac{q_i q_j \mathrm{erfc}\left(\eta^{1/2} r_{ij}\right)}{r_{ij}} \tag{2.7}$$

Here, erfc is the complementary error function. r_{ij} is the scalar distance between atoms i, in the reference unit cell, and j, in the l-th unit cell. $\mathbf{G}$ is a reciprocal lattice vector. η is a free parameter, which defines the distance for which interactions are regarded as either short-range or long-range.

There are two further issues thrown up by this, however. Firstly, in the cell where $l = 0$, i and j are the same atom. Therefore, the summation form above contains a spurious self-interaction, which is necessary to make the reciprocal-space transformation; this must be accounted for, and we can do so by by taking the limit of the above equations as $\mathbf{r}$ tends to zero. The resultant correction is known as the *self term*.

Secondly, we need to consider what happens when $\mathbf{G} = 0$. It turns out that in the limit of $\mathbf{G}$ tending to zero, the value of the long-range term depends on the direction $\mathbf{G}$ tends from; the $\mathbf{G} = 0$ term is therefore regarded as a macroscopic

field, and neglected in the energy calculation. However, this is particularly relevant in the case of the SIESTA code; it formulates the nucleus-nucleus interaction in a manner which is strictly short-ranged [20], and therefore the Ewald summation is only used in the context of the electronic calculation, Here, η can be set to zero and the entire calculation performed in reciprocal space; therefore, this provides a mechanism by which a homogeneous background charge can be imposed, as by construction it corresponds to the $\mathbf{G} = 0$ term in reciprocal space. We use this result when considering WO_3 later in this thesis.

The Ewald sum scales, computationally, as $O\left(N^{3/2}\right)$ with correct choice of η; this is covered in some detail in Gale's paper on GULP [21] and the GULP manual [25]. The short-range potential will converge rapidly in real space, and the long-range in reciprocal space; hence, both can be truncated with acceptable loss of precision.

2.3 Classical force-field simulation

In what I have termed a *classical* simulation, the system is treated as being composed of point atoms of defined charge. H($\mathbf{x}$) is then composed of two parts;

$$H(\mathbf{x}) = S(x) + C(x)$$

where $S(x)$ is the electrostatic part of the energy, and $C(x)$ is the contribution from additional terms; in essence, the bonding interactions between atoms in the system are broken down into an electrostatic term and everything else. Consider each of these separately; for the electrostatic interactions one can use the Ewald sum, as above, but that leaves all other forms of atom-atom interaction.

2.3.1 Other interactions

In this thesis, in the materials simulated using a classical forcefield – aluminosilicates – there is an interplay between electrostatic factors, particularly with respect to dopant species, and the predominantly covalent interactions within the (Si, Al)–O framework.

To represent these interactions, the Buckingham potential is used, which has the functional form

$$V(r_{ij}) = A_{ij} e^{\left(-\frac{r_{ij}}{\rho_{ij}}\right)} - \left(\frac{C_{ij}}{r_{ij}}^6\right)$$

where i, j are the indices of the two interacting atoms. This is physically motivated, in that the two parts of the expression represent short-range repulsion and long-range attraction respectively, but is essentially a phenomenological expression; values are fitted to A, C and ρ in order to best reproduce the structure and physical properties (bulk modulus, elastic constants, etc) of a range of structures related to the material or materials which one wishes to study.

Many other potentials with different functional forms exist, such as the well-known Lennard-Jones potential, and are commonly supported by a wide range of codes (including GULP), but are not used in the present study.

In the present work, parameterisation developed by Calleja et al. [26] is used, adding Li–Al and Li–Si interactions (by fitting to α–spodumene) to the forcefield developed for silica glass by van Beest et al. [27]; I go into this in more detail in due course. One weakness of which we should be aware is that the forcefield neglects the polarizability of oxygen; it is possible to extend this class of simulation by representing oxygen as a positively-charged point core connected to a spherical shell of negative charge by a fictional spring, but there are a number of reasons why this is not necessarily a desirable approach. Firstly, it is somewhat more complicated to deal with conceptually when designing algorithms, as the degrees of freedom of the nuclei and

shells are necessarily very strongly coupled. Secondly, it is more computationally intensive, as it would add three extra degrees of freedom for each oxygen atom in our system. Thirdly, in any case, the potentials used are, even neglecting polarizability, well-characterised and well-understood, and they give at least qualitatively accurate physical behaviour for the systems this study is interested in.

2.4 Effective medium theory

The effective-medium theory [22] (EMT), as developed by Karsten Jacobsen and collaborators, is in some sense a middleground between entirely empirical, phenomenological forcefields (like those found in GULP) and fully self-consistent quantum-mechanical methods. The total energy of any particular atom in a system is determined by its local environment – what it bonds with or interacts with electrostatically. Therefore, a reasonable approximation is to consider the atom as being embedded in a homogeneous free electron gas, constructed from the electron density of neighbouring atoms;

$$\Delta E_{tot} = \sum_{i=1}^{N} E_{c,i}(\bar{n}_i)$$

where $E_{c,i}(\bar{n}_i)$ gives the embedding energy of atom i in a gas of density $\bar{n}_i$; $\bar{n}_i$ is given by the average density over the neighbouring atoms over the volume of space occupied by atom i.

One can therefore attempt to propose a universal energy function $E_{c,i}(n)$, describing to a first approximation atom i's bonding; this will only hold in situations where the electronic structure of the system is reasonably well-described by the free electron gas, which (in short) means conductors - particularly $3d$ and $4d$ transition metals.

The full derivation of EMT is some distance outwith the scope of this thesis; in the present work, it has been used solely as a test potential for calculations on the Pt(001) surface, for which it is at least qualitatively reasonable. The EMT potential used is embodied in the ASAP code, part of the CAMPOS suite of materials-simulation tools.

2.5 Density functional theory

Of the many quantum mechanical methods available to researchers in the solid-state sciences, arguably density functional theory (DFT) is the most popular. In the form used in this research, it possesses an excellent balance between computational cost and accuracy of calculation, and is as such ideal for use as a probe of the behaviour of valence electrons.

2.5.1 Born-Oppenheimer, the Hohenberg-Kohn theorem, and the Kohn-Sham equations

In the previous section, it was assumed that atoms were embedded in a homogeneous electron gas (in the case of effective medium theory), or that electrons and nuclei were effectively fused into single entities (in classical lattice dynamics). In both these cases, therefore, the effective position of electrons is defined by the nuclei. To move to a fully quantum-mechanical treatment of the interactions between electrons and nuclei, we must therefore consider the motion of the electrons – any model without this cannot be better than phenomenological.

The quantum dynamics of the electrons and nuclei must be faced. This is a substantial problem; however, it can be simplifed greatly by using the Born-Oppenheimer [28] approximation. Note that nucleons (protons and neutrons) are around 1836 times the mass of the electron; therefore, in response to a thermal pertubation of the

system, the electrons reach thermal equilibrium much faster than the nuclei. It follows that, under conditions relatively close to equilibrium, one can consider the motions of electrons and nuclei separately; they move on different timescales, and therefore (in particular) one can consider the motion of the nuclei in an effective field of electron density defined by whatever the ground-state is for the electrons at that instant in time. In a nutshell, we assume that the electronic system is always at equilibrium. As a further approximation, the nuclei can be treated classically; quantum behaviour of nuclei is negligible under the conditions (and for the species) being studied in the present work.

The most popular quantum mechanical method by some distance for solid-state simulation is density functional theory (DFT). The central theorems of DFT were developed in two famous papers by (firstly) Hohenberg and Kohn [9], and (secondly) Kohn and Sham [8]. Making the Born-Oppenheimer [28] approximation, Hohenberg and Kohn proved the existence of a unique functional which takes the ground-state electron density, $n(\mathbf{r})$, as its free variable and returns the energy of the system;

$$E = E[n(\mathbf{r})] \tag{2.8}$$

and that the ground-state electron density can be obtained variationally. It follows from this that the ground-state density also uniquely defines the many-electron ground-state wavefunction;

$$\Psi_0 = \Psi_0 \left[n_0\right] \tag{2.9}$$

Therefore, any observables of the ground-state wavefunction can be rewritten as functionals of the ground-state electron density – most notably, the ground state energy. Given such a functional, one could therefore invoke the variational prin-

ciple to directly solve for the ground state electron density. In principle we have collapsed the many-variable wavefunction problem into the single-variable problem of the ground state electron density – but the Hohenberg-Kohn theory only proves the *existence* of an energy density functional, rather than giving us any information as to what it is!

Indeed, the exact form of the energy functional is not known; the best we have are various approximations to the exact functional. The form of the approximations used date back to earlier non-self-consistent approaches (such as Thomas-Fermi and the X_α method), but were developed in their present form in the work of Kohn and Sham.

Returning to the Hamiltonian - the Schrödinger [29] equation may be written as follows:

$$H\Psi = [T + V + U]\,\Psi$$

where T is the kinetic energy operator, U is the electron–electron interaction, and V is the static (thanks to Born–Oppenheimer) potential in which the electrons are moving. T and U here are system-independent, whereas V is system–dependent - however, once we fix a particular system, both T and V are straightforward to calculate. U, which is the difference between the single-particle and many-particle case, is the part of the many-body Hamiltonian which causes us difficulty. One approach is to expand the wavefunction in terms of single-electron wavefunctions; the simplest approach along these lines is Hartree–Fock theory, which exactly includes electron exchange but neglects electron correlation; so-called post-Hartree-Fock methods (typified by the method of Configuration Interaction) include this effect, but at a computational cost which scales substantially worse than the cube of the number of electrons in the simulation, rendering them impractical for simulations of the size undertaken here. Configuration Interaction, in particular, scales exponentially in its

full form, as it amounts to a combinatorial expansion of the possible states of the system.

Instead, using Hohenberg-Kohn, we know that the energy can be rewritten in terms of the ground-state electron density, n;

$$E[n] = T[n] + U[n] + V[n]$$

The external potential is expressible directly in terms of the ground-state density;

$$V[n] = \int V(\mathbf{r}) n(\mathbf{r}) d^3\mathbf{r}$$

One therefore needs to minimize this expression with respect to $n(\mathbf{r})$, the ground-state density (invoking the variational principle, as mentioned above). Consider a fictitious non-interacting system with energy E_s and corresponding wavefunction Ψ_s;

$$\langle \Psi_s[n] \left| T_s[n] + V_s[n] \right| \Psi_s[n] \rangle$$

Setting $V_s[n] = V + U + (T - T_s)$, it follows that for this non-interacting system, the ground-state electron density $n_s(r)$ will be equal to $n(r)$, the ground state density of the interacting system; one can therefore solve the non-interacting system self-consistently using iterative, self-consistent field methods, and obtain the ground-state electron density of the fully-interacting, real system. Expanding the non-interacting system in terms of single-electron functions ψ_i, known as molecular orbitals, we write;

$$-\left[\frac{\hbar^2}{2m}\nabla^2 + V_s(\mathbf{r})\right]\phi_i(\mathbf{r}) = \epsilon_i\phi_i(\mathbf{r})$$

These are known as the Kohn-Sham equations. Expanding V_s;

$$V_s = V + U + (T - T_s)$$
$$V_s = \int V(\mathbf{r})n(\mathbf{r})d^3\mathbf{r} + \int \frac{e^2 n_s(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{XC}\left[n_s(\mathbf{r})\right]$$

where the first term is the interaction of the electron density with the static external potential of the atomic nuclei (which is known), the second term is the Coulombic electron-electron interaction (also known exactly), and the third term is the so-called exchange-correlation term; this third term includes all the many-body interactions, but is not known exactly. Therefore, one must approximate V_{XC}.

There are many approximations used for this part of the functional. The most common strategy to use is to start, at least, with a parameterization of the exchange and correlation of the free electron gas, which can be calculated to a very high degree of accuracy using methods such as quantum Monte Carlo [30]; if one makes the approximation that the exchange and correlation energy at any point in space is a functional of the density at that point in space only – the *local-density approximation* [8] – we therefore have developed an approximate functional which can actually be solved. Extended approximations, either by hybridising the functional with exact exchange from Hartree-Fock (as in the B3LYP [31, 32] functional, very popular in quantum chemistry) or by extending the local-density approximation by considering the gradient of the electron density as well (*generalized gradient* approximations), exist, but because there is no systematic way of improving the accuracy of a functional for any given system – which is a major drawback compared

to the post-Hartree-Fock methods mentioned above – one cannot assume that any given functional will give one better results than any other, and it is furthermore impossible to estimate the errors arising from any given functional without recourse to calculations using another method. The Perdew-Zunger [33] parameterization of the local-density approximation is used for all quantum mechanical simulations within this thesis; experience derived from many prior simulations by other authors indicated that it was adequate for present needs.

2.5.2 Pseudopotentials

In any quantum-mechanical computation, the amount of CPU time taken for a calculation to converge will be correlated, to some order, with the number of free electrons; in the case of conventional DFT, the time taken scales as $O(n^3)$. As such, the fewer electrons in the calculation, the better from the perspective of performance. For heavy atoms, such as tungsten, the majority of the electrons play no part in bonding; only the $6s^2$ and $5d^4$ electrons are relevant in this particular case.

Therefore, it would be good to design our simulation in such a way as to not include these less-important electrons in our wavefunction. The pseudopotential approximation provides us with a way of achieving this. In essence, one divides the electrons on a given atom into two groups, the *valence* and the *core*, and replace both the potential arising from the bare nucleus of the atoms and the core electrons with a single potential obtained from a calculation of the atom in isolation; thus, in the core region, our valence wavefunction is replaced with a (smoother) pseudo-wavefunction, with the restriction that the eigenvalues of the Schrödinger equation for the pseudo-wavefunction acted on by the pseudopotential must be identical to those for the full potential and true wavefunction. This approximation is acceptably accurate as long as the core species really are uninvolved in the bonding scheme.

In the case of tungsten, the groundstate electron configuration is $[Xe]4f^146s^25d^4$, where $[Xe]$ represents the closed-shell electronic configuration of xenon. Here, the

$4f$ orbital is closed, and therefore can be regarded as part of the core; our valence is therefore six electrons, as mentioned earlier, rather than 74. An all-electron calculation for an isolated atom of W (in the gas phase) is performed, and at some radial cutoff – which must be substantially shorter than any bonding interactions – the pseudopotential is constructed so that this fictional potential matches the real potential which would be observed at this cutoff, and so that the pseudopotential has continuous derivatives at this point. Furthermore, the pseudopotentials used in the present study are *norm-conserving* - in other words, up to this cutoff, r_c, the following integral is identical over both the real and pseudo-wavefunctions;

$$\int_0^{r_c} \psi^*_{AE}(\mathbf{r})\psi_{AE}(\mathbf{r})d\mathbf{r} = \int_0^{r_c} \psi^*_{PS}(\mathbf{r})\psi_{PS}(\mathbf{r})d\mathbf{r}$$

Therefore, outside r_c (and therefore in the region where bonding occurs), the real and pseudo wavefunctions should be identical, and will generate identical charge densities.

There is much more which could be written on this subject, as with much else in this section, but pseudopotentials are not the main thrust of our research; in the present work, norm-conserving Trouiller-Martins [34] pseudopotentials were used, with partial core correction, in their fully non-local form, These were calculated by the `atm` code of Alberto Garcia. It is necessary to use the same exchange and correlation functional to generate the pseudopotential as is used in the simulation, and therefore here the Perdew-Zunger LDA has also been used.

2.5.3 The SIESTA method and code

All the DFT simulations in the present thesis were performed using the SIESTA (Spanish Initiative for Electronic Simulation with Thousands of Atoms) code, which itself implements the SIESTA [20] method. SIESTA can perfom DFT simulations

in two different modes; using a conventional self-consistent field/diagonalization method to solve the Kohn-Sham equations, or by direct minimization of a modified energy functional [35]. In the latter case, the entire calculation simulation is linear-scaling – the amount of time it takes is a linear multiple of the number of electrons in the system, whereas in the former case the performance of the calculation scales as the cube of the number of electrons, but with a smaller prefactor. (However, it should be noted that in both cases, the construction of the exchange and overlap matrices takes place in linear time – this arising from the use of a strictly localised basis set.) In terms of performance, therefore, there is a crossover for medium-size systems (low hundreds of atoms); as these are the size of systems I am investigating in this thesis, the conventional solver was used for the simulations in the present study.

As mentioned above, the heritage of SIESTA as a linear-scaling code shows in its choice of basis set; it uses a strictly-localised set of pseudoatomic orbitals, originally proposed by Sankey and Niklewski [36] for the FIREBALL code, and performs overlap (and some of the Hamiltonian) integrals between these non-orthogonal orbitals on a real-space grid. The effect of that is that when a given atom moves by a distance smaller than the space between two points on the grid, the points where the nucleus (and, hence, the pseudopotential) and the orbitals overlap this real-space grid change, leading to a periodic variation in the value of the integrals depending on the grid; this variation is evident in both the total energy of the calculation and the atomic forces, and is known as the *eggbox effect*.

The use of a real-space grid also affects choice of pseudopotential for a SIESTA simulation; the smoother the pseudopotential used, the smaller the eggbox effect and the coarser a real-space grid can be used, reducing the computational difficulty of the calculation.

2.6 Dynamic languages and XML; rapid algorithm design

2.6.1 Introduction

It is fair to say that most scientists lag some way behind the cutting edge of software engineering technique. There are very good reasons for this; at present, the *lingua franca* among scientific programmers is usually one of the dialects of FORTRAN. In its favour, it has excellent library support and very good performance for its intended purpose, heavyweight numerical algorithms, but by the standards of most other commonly-used programming languages has I/O and network-programming facilities which vary from somewhat lacking to non-existent.

Rather than attempting to extend the language to address these shortcomings, in recent times there has been a move towards mixed-language development. FORTRAN has undoubted strengths - excellent compilers, support for parallelization and vectorization, around thirty years of experience in performance tuning - and it makes sense to make use of that, whilst writing non-performance-critical code in a more expressive, higher-level language with the concomitant advantages in concision, understandability and maintainability of code.

In this chapter, two areas are briefly introduced in which new technologies have had a substantial impact on the work in this PhD.

2.6.2 Python and high-level programming languages in numerical computation

Python [37] is an interactive, object-oriented programming language originally designed by Guido van Rossum at CWI in the Netherlands in the early 1990s. From that base, it has become one of the most popular scripting languages (alongside

Perl and TCL), particularly on Unix systems - however, it has been ported to every major (and most minor) operating systems, from the Palm Pilot upwards.

Unlike C and Fortran, but similarly to Perl and TCL, Python is often classed as an *interpreted* rather than *compiled* language; one runs Python programs directly with a Python interpreter rather than compiling them to machine language first with a compiler. The advantage of this is that you no longer need to recompile a program after making a change in it; the trade-off for this gain in flexibility is that interpreted languages tend to have poorer performance than compiled ones.

In fact, Perl and Python are typically compiled on-the-fly to an intermediate representation (called a *bytecode*), and it is this bytecode which is actually executed by the interpreter. Two other languages which take this intermediate-bytecode approach are Java and Visual Basic, but in both of those the bytecode compilation and interpretation are decoupled.

Although Python has remarkably good numerical performance for a language of its class (particularly when using the Scientific Python [38] extension libraries, which provide Fortran-like array operations), it cannot directly compete with Fortran in terms of raw computational power. However, it excels in other areas; by writing the speed-critical parts of one's code in a compiled language, and the remainder in Python, one can gain the advantages in simplicity and expressiveness arising from Python's richer libraries, syntax, and object system as well as the ease-of-development of an interpreted language. In particular, in the present work, there is a very clear divide between the computationally intensive part of our work (evaluation of the total energy and forces of a given atomic configuration) and everything else, which makes this kind of strategy extremely attractive.

One project which uses this to its advantage is the CAMPOS [39] suite of open-source software, primarily written by a group of researches at the Technical University of Denmark in Lyngby. The model taken by CAMPOS is to encapsulate the energy-evaluation method in a Python object, and to implement everything else (optimization algorithms, workflow control, input/output, etcetera) in Python. This is

done in such a way that the energy evaluator conforms to a generic interface regardless of what it actually is - it can be changed from a completely fictional potential to a cutting-edge ab initio electronic structure package without the need to make any alterations to any other algorithm in the simulation.

Therefore, the energy evaluator is reduced to being purely that; it is, effectively, a function which takes a parameterized model of the system (atomic positions, unit cell size, any necessary forcefields or pseudopotentials, etcetera) and returns the energy of the system and the forces acting on any atoms within it. Control is then returned to the Python program, which acts on this new information, typically by deciding where to move atoms to for the next iteration; this can easily be written in Python, as in general this process is many orders of magnitude less computationally intensive than the energy evaluation itself. Therefore, the greater performance of Fortran is irrelevant in this case; the optimization algorithm being used takes an insignificant amount of the execution time of the entire run, so even if it takes ten times longer than it would written in a more efficient language, that is more than made up for by the time saved by the programmer by using a higher-level language.

Decoupling the energy-evaluation step from the optimization algorithm in this way is extremely powerful for several reasons.

Firstly, it permits one to develop algorithms in a completely general, code-independent way. In the case of the present work, it allowed me to write, again in Python, low-dimensional analytically soluble energy surfaces which could be used to test my trial transition-state discovery algorithms on, and then subsequently use *exactly* the same code in order to try the algorithms on real systems using multiple different simulation methods (classical forcefields, effective medium theory, up to DFT). This dramatically speeds the development and debugging of implementations of algorithms, and encourages a sort of "experimental programming" which would be much more arduous if one needed to go round the compile-debug-run-recompile cycle.

Secondly, it is good software-engineering practice; by encapsulating the details of the atomistics in one place and the algorithm in another, it is easier to edit and

correct either one in isolation. This, again, has obvious benefits in debugging and ease of development.

Thirdly, it encourages code reuse; implementing a good, robust minimizer, molecular dynamics, or similar is very difficult and detail-intensive. It is very desirable not to have to reinvent the wheel each time one writes a new code; by modularizing the design of a simulation in this way, one has an instant library of battle-tested routines which can be pulled together on the fly to try out a new idea.

Fourthly, although this is not attempted in the present work, it becomes substantially easier to write the necessary framework for a mixed-evaluator run (say, preconditioning a system by energy minimization with a classical forcefield method before switching to using an ab-initio DFT code for final optimization); the same atomic configurations, data, and algorithms can be used immediately within the same run, without having to interrupt a run and transcribe the output of one program into the input format of another (either by hand or programmatically). This both makes this kind of approach much more practical and makes it much less error-prone; even conversion programs can fail - for example, in the event of a seemingly minor change in the output format of one of the codes being used.

In the present work, calculations have been performed using simple analytic potentials, the ASAP effective medium theory code, the GULP lattice-dynamics code, and the SIESTA LCAO-DFT ab-initial total energy program; exactly the same optimization algorithms were used in all of these cases. The interface to the ASAP code (and the ASAP code itself) comes with the CAMPOS distribution; I modified the SIESTA interface in CAMPOS heavily for my needs, as will be touched on in the next section, and the GULP interface was written by the present author.

It is particularly telling that the versions of GULP and SIESTA used in this work were written in Fortran 90; ASAP in C++; and analytic potentials in Python – yet using the CAMPOS framework lets us switch between any and all of these transparently, without adapting the algorithms used. The immediate benefit of this is that one can be certain that there are no differences in the implementation of our

algorithm when used with each of these force evaluators; traditionally, it would have been necessary to add code to each of these programs to carry out the algorithm separately, opening up a much-increased risk of introducing new bugs in each of these reimplementations. Also, it is much less effort - both because there is no need to restate ourselves initially, but moreso in that any optimizations or bugfixes found with one force evaluator are immediate available to all of them, and do not need to be carried across to each code by hand.

2.6.3 Chemical Markup Language

The Chemical Markup Language (CML) [40] is a particular example of a family of languages called XML [41] – Extensible Markup Languages - which also includes, for example, XHTML [42] (the XML representation of the Hypertext Markup Language, in which webpages are written).

These languages allow semantic markup of a document. For example, in the thesis you are currently reading, text can be classified into various kinds - headings, footnotes, figure and table captions, body text, and so on – and each of these has a different semantic context. This is independent of the layout of the document; the semantics may be used to help define how an essay is typeset or not, but regardless of that they exist independently of any particular visual representation.

In the case of CML and this project, there is a set of core concepts (atoms, unit cells, and so forth and so on) which are common to all of the simulation codes used – basically, the same core concepts which are encapsulated in the CAMPOS framework. Therefore, for CMLised codes, it becomes much easier to write the "glue code" – particularly, output parsers – necessary to tie it in to our simulations.

The critical problem is that many programs, particularly those written in FORTRAN, have somewhat arcane input and output formats; FORTRAN 77-style fixed-column layouts (where misplaced whitespace can break your input file) are still

common, for example, and the documents are not structured so as to be machine-readable. In practice, one has to write a parser for the particular idiosynracies of every program's input and output formats (usually by matching specific strings in the output using regular expressions) – and any change, even a seemingly very minor one, can break this in such a way as to break one's programs (or, worse, cause them to give inaccurate or misleading results). Also, as every code has its own conventions, the glue-code written for one case is rarely, if at all, reusable. Essentially, one has to start from scratch with every program.

XML, whilst not being a standard file format in the conventional sense, does have a standard structure; it is composed of pairs of tags, of the form

```
<tag attribute1='value' attribute2='value'> data </tag>
```

where tags may be nested (i.e. `<t1><t2></t2></t1>` is valid, but `<t1><t2></t1></t2>` is not. A document which conforms to this (and the other rules of XML) is deemed to be "well-formed".

This property means that regardless of the meaning assigned to tags, which is defined in CML, the mechanical process of reading an XML document – essentially, a tree structure – can be abstracted into a library very easily. Unsurprisingly, this has happened for every major programming language. This removes most of the difficulty alluded to above; whitespace is not syntactic, for instance, so only changes to the definition of the Chemical Markup Language need be accounted for, not trivial changes in output format.

Two of the codes used in this work, GULP and SIESTA, can output CML, although neither can read it; in practice, however, input file formats are substantially easier to construct programmatically than arbitrary output file formats are to read. One of these (GULP) was not previously supported by the CAMPOS framework, and therefore it was necessary to write an interface to it; this was radically simplified by the use of CML; also, the parser thus written for GULP's CML output is easily portable to SIESTA's CML output, and indeed to any future code using the

language.

CML's value lies in this extension of the toolset available; whereas now, visualization programs (for instance) are often closely tied to one code in particular, meaning that one needs to learn how to drive the software to interpret results for each code independently, CML points towards a future where common semantics are defined across a range of codes; this will make it easier to compare results, exchange and cross-check work, and set up (again, for example) runs which switch between completely different levels of theory at different stages of the calculation.

Chapter 3

Charge doping in WO_3

3.1 Introduction

Tungsten trioxide, WO_3, is a technologically-significant ceramic: it is electrochromic [11], and the possibility of ion intercalation and deintercalation gives rise to several potential applications in devices (for instance, as a cathode of rechargable batteries [43]). In particular, it is often used in optical applications due to the fact that the color can be changed by doping with electrons. (This was first observed in 1815 by Berzelius [44] in H_xWO_3). Bulk stoichiometric WO_3 is yellow-green in hue, but the sodium-doped tungsten bronzes, Na_xWO_3, exhibit most colors of the visible spectrum on varying Na concentration [45]. In addition, the optical absorbance and reflectivity of the material can be modulated by injection or extraction of electrons and ions, giving excellent control over and tunability to the optical properties [11, 12].

In addition to strongly affecting the color of the material, the incorporation of electron donating ions also has a strong effect on the structure. For example, complex structural behaviour is observed on doping with electron donors, such as Na, Li or H. Such dopants occupy the vacant perovskite A site and therefore they can be easily introduced over a wide range of concentrations, from trace concentrations in

the so-called ϵ phase up to the limiting case of the tungsten bronze, $NaWO_3$. In $NaWO_3$, there is one donor per formula unit, and an aristotypic perovskite structure is adopted [46]. Its structure is the same as that of ReO_3, with both solids retaining the ideal cubic perovskite structure at all temperatures. Indeed, assuming that in $NaWO_3$ the valence electron of Na is donated into the *5d*-type conduction/antibonding band of WO_3, the two materials are isoelectronic. In contrast, WO_3 displays both off-centring of the W ion from its ideal centrosymmetric position, and Glazer [15]-type tilting transitions. Studies on intermediate concentrations of Na_xWO_3 [47] suggest a complex phenomenology, involving a consistent reduction in the degree of polyhedral tilting with increasing dopant concentration. Notably, Clarke observed that the room-temperature structures of Na_xWO_3 ($0.62 < x < 0.94$) are slightly distorted from the high-symmetry $Pm\bar{3}m$ aristotype, and proposed a preliminary phase diagram to explain the observed diffraction data [47].

There have been several prior computational studies on WO_3, carried out at various levels of theory. Here I summarize the results of the density functional theory (DFT) studies from the literature. Perhaps the most relevant for this study is the work by Corà et al. [48, 49] on the electronic structure of cubic WO_3, ReO_3 and $NaWO_3$. Using the full potential linear muffin tin orbital method, with the local density approximation, they showed that the band structures of the three compounds are very similar, with the extra electron in ReO_3 and $NaWO_3$ occupying the antibonding conduction band and decreasing the metal-oxygen bonding strength. They also showed that displacement of the transition metal from its centrosymmetric position towards an oxygen ion causes increased overlap of the transition metal $5d$ t_{2g} orbitals at the bottom of the conduction band with the O $2p$ orbitals at the top of the valence band. This in turn lowers the energy of the top of the valence band, and raises that of the bottom of the conduction band, explaining why metallic ReO_3 and $NaWO_3$ remain cubic, whereas "d^0" WO_3 [50] has an off-centring distortion. Hjelm et al. [51] used the same method to establish that, in $LiWO_3$ and $NaWO_3$, rigid band filling of the WO_3 conduction band occurs, whereas in HWO_3 the hydrogens form hydroxide units with the oxygen atoms and change the electronic structure. de Wijs

et al [2] calculated the electronic properties of the various experimental structural phases of WO_3 using a plane wave ultra-soft pseudopotential implementation of DFT within both the generalized gradient and local density approximations. They found that increases in pressure are readily accommodated by tilting of the octahedra, explaining the small experimental bulk modulus. Also, they observed large W displacements, accompanied by strong rehybridization and changes in the electronic band gaps. This latter observation is consistent with the anomalously large Born effective charges calculated for cubic WO_3 by Detraux et al. [52].

This chapter builds on these earlier theoretical studies by treating *fractional* doping effects (i.e., the effect of electron concentrations between 0 and 1 per W ion) in a fully self-consistent manner. My computational approach allows the isolation of the effects of additional valence electrons from other factors, such as the presence of donor cations and structural disorder at intermediate concentrations. Indeed, it has already been shown that simple model calculations including only electronic effects can reproduce some aspects of the observed structural behavior [48]. The main result is that electronic effects are able to account fully for the experimentally observed structural phase transitions.

3.2 Methodology

Calculations were performed using the SIESTA implementation [53, 20], of density-functional theory [9] (DFT), within the Perdew-Zunger [54] parametrization of the local-density approximation [8]. Core electrons were replaced by norm-conserving Troullier–Martins [34] pseudopotentials, factorised as prescribed by Kleinman and Bylander. [55] The valence electrons were taken to be $2s^2 2p^4$ for O, and $6s^2 5d^4$ for W. Wave-functions were expanded in a basis of numerical atomic orbitals of finite support [56]. A double-ζ polarised (DZP) basis set was used, whereby two basis functions per valence atomic orbital are included, plus extra shells of higher angular momentum to allow polarization. This size of basis has been found, in prior studies,

Species	n	l	$r(\zeta_1)$	$r(\zeta_2)$
W	6	0	6.50	5.00
	5	2	6.49	3.75
	6	1*	6.49	n/a
O	2	0	3.31	2.51
	2	1	3.94	2.58
	3	2*	3.94	n/a

Table 3.1: Cutoff radii, $r(\zeta)$, for the basis sets corresponding to W and O species; as the basis set is double–ζ polarized, there are two basis functions (and therefore radii) per hypothetical atomic orbital, as specified by quantum numbers n and l. All radii are given in bohr. The starred entries are polarization orbitals.

Species	s	p	d	f
W ($5d^4 6s^2$)	2.85	3.03	2.25	2.25
O ($2s^2 2p^4$)	1.15	1.15	1.15	1.15

Table 3.2: Confinement radii for each of the s, p, d, and f channels of the W and O pseudopotentials. All radii are given in bohr. The W pseudopotential used also included a partial-core correction, with a radius of 1.30 bohr, according to the scheme of Louie et al [1].

to give results of similar quality to typical plane-wave basis sets in other major codes. Details for the pseudopotentials and the basis sets are given in Tables 3.1 and 3.2.

For k-point sampling, a cutoff of 10 Å was used [57]; this gave 24 independent k-points in the first Brillouin zone of the low-temperature monoclinic phase, corresponding to a 4×3×3 mesh. This is equivalent to 6 independent k-points within the first Brillouin zone of a 64-atom supercell of the same structure, corresponding to a 2×2×3 mesh (if one were to neglect degeneracy). Both k-point meshes were generated by the method of Monkhorst and Pack [58]. The fineness of the real-space grid used for numeric integration was set to correspond to an energy cutoff of 200 Ry. A grid-cell sampling [20] of four points (arranged on a face-centred cubic lattice) was used to reduce the space inhomogeneity introduced by the finite grid. All calculations were performed using variable-cell relaxation, with the convergence criteria set to correspond to a maximum residual stress of 0.01 GPa, and maximum

residual force component of 0.04 eV/Å.

In order to test the quality of the pseudopotential and basis set used, I first performed calculations on the same phases as investigated by de Wijs et al. Results of comparable quality to theirs were obtained (see table 3.3), although my results tend to give the expected underestimation of the experimental lattice parameters whereas theirs, even in the LDA, tend to overestimate.

In order to investigate the effect of additional dopant charge on the bonding character and structure of WO_3, a 2×2×1 supercell of the low-temperature monoclinic (ϵ) phase was used (the aforementioned 64-atom supercell). This contains sixteen W centres, and can be considered to be a $2\sqrt{2} \times 2\sqrt{2} \times 2$ supercell of the aristotypic perovskite structure. As such, it is sufficiently large to encompass all ground-state symmetries experimentally observed in pure WO_3. Additional charge was added by the injection of extra electrons into the system, with charge neutrality over space being ensured by a corresponding homogeneous positively-charged background; this can be performed using the corrections to the Ewald sum outlined in Chapter 2.

3.3 Calculated phase diagram

I performed calculations for numbers of dopant electrons between 1 and 16 per supercell, (corresponding to electron concentrations in the range $0 < x < 1$ per W centre) in single-electron steps. Full, variable-cell, structural relaxations were performed for each doping level. A series of five phases was observed over the doping range. Here, the five structures in turn are described, starting with the highest doping level. I refer to figure 3.1, which shows (a) my calculated lattice parameters (c' is the c lattice parameter multiplied by $\sqrt{2}$ in order to normalize it with respect to the a and b parameters), (b) my calculated unit cell volumes, and (c) our W-O bond lengths. Suggested phase boundaries are also shown in figure 3.1 (a) as solid vertical lines.

Table 3.3: Lattice parameters, experimental, from my simulations, and from the work of de Wijs [2] for (hypothetical) cubic, low temperature monoclinic, room temperature monoclinic, high-temperature tetragonal and room temperature triclinic phases of WO_3. All distances are given in Å. Note that de Wijs' calculations were performed under the local-density approximation for monoclinic phases, and under the generalized gradient approximation for triclinic and tetragonal phases.

Phase		a	b	c	α	β	γ
Cubic							
	This work	3.81	3.81	3.81	90.0	90.0	90.0
Monoclinic	LT						
	This work	5.15	5.05	7.63	90.0	92.6	90.0
	de Wijs	5.34	5.31	7.77	90.0	90.6	90.0
	Experiment [59]	5.28	5.15	7.66	90.0	91.8	90.0
Monoclinic	RT						
	This work	7.30	7.49	7.32	89.8	90.1	90.0
	de Wijs	7.37	7.46	7.64	90.0	90.6	90.0
	Experiment [60]	7.31	7.54	7.66	90.0	91.8	90.0
Tetragonal							
	This work	5.31	5.31	3.91	90.0	90.0	90.0
	de Wijs	5.36	5.36	3.98	90.0	90.0	90.0
	Experiment [61]	5.27	5.27	3.92	90.0	90.0	90.0
Triclinic	RT						
	This work	7.18	7.36	7.63	88.0	90.5	90.4
	de Wijs	7.54	7.64	7.84	89.7	90.2	90.2
	Experiment [62]	7.30	7.52	7.69	88.8	90.9	91.0

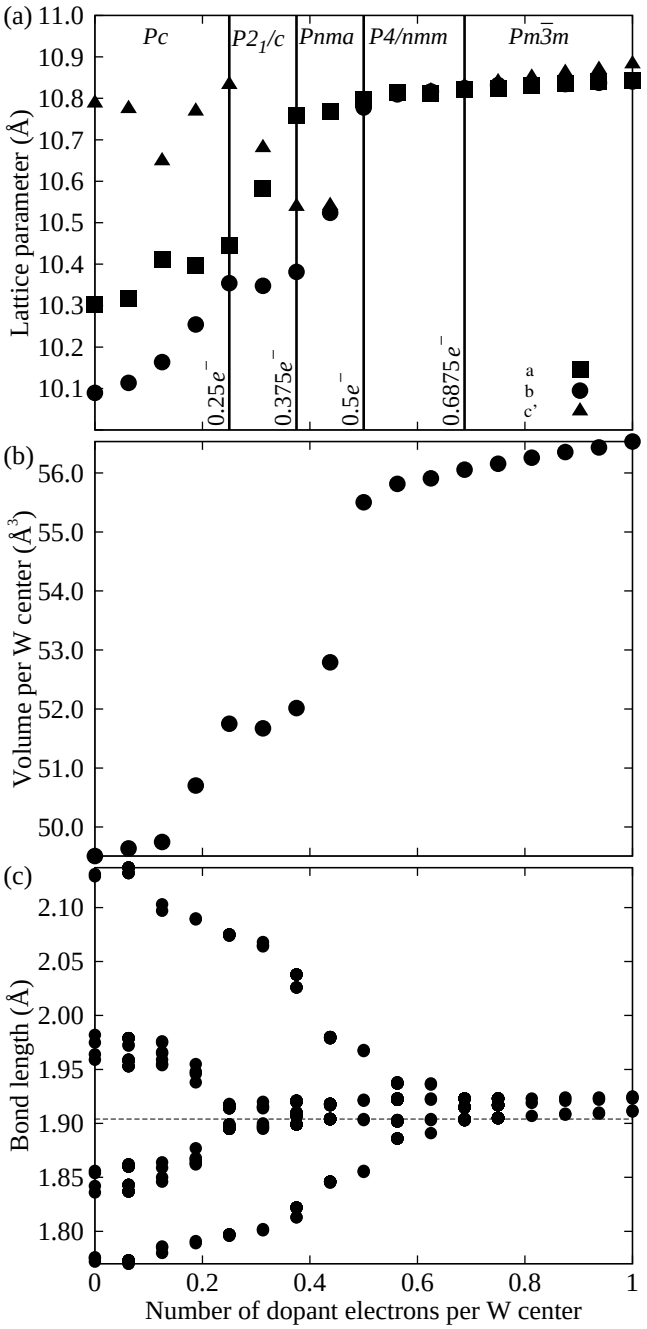


Figure 3.1: (a) Lattice parameters (for normalized cubic supercell with respect to the ϵ phase) of WO_3. Phase boundaries are denoted by solid lines, and labelled with their position (in electrons per centre); (b) Volumes of $2\times2\times2$ supercell of WO_3; (c) Bond lengths within WO_6 octahedra, all versus electron dopant concentration.

At maximal doping, the perovskite aristotypic phase is obtained (i.e. the highest symmetry phase in the series: in the case of perovskites, primitive cubic, $a = b = c$, one formula unit per unit cell) structure, with space group $Pm\bar{3}m$. The three lattice parameters are equal, as are the W-O bond lengths.

With decreasing dopant concentration the unit cell volume decreases slightly, until at $x \approx \frac{11}{16}$, a symmetry-breaking transition is found – to an antiferroelectric tetragonal phase with $P4/nmm$ symmetry. The W atom moves off the centre of its O_6 octahedron in the [001] direction [as shown in figure 3.2 (a)], resulting in different W-O bond lengths and a lowering of the space group of the system. Interestingly, all three lattice parameters continue to have the same effective length with respect to the supercell (within the accuracy of the computations) in spite of the inequivalence of one of them due to the reduction in symmetry. This can be ascribed to the fact that the oxygen framework distorts very little in these transitions, and it is the oxygen framework that is largely responsible for the volume and shape of the unit cell. The antiferroelectric tetragonal phase persists down to a dopant concentration of approximately $\frac{1}{2}$ electron per W, with a gradual increase in off-centring with decreasing dopant concentration. By $x = \frac{1}{2}$, the off-centring of the W along the [001] direction has increased to around 0.08 Å.

The antiferroelectric tetragonal phase becomes unstable under $\frac{1}{2}$ electrons per W, and an orthorhombic phase is stabilised. There is also a clear discontinuity in the volume per unit cell at this doping concentration. This arises from a tilt of the WO_6 octahedra towards [110]-type directions which in turn causes the loss of the four-fold rotation in the c direction, lowering the space group to $Pnma$. Initially, the c' and a lattice parameters (within the pseudocubic supercell), despite being symmetrically inequivalent, remain essentially equal in length; this can be ascribed to the absence of distortion of the WO_6 octahedra, which originally tilt (around an axis parallel to b) as a rigid unit without substantial deformation. These parameters become unequal at $x \approx \frac{7}{16}$; however, I do not find a lowering of the space group caused by this.

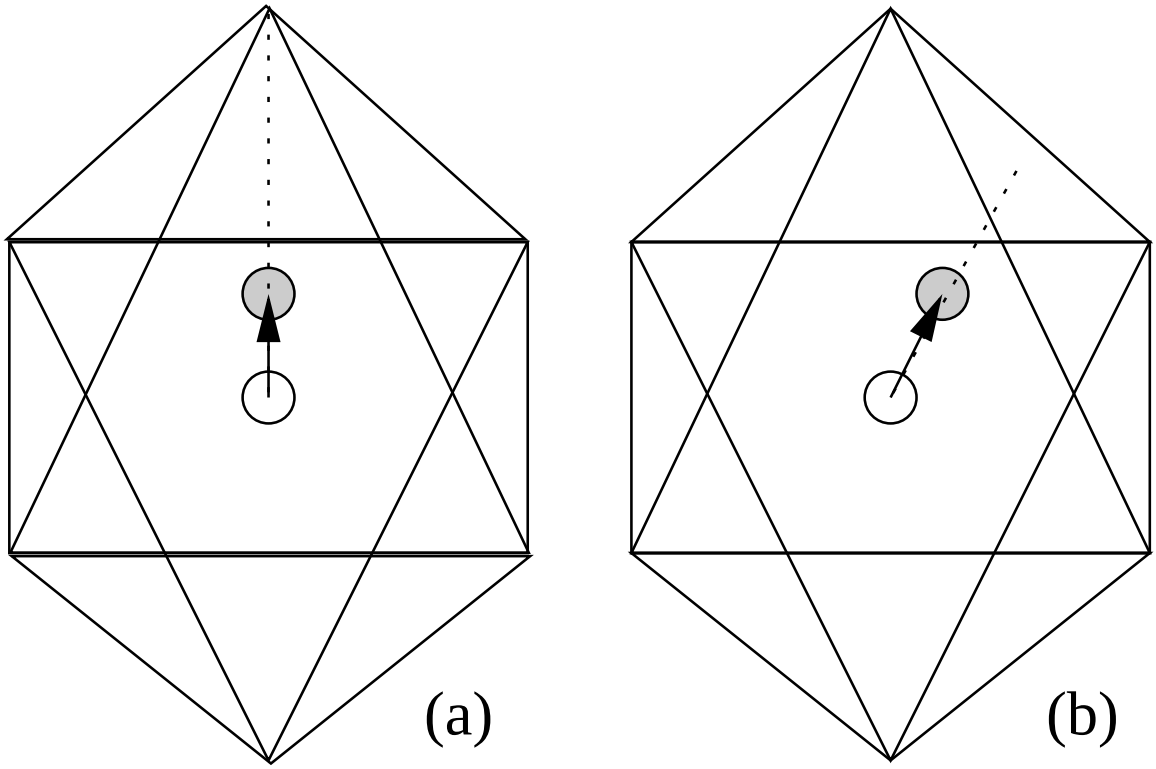


Figure 3.2: Offcentring in WO_6 octahedra; (a) along [001], (b) along both [001] and [110]-type directions

Below $x \approx \frac{3}{8}$, a monoclinic $P2_1/c$ phase (similar to that identified by de Wijs [2]), related to the low-temperature ϵ phase in WO_3, is stable. The phase boundary from the orthorhombic to monoclinic phase is characterized by an increase in the W-O displacement along [001] (seen in the increased bond splitting in figure 3.1 (c)) as well as a marked rotation around z (loosely speaking in the $x-y$ plane). The Glazer tilt is expressible as $a^-b^-c^-$. The monoclinic symmetry persists through the remainder of the phase diagram; however an additional phase boundary exists at $x \approx \frac{1}{4}$. Here an additional splitting in the W-O bond lengths indicates offcentring of the W atom in a [110]-type direction (as shown in figure 3.2 (b)) and there is also a second discontinuity in the cell volume. This corresponds to a loss of screw axes along **b**, and hence a further lowering of the space group to Pc, the ϵ phase.

Order parameters can be defined for the transitions between each of these phases as shown in figure 3.3. The order parameter for the Pc to $P2_1/c$ transition (figure 3.3 (a))is the difference between the W-O bondlengths in the $x-y$ plane (xy_1 and xy_2). The order parameter for the $P2_1/c$ to $Pnma$ transition is, as expected, the deviation of the β angle from 90 °; this goes sharply to zero at $x \approx \frac{3}{8}$ (figure 3.3 (b)). The order parameter for the $Pnma$ to $P4/nmm$ phase is the length difference between the two remaining unequal lattice constants, $|b-a|$ (figure 3.3 (c)). Finally, the ideal cubic $Pm\bar{3}m$ perovskite structure is reached when the W moves to its centrosymmetric position, and the order parameter for the tetragonal to the cubic phase is the magnitude of the W off-centre displacement along [001] (figure 3.3 (d)). As such, all the phase boundaries are well-defined, although further work could be undertaken to locate them with greater precision. In the next Section I discuss the effects giving rise to this phase structure in greater depth.

3.4 Discussion

There are two types of structural distortion occurring in WO_3; rotation of the (almost rigid) octahedra, and displacement of the W ion from the centre of its oc-

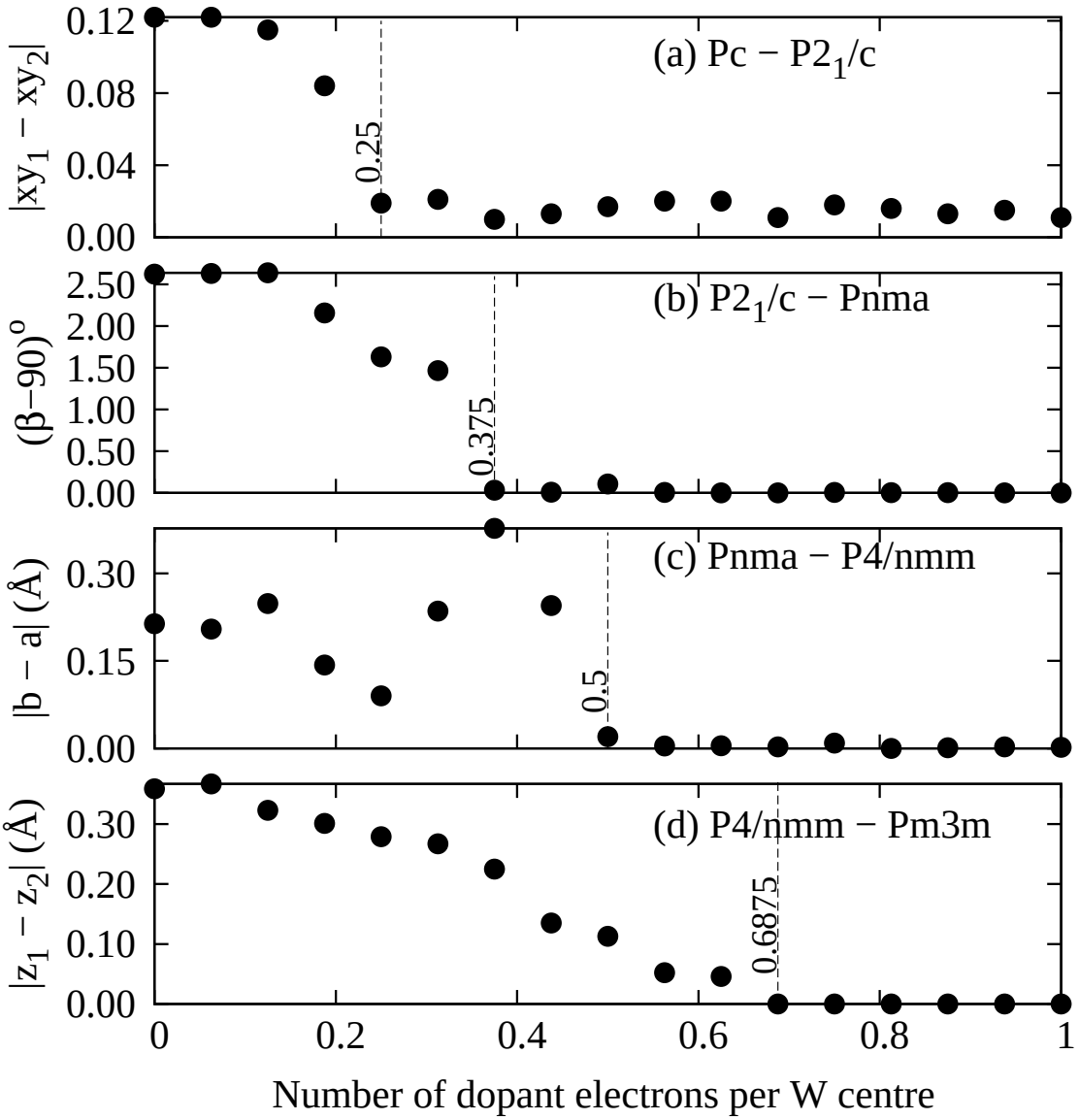


Figure 3.3: Order parameters for the doping-induced phase transitions in WO_3. (a) Pc – $P2_1/c$, in Å, where xy_1 and xy_2 are the inequivalent W-O bond lengths in the the $x-y$ plane; (b) $P2_1/c$ – $Pnma$, in degrees; (c) $Pnma$ – $P4/nmm$ (in Å); (e) $P4/nmm$ – $Pm\bar{3}m$ (in Å), where z_1 and z_2 refer to the two W–O bonds aligned predominantly with the z axis. Dashed vertical lines denote predicted phase boundaries, and are labelled with their positions in electrons per centre.

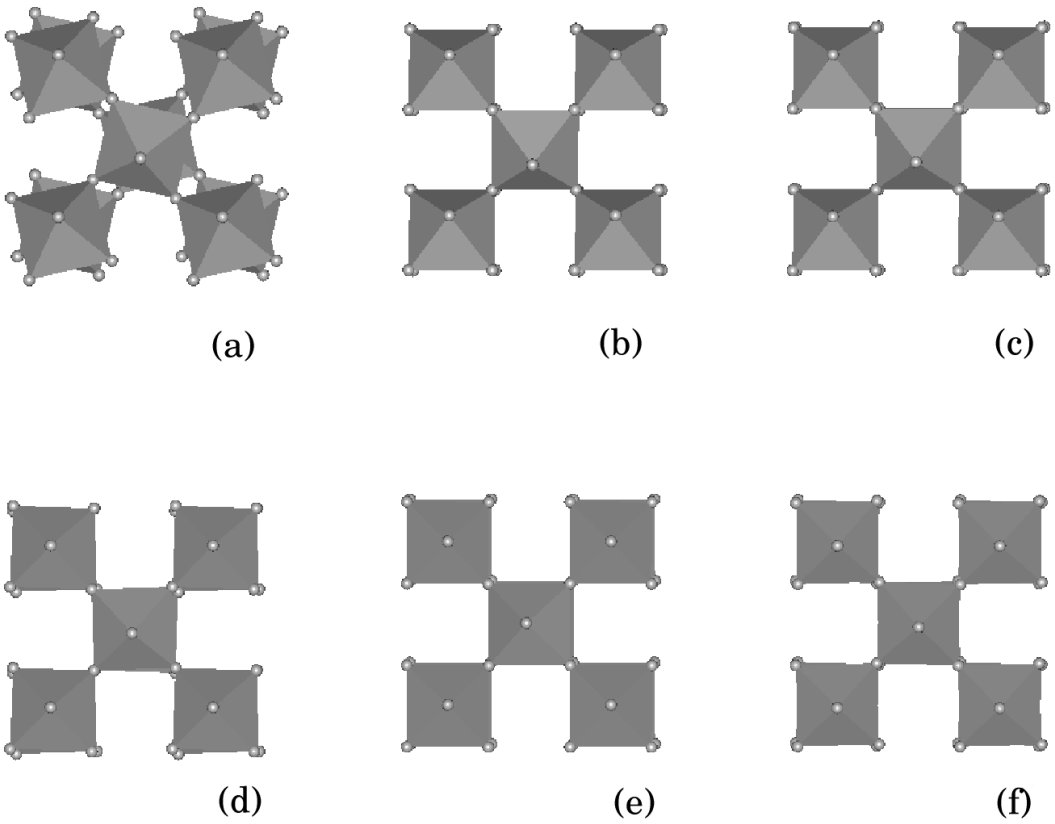


Figure 3.4: Polyhedral tilts in WO_3, as varying with dopant charge. The charges in each case are, in electrons per W atom, are; (a) 3/16, (b) 5/16, (c) 7/16, (d) 1/2, (e) 5/8, (f) 1; these correspond to, respectively, the monoclinic Pc phase; the $P2_1/c$ phase, approaching the $Pnma$ boundary; the middle of the orthorhombic phase; the $Pnma$ – $P4/nmm$ boundary; within the stability range of the tetragonal phase; and the aristotype. It is clear that the distortions between some phases are extremely small. This poses difficulties in ascertaining exactly where they are located in composition space.

tahedron. The occurrence or absence of W offcentring is known to be determined by a balance between electronic Coulomb repulsions (which are minimized for the centrosymmetric structure) and additional bonding considerations which might stabilize the non-centrosymmetric phase [50]. In the case of WO_3, off-centring of the W ion results in additional overlap that lowers the energy of the O $2p$-like valence band compared with that of the centred structure, and raises the energy of the predominantly W $5d$ t_{2g} conduction band [48]. As the doping level is increased, and more electron density is added to the conduction band, there is no energetic advantage to this additional orbital overlap, and the W ion moves back to its centrosymmetric position.

The fact that the small off-centring in the [110] direction is quenched first implies that the antibonding orbitals corresponding to this covalent interaction are low-lying in the conduction band and therefore filled first. The offcentring in the z direction persists to higher doping concentrations and is only quenched out at $\approx \frac{3}{4}$ e^- per centre. This 75% concentration is clearly less than the value of 0.98 e^- suggested by Corà et al.'s analysis of one-electron energies for the displacement of Re along [100] in ReO_3 [48]; this accords with their suggestion, however, that crystal–field effects would favour cubic structures, and thus any analysis under a rigid-band approximation (or similar) would overestimate the degree of doping necessary to cause the onset of cubicity.

There is strong evidence of a relationship between the polyhedral rotation and off-centring mechanisms. In figure 3.4 sketches are shown of the structures at different doping concentrations to illustrate the polyhedral tilting. At $\frac{3}{16}$ doping (figure 3.4 (a)) one can see that the rotations are large around both axes. However in figure 3.4 (b) (by which the offcentring in the [110] direction has been essentially quenched ($x \approx \frac{5}{16}$)) one can see that the rotation of polyhedra in that plane also disappears, whilst rotation of polyhedra between connections along the z axis persists, as does the offcentring in that direction. This persists, with reducing magnitude, to $x \approx \frac{7}{16}$ (figure 3.4 (c)), but by 50% doping, (figure 3.4 (d)) any remaining tilt up the z axis (out of the page) is extremely small. However, at this concentration, offcentring

along [001] persists. Notice also that the majority of volume change in the unit cell, which in perovskites is caused by changing of the rotation angles, has also occurred by this doping level. Finally, any tilt systems in figures 3.4 (e) and (f), corresponding to concentrations with centred W ions, are too small to observe in the figure.

It should be noted that the very small [100]-offcentring - less than 0.01 Å in magnitude - which persists at high doping concentration does not appear to be a real effect. In fact, both this and the small Jahn-Teller distortion at 100% doping appear to be numerical artifacts related to the real-space mesh cutoff used: when the mesh is increased to 500 Ry (from 200 Ry), both effects disappear. As it is computationally expensive, one cannot justify this level of convergence for all calculations, particularly given that the order parameters delineating the phase boundaries in the system are well-defined at the present level of accuracy. Even so, it is clear that the system is very near the limits where Jahn-Teller distortion and/or offcentring of the W atom become thermodynamically stable. As expected by analogy with ReO_3, WO_3 doped with 1 electron is entirely cubic, with no Jahn-Teller distortion or W off-centring. Since the transition metal, in both cases, has a formal d^1 electron configuration (and d^1 ions in octahedrally coordinated complexes do not have a second-order Jahn-Teller distortion) the centrosymmetricity is easy to explain. The lack of a true Jahn-Teller distortion (which is predicted for such a d^1 ion in octahedral coordination) is more subtle however, and is determined by a competition between energy lowering through hybridization (strongest in the cubic structure) and energy lowering through gap formation during the distortion. In bulk ReO_3, the broad bands and low density of states at the Fermi energy result in the cubic phase having the lowest energy. In contrast, in a molecular complex the bands are infinitely narrow and the Fermi energy density of states is infinitely large (favouring Jahn-Teller distortion).

In figure 3.5 the calculated change in energy is plotted, as a function of the amount of *both* Jahn-Teller distortion and transition metal off-centring, for WO_3 doped with one electron per W centre. In both cases a constant volume is maintained. The x axis shows the fractional change in bond length; for the Jahn-Teller distorted

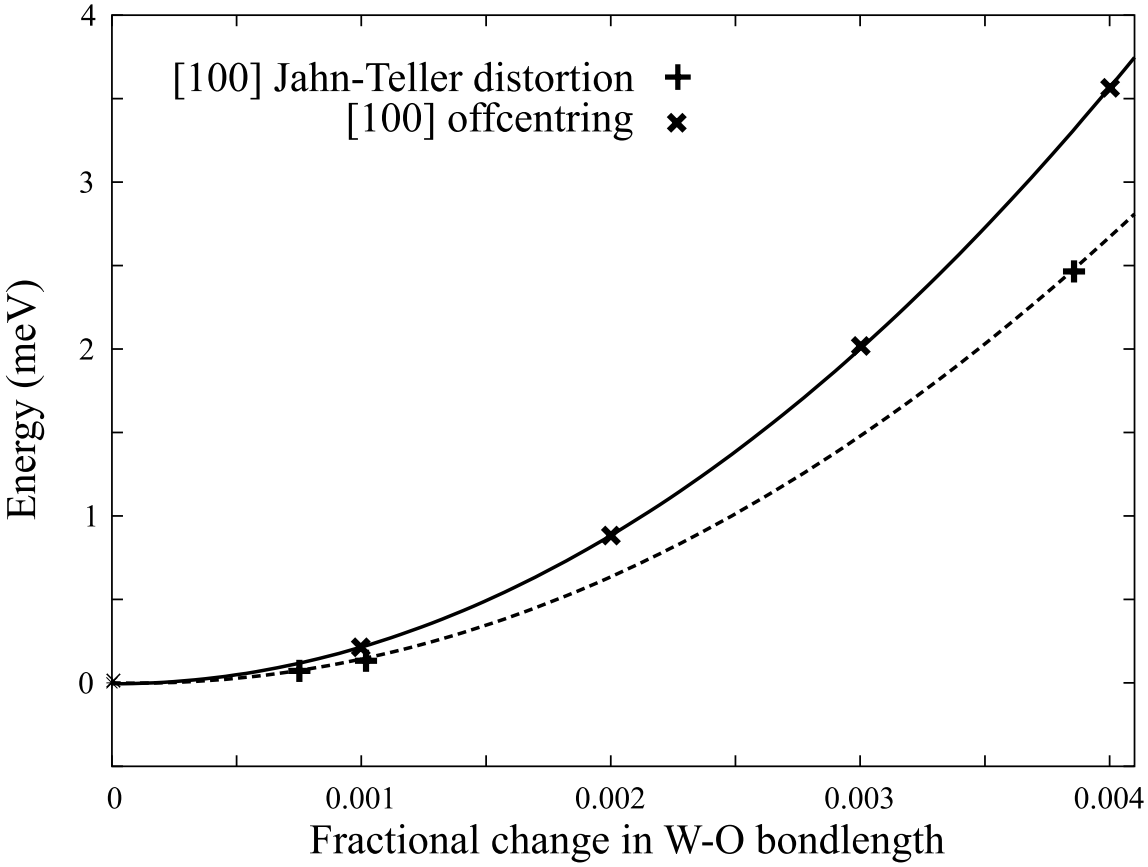


Figure 3.5: Increase in energy per formula unit as a function of fractional change in W-O bondlength, for both W off-centring and Jahn-Teller distortion along a cubic axis.

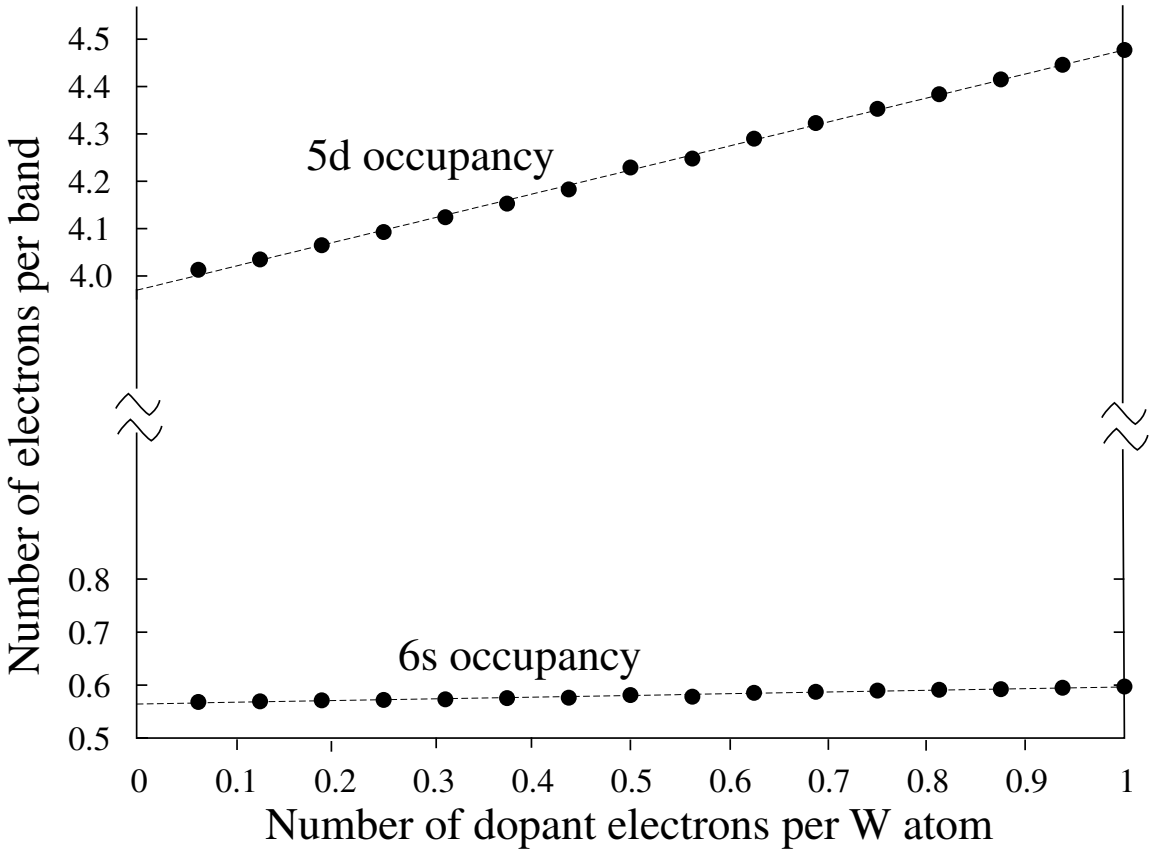


Figure 3.6: Mean occupancy per W atom of $5d$ and $6s$ orbitals versus dopant charge.

structure both W-O bondlengths increase by this amount, and for the off-centred structure one increases and the other decreases. Note that, by this measure, the structure is stiffer to Jahn-Teller distortion than it is to off-centring. Also note that the curvature for off-centring is very close to that of Cora et al. [48] for ReO_3 (which in turn is around half that for $NaWO_3$).

A Mulliken analysis of the occupancy of the three families of bands due to W – $6s$, $5d_{e_g}$, and $5d_{t_{2g}}$ – suggests that the occupancy of the $5d$ band is far from d^0 in undoped WO_3, and far from d^1 in the fully doped system. The occupancy of the d-band is closer to four in undoped ϵ–WO_3 (see figure 3.6) and increases by approximately 0.5 e^- in the fully doped system. The remaining half of an electron, unsurprisingly,

fills the antibonding bands formed by O($2p$)–W($5d$) overlap. Note that the mean occupancy of the $6s$ orbital, which is ≈ 0.58 electrons per orbital in the undoped system, hardly increases on addition of electrons. This implies a typical W valence population of between 4.5 and 5 electrons, in contrast to the idealised purely-ionic picture of a bare W^{6+} ion (with no valence electrons); this is yet further evidence of a system remarkably dominated by covalent W–O interactions. It is clear from further analysis that above 50% doping, dopant charge preferentially fills the $5d_{xz}$ and $5d_{yz}$ bands; the other bands experience only a small change in total occupancy. This accords well with the suggestion that bonding–antibonding splitting occurs through O($2p$)-W($5d$) π-like overlap; the absence of change in xy bonding being demonstrated by the absence in change of occupancy of the $5d_{xy}$ band. However, given the tilting transitions and highly-deformed structures, there is great difficulty in assigning the occupancy of $5d$ orbitals; therefore, further studies of the influence of band formation and orbital occupation on both first- and second-order Jahn-Teller distortions are ongoing.

In terms of reconciling my calculations with previous experimental data [47], one must be somewhat cautious. Clarke analyzed his results under the presumption that there would not be significant offcentring of the W ion within WO_6 octahedra, which I believe to be incorrect: nevertheless, I believe that the calculated tetragonal phase is consistent with collected data for his proposed cubic phase at high doping. Furthermore, my calculations deliberately neglect the effects of defects, chemical disorder, and other imperfections in the lattice.

Finally, it should be mentioned that the original motivation of our study was to investigate possible self-trapping behaviour of charge in this material. Excess electrons in the ϵ (low-temperature monoclinic) phase have been shown experimentally to self-trap, forming polarons [59]. No localisation of charge or deformation (polaron formation) has been found in any of my calculations. This null result is of some interest, but should not be taken to mean that polarons do not form in this material; it is arguable that either the supercell used is too small to observe polaron localisation, or that the LDA will underestimate the binding energy of a polaron

(thus causing it not to be a stable state of the system): further study is needed in this area.

3.5 Conclusions

The main conclusion that can be drawn from these calculations is that the experimentally observed structural distortions induced in Na_xWO_3–type bronzes by increasing doping are predominantly electronic in nature. By using my methodology of adding electrons to WO_3 without also adding Na atoms, I have removed any possible structural/disorder effect caused by the Na^+ cations on the A sites within the structure. Therefore, given that my calculations reproduce the experimental observed sequence of symmetries upon doping ((i) Monoclinic (Pc and $P2_1/c$); (ii) Orthorhombic ($Pnma$); (iii) Tetragonal ($P4/nmm$); (iv) cubic aristotype ($Pm\bar{3}m$)) I conclude that the effect of the Na^+ cations on the structure must be small. Indeed, this is not unexpected. Following the classification of Robin and Day [63], as mentioned by Bersuker [64], $NaWO_3$ is a good example of pure electronic doping; there is complete transfer of the donated Na electron over to the WO_3 sublattice since the $3s$ band of Na lies well above the $5d$-antibonding band from W.

Finally, I propose that these materials present interesting opportunities for future experimental and theoretical study, given the degree of structural control that can be gained from doping or substitution. In particular, the "quenching out" of polyhedral rotation (and hence the opening of channels within the structure) may have effects on diffusion and ion intercalation in these structures.

Chapter 4

Constrained Linear Maximization

4.1 Introduction

Within harmonic transition state theory (hTST) [65], there are two sets of key points on the potential energy surface (PES) of any given system; the energy minima, corresponding to the metastable states of the system, and the first-order saddle points connecting them. Given these, it is possible to reconstruct every reaction pathway in the system, and thus determine thermodynamic and kinetic behaviour; in many ways, this complete prediction of reactive properties is one of the key challenges for simulations in materials research.

It is therefore not surprising that many techniques have been proposed in order to identify these key points. The techniques of energy minimization are very well established; identification of saddle-points is substantially harder, and many innovative techniques have been proposed. These can, essentially, be divided into three varieties: biased dynamics (including Voter's hyperdynamics [66] and Micheletti and Laio's history-dependent metadynamics [67] methods), which are primarily methods for performing long timescale simulations, but in so doing identify transition states; two-basin direct searches, from the historical Intrinsic Reaction Coordi-

nate [68] method to more modern Nudged Elastic Band [69] and Linear/Quadratic Synchronous Transit approaches, which aim to locate the transition between two known metastable states; and single-basin searches, which aim to find saddle points based on no knowledge of any other stable states. Therefore, in terms of specific motivation, single-basin direct searches and metadynamics methods have most in common: they point towards a generalized mapping of the salient features of the PES, rather than discovery of the most energetically favourable mechanism for one specific transition on the surface.

I propose an extremely simple single-basin algorithm, derived from the old idea of the Intrinsic Reaction Coordinate [68], for the location of first-order saddle points on a PES, given only the ability to calculate the energy, and its first partial derivatives with respect to position, of a given configuration of the system. By my method, the Intrinsic Reaction Coordinate method is extended and generalised to the single-basin case, and by so doing, a controllable, conceptually simple, and highly parallelisable method for energy surface exploration is developed.

4.2 Constrained Linear Maximization

The algorithm is named *constrained linear maximization* (CLM).

Represent the configuration space of the system (containing N atoms) by a $3N$-dimensional state vector, $\mathbf{x}$, composed of a list of the coordinates of each of the atoms in the system (in some appropriate coordinate basis; for the present work, the Cartesian basis is used). Assume that the total energy of the system is representable as an (exact and deterministic) function, $E(\mathbf{x})$.

In a chemical reaction, the path of least work (the path with the lowest maximum in energy) is taken between the two basins. This path must pass through a first-order saddle point on the PES, and therefore is at a maximum along one of the Hessian's

eigenvectors and a minimum along all the others. By identifying this point, one can recover the entire reaction pathway.

Often, one cannot achieve this practically by direct calculation of the Hessian. For extended systems, the Hessian is much too large to evaluate repeatedly; even a single evaluation of the Hessian directly takes $6N$ force evaluations for a given system.

However, some number of atoms *must* participate in a given class of activated processes (for instance, the diffusing atom in a diffusion process; other atoms may also participate, but one can be certain of the involvement of some atoms), one selects an "active subspace" within the system (consisting, for instance, of the degrees of freedom of any atom within a fixed radius of the diffusing species).

This immediately suggqests an algorithm based on the above argument; choose a random normalized vector within the active subspace, $\mathbf{v}$, and choose the distance along that direction, γ, which maximizes energy subject to the constraint that energy will be simultaneously minimized along all the orthogonal degrees of freedom. This can be achieved very simply by use of a simple hill-climbing search for a modified potential.

To express this algebraically, define a space composed of the complement of the complete configuration space and V; this can be represented by a basis of size 3N-1, $\mathbf{q}_0 ... \mathbf{q}_{3n-1}$. Within this, one can easily define the modified potential:

$$E(\gamma) = \min_{q_1 ... q_n} E\left(vx_0 + \gamma \mathbf{v} + \sum_{n=0}^{3n-1} (k_n \mathbf{q}_n) \right) \tag{4.1}$$

Therefore, for a transition state, one can argue that:

$$E_{TS} = \max_{\gamma} \left[\min_{q_1 ... q_n} E\left(vx_0 + \gamma \mathbf{v} + \sum_{n=0}^{3n-1} (k_n \mathbf{q}_n) \right) \right] \tag{4.2}$$

The best way to visualize this proposal is by considering its behaviour near a saddle point on an idealized quadratic potential. In this spirit, consider an idealized potential of the following form:

$$E\left(x_1 \ldots x_n, y\right) = \sum_{1}^{n} \left[a_n \left(x_n - x_{n_0}\right)^2\right] - b(y - y_0)^2 \tag{4.3}$$

which has a transition state at $\mathbf{r_0} = (x_{1_0}, x_{2_0}, x_{3_0} \ldots y_0)$. Choosing an arbitrary direction $\mathbf{v}$ as above, any position $\mathbf{r}$ on the surface can be expressed as

$$\mathbf{r} = \gamma\mathbf{v} + \sum_{1}^{n} \left(\mathbf{q_n}\right) \lambda_n \tag{4.4}$$

where $\{\mathbf{q_{1\ldots n}}\}$ comprises the set of unit vectors normal to $\mathbf{v}$. One can recast equation 4.1 using Lagrange multipliers to embody the minimization constraint; noting that

$$\mathbf{r}.\mathbf{v} = \gamma \tag{4.5}$$

where γ is the free parameter for maximization, it follows that

$$\min_{\mathbf{r},\Delta} F_\gamma\left(\mathbf{r}, \Delta\right) = E\left(\mathbf{r} - \Delta\left(\mathbf{r}.\mathbf{v} - \gamma\right)\right) \tag{4.6}$$

Δ being the Lagrange multiplier, with partial derivatives

$$\left.\frac{\partial E}{\partial x_i}\right|_{min} = \Delta v_{x_i} \tag{4.7}$$

which allows one to consider the potential under evaluation. Taking the partial derivatives of Equation 4.3 with respect to x_i and y, and equating to 4.7:

$$\frac{\partial E}{\partial x_n} = 2a_n \left(x_n - x_{n_0}\right) = \Delta v_{x_n} \tag{4.8}$$

$$\frac{\partial E}{\partial y} = -2b \left(y_n - y_0\right) = \Delta v_y \tag{4.9}$$

It follows from this that

$$\Delta v_{x_n} = 2a_n x_n - 2a_n x_{n_0} \tag{4.10}$$

$$\Rightarrow \frac{\Delta v_{x_n}}{2a_n} = x_n - x_{n_0} \tag{4.11}$$

$$\Rightarrow x_n = \frac{\Delta v_{x_n}}{2a_n} + x_{n_0} \tag{4.12}$$

$$\Rightarrow x_n v_{x_n} = x_{n_0} v_{x_n} + \frac{\Delta v_{x_n}^2}{2a_n} \tag{4.13}$$

and, similarly,

$$y v_y = v_y y_0 - \frac{\Delta {v_y}^2}{2b} \tag{4.14}$$

Bearing mind the constraint in equation 4.5, one observes that:

$$\gamma = \mathbf{r}.\mathbf{v} = \sum_{1}^{n} \left[x_{n_0} v_{x_n} + \frac{\Delta {v_{x_n}}^2}{2a_n} \right] + v_y y_0 - \frac{\Delta {v_y}^2}{2b} \tag{4.15}$$

$$= \mathbf{r_0}.\mathbf{v} + \sum_{1}^{n} \left[\frac{\Delta v_{x_n}^2}{2a_n} \right] - \frac{\Delta {v_y}^2}{2b} \tag{4.16}$$

$$= \mathbf{r_0}.\mathbf{v} + \frac{\Delta}{2} \left[\sum_{1}^{n} \left(\frac{v_{x_n}^2}{a_n} \right) - \frac{{v_y}^2}{b} \right] \tag{4.17}$$

$$\Rightarrow 2\left(\gamma - \mathbf{r_0}.\mathbf{v}\right) = \Delta \left[\sum_{1}^{n} \left(\frac{v_{x_n}^2}{a_n} \right) - \frac{{v_y}^2}{b} \right] \tag{4.18}$$

Making the substitutions derived above for x_n and y into the potential, given in equation 4.3):

$$E(\gamma) = \sum_{1}^{n} \left(x_n - x_{n_0}\right)^2 - b\left(y - y_0\right)^2 \tag{4.19}$$

$$= \sum_{1}^{n} a_n \left[\frac{\Delta v_{x_n}}{2a_n} + x_{n_0} - x_{n_0} \right]^2 - b \left[y_0 - \frac{\Delta v_y}{2b} - y_0 \right]^2 \tag{4.20}$$

$$= \sum_{1}^{n} \left[\frac{\Delta^2 {v_{x_n}}^2}{4a_n} \right] - \frac{\Delta^2 {v_y}^2}{4b} \tag{4.21}$$

$$= \frac{\Delta^2}{4} \left[\sum_{1}^{n} \left(\frac{v_{x_n}^2}{a_n} \right) - \frac{{v_y}^2}{b} \right] \tag{4.22}$$

Therefore, substituting for the expression in square brackets:

$$2(\gamma - \mathbf{r_0}.\mathbf{v}) = \Delta \left[\sum_1^n \left(\frac{v_{x_n}^2}{a_n} \right) - \frac{{v_y}^2}{b} \right] \tag{4.23}$$

$$\Rightarrow \frac{2}{\Delta}(\gamma - \mathbf{r_0}.\mathbf{v}) = \left[\sum_1^n \left(\frac{{v_{x_n}}^2}{a_n} \right) - \frac{{v_y}^2}{b} \right] \tag{4.24}$$

$$\Rightarrow E(\gamma) = \frac{\Delta^2}{4} \times \frac{2}{\Delta}(\gamma - \mathbf{r_0}.\mathbf{v}) \tag{4.25}$$

$$= \frac{\Delta}{2}(\gamma - \mathbf{r_0}.\mathbf{v}) \tag{4.26}$$

At a stationary point, $\partial E/\partial\gamma = 0$ and $\partial^2 E/\partial\gamma^2 < 0$. It follows that:

$$2(\gamma - \mathbf{r_0}.\mathbf{v}) = 0$$
$$\Rightarrow \gamma = \mathbf{r_0}.\mathbf{v}$$

but as $\gamma = \mathbf{r_{max}}.\mathbf{v}$ from equation 4.5, it follows that:

$$\mathbf{r_{max}}.\mathbf{v} - \mathbf{r_0}.\mathbf{v} = 0$$
$$\Rightarrow (\mathbf{r_{max}} - \mathbf{r_0}) \perp \mathbf{v}$$
$$\Rightarrow \mathbf{r_{max}} = \mathbf{r_0}$$

here. Thus, we have located the stationary point on the energy surface at $\mathbf{r_0}$. Of course, one is then not certain of the character of the stationary point is, so we go on to this problem next.

Consider a potential of the form

$$E(\mathbf{x}) = \sum_{1}^{n} a_n {x_n}^2 \tag{4.27}$$

where

$$\mathbf{x} = \mathbf{x_0} + \gamma \mathbf{v} + \sum_{1}^{n-1} \lambda_p \mathbf{q_p} \tag{4.28}$$

analogously to above, which is equivalent to the above problem up to translation of basis as long as $a_1 < 0$ and $a_{2\ldots n} > 0$.

Choose an (arbitrary) component of $\mathbf{x}$, x_n; it follows that

$$x_n = x_{n_0} + \gamma v_n + \sum_{1}^{p} \lambda_p q_{p_n} \tag{4.29}$$

Substituting this into our potential, one finds that:

$$E = \sum_{1}^{n} a_n \left(x_{n_0} + \gamma v_n + \sum_{1}^{n-1} \lambda_p q_{p_n} \right)^2 \tag{4.30}$$

Therefore:

$$\frac{\partial E}{\partial \gamma} = \sum_{1}^{n} \left(2 v_n a_n \left(x_{n_0} + \gamma v_n + \sum_{1}^{n-1} \lambda_n q_{p_n} \right) \right) \tag{4.31}$$

$$\Rightarrow \frac{\partial E}{\partial \gamma} = \sum_{1}^{n} \left(2\gamma a_n {v_n}^2 + 2 a_n v_n x_{n_0} + 2 v_n \sum_{1}^{n-1} \lambda_n q_{p_n} \right) \tag{4.32}$$

$$\Rightarrow \frac{\partial^2 E}{\partial^2 \gamma} = \sum_{1}^{n} 2 a_n {v_n}^2 \tag{4.33}$$

Therefore, if there is to be a maximum in energy with variation of γ, then $\frac{\partial^2 E}{\partial^2 \gamma} < 0$; therefore

$$-a_1 {v_1}^2 > \sum_{2}^{n} a_n {v_n}^2 \tag{4.34}$$

This is a central result – it allows one to calculate, for this potential, demonstrating the range of $\mathbf{v}$ for which IRC-style simple-minded maximization along a random vector will be effective. In other words, this allows us to quantify the angle of approach to the transition state for which the maximization path through the transition state will be continuous. One simple case is where $-a_1 = a_2 = \ldots = a_n$;

$$ {v_1}^2 > \sum_{2}^{n} {v_n}^2$$

but noting that (by definition),

$$\begin{aligned} v_1^2 + \sum_2^n {v_n}^2 &= 1 \\ \Rightarrow \sum_2^n {v_n}^2 &= 1 - v_1^2 \\ \Rightarrow {v_1}^2 &> 1 - v_1^2 \end{aligned}$$

It follows that:

$$\Rightarrow {v_1}^2 > 1/2 \tag{4.35}$$

Of the two roots here, one points towards the transition state, and one away from it; therefore one need consider only the positive root:

$$v_1 > \frac{1}{\sqrt{2}} \tag{4.36}$$

This lets one simulate the scaling of the success-rate of CLM with increasing dimensional degrees of freedom, simply by generating normalized trial vectors and testing them against this condition, as can be seen in Figure 4.2.

This is a more general result than it seems, given that all transition states tend (within harmonic transition state theory) to being quadratic the closer you get to the transition state. Therefore, this result shows that random selection of trial vectors is not going to be good enough – Figure 4.2 shows that the scaling of probability of sucessess with increasing dimensionality is such that the probability of success rapidly tends towards zero.

In fact, so far, the method is exactly identical (mathematically) to the reaction

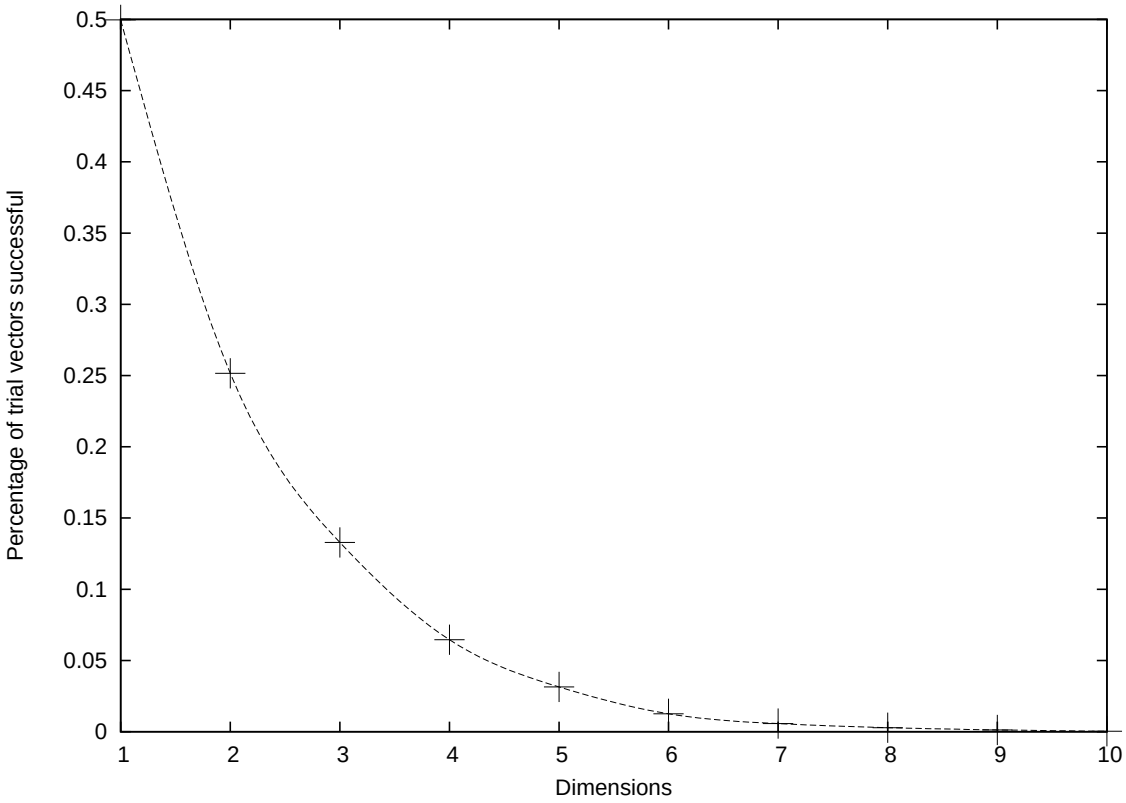


Figure 4.1: Converged CLM runs on a 2D harmonic potential.

coordinate (IRC) methods I have referred to above; where we differ is, firstly, in selection of initial search direction. In IRC, the search direction used is a linear interpolation between reactant and product basins. As I have demonstrated, this is insufficient in many common cases; if the true reaction coordinate lies some distance off this interpolated line (that is to say, the reaction coordinate is curved) then the trajectory taken through the system becomes discontinuous as noted above.

Additionally, one does not necessarily, *a priori*, know the locations of two minima on our PES; therefore in these cases some other process for selection of an initial trial vector must be used. In most systems, one is most interested in the behaviour of one particular atom or set of atoms (for instance, if one is investigating diffusion, the position of the diffusing species is necessarily going to be involved in the reaction coordinate); using this kind of knowledge of the system one can bias one's initial selection to fall within a small subspace.

However, it is still likely that the trajectory taken through the system will be discontinuous (if it is not, we have successfully discovered a transition state). Therefore, one needs to consider what happens in the case where my algorithm fails.

One can compare two paths (for increasing γ) on a simple 2D potential of the form $E(x, y) = x^2 + (y^2 - 1)^2$. This has a transition state at $(0, 0)$ and minima at $(0, -1)$ and $(0, 1)$; by starting the runs in the minimum at $(0, 1)$, one produces trajectories which can be seen in figure 4.2.

One of these runs is successful, passing through the transition state; for the other, there is a discontinuity over the location where the transition state lies. This is visible in plots of the system's energy as a variable of γ for these two runs (figure 4.3); in the successful case, $E(\gamma)$ is maximised at the transition state, whereas in the unsuccessful one there is a cusp.

However, we now have information which can enable us to orient our search vector towards a direction which will converge on the desired transition state.In the discontinuous case, the true transition state typically lies between the two end-points of

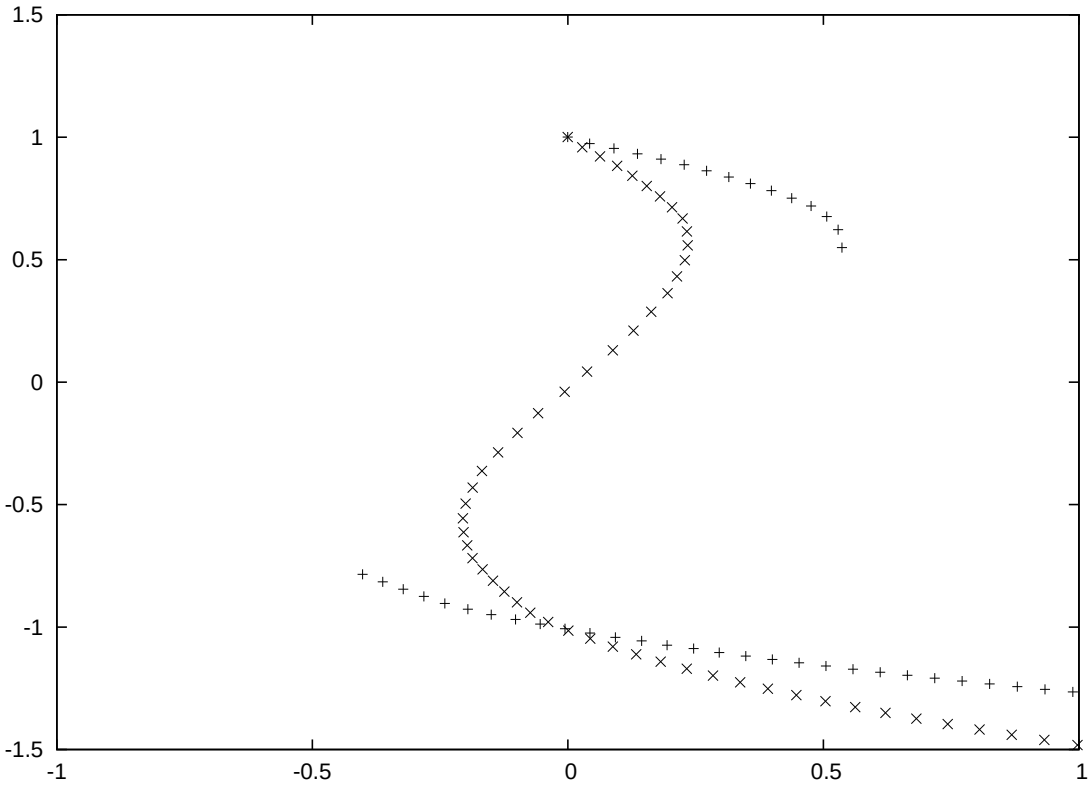


Figure 4.2: System trajectories for a quadratic/quartic 2D potential.

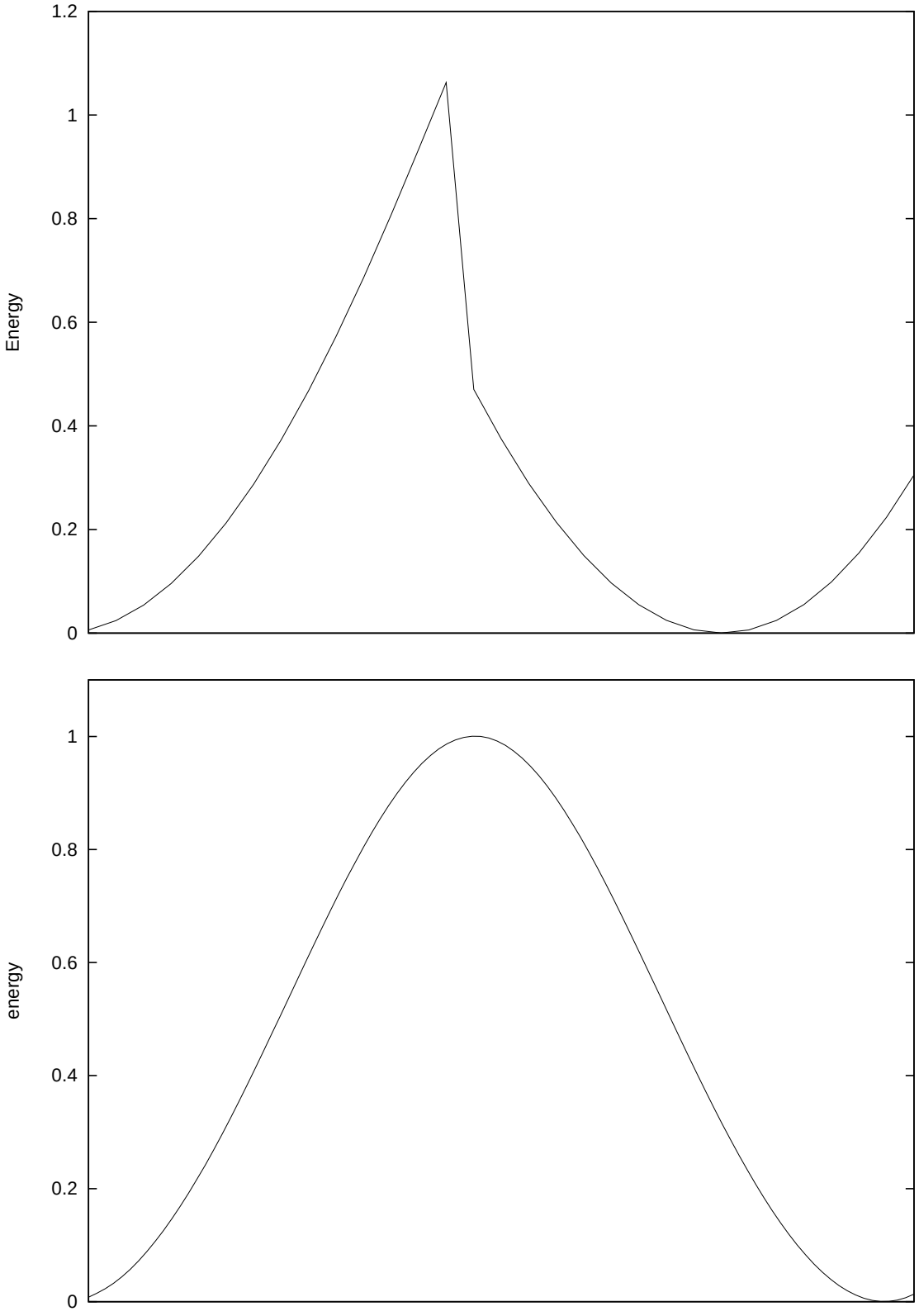


Figure 4.3: System energy profiles for a quadratic/quartic 2D potential, in one successful (the lower plot) and one unsuccessful (the upper) case.

the discontinuity (in the coordinate space of the system). This suggests a strategy.

Let the normalized vector between the discontinuity endpoints be $\mathbf{k}$; therefore, take our new trial vector to be:

$$\mathbf{v}' = \frac{\mathbf{v} + \delta\mathbf{k}}{((\mathbf{v} + \delta\mathbf{k}).(\mathbf{v} + \delta\mathbf{k}))^{1/2}} \tag{4.37}$$

This will rotate our search direction towards the transition state in the vicinity of the transition. As was demonstrated by the critical-angle analysis above, this is the central idea of this new method; the necessary convergence condition for my method is that the search direction is oriented closely enough to the minimum-curvature eigenvector at the TS.

As might be expected, the efficiency of this method depends strongly on the choice of δ; this shall be examined in the next section, but it should be noted that (assuming δ is not excessively large) the method is self-correcting; as $\mathbf{v}'$ approaches the correct search direction, $\mathbf{v}'.\mathbf{k}$ (for normalised $\mathbf{k}$) approaches unity, and therefore the magnitude of search vector rotation will decrease (leading to convergence).

The results of this can be seen in figure 4.4, and show that this is a viable strategy; CLM with search update by interpolation is shown to be both effective and to scale approximately linearly with degrees of freedom.

4.3 Constrained Linear Maximization: Further Details

This is no use, however, unless it is easy to implement as a practical search algorithm. This is, thankfully, the case. The algorithm decouples neatly into two stages,

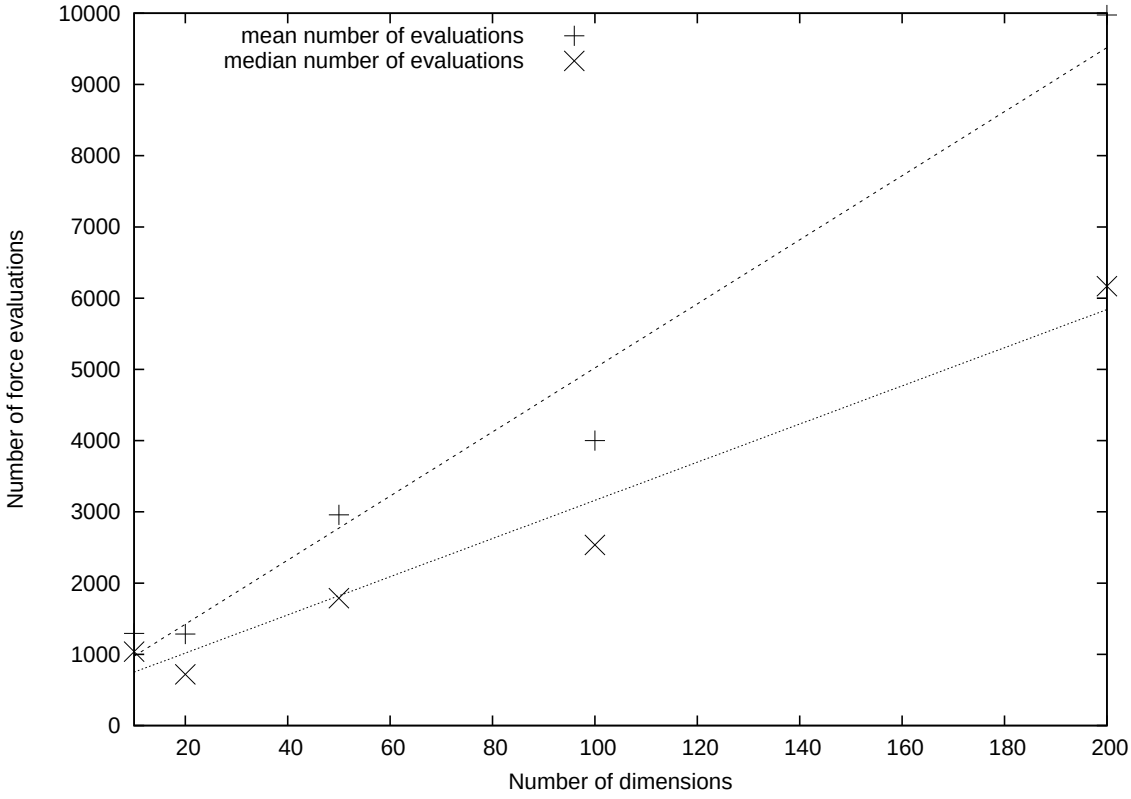


Figure 4.4: Dimensional scaling of the interpolation strategy for a potential of the form $E(x_1, x_2 \ldots x_n) = (1 - x_1^2)^2 + x_2^2 + \ldots + x_n^2$

consisting of an outer maximization algorithm around an inner minimization (or, to look at it another way, maximization of a modified functional which contains a minimization stage).

First, let us consider the minimization stage. We must prevent energy minimization along one specific direction (i.e., in the space of $\mathbf{q_1} \ldots \mathbf{q_{3n-1}} \perp \mathbf{v}$), which can easily be achieved by firstly applying a projection operator to the system's forces;

$$\mathbf{F}'(\mathbf{x}) = \mathbf{F}(\mathbf{x}) - (\mathbf{v}.\mathbf{F}(\mathbf{x}))\mathbf{v} \tag{4.38}$$

and, then, performing conjugate-gradient minimization of the energy using these modified forces.

Using this, one can then maximize the value of equation 4.1 – the fundamental equation of the algorithm, or in other words maximization of the energy of the system along a given vector in space, here $\mathbf{x}$, whilst minimizing in its tangent space – with respect to γ. One approach would be to use bisection search; this is effective, but is less efficient than an even simpler approach - simply taking the end-point of a previous tangent minimization and iteratively displacing the system by some small distance along the maximization direction. In this latter case, the system starts much closer to the constrained trajectory, and therefore far fewer steps are needed in the tangent space to minimize energy.

A further optimization is that when a discontinuity is detected, one can obtain a good initial guess for the next maximisation iteration by first minimizing the discontinuous endpoint closer to the origin in the subspace defined by the new maximization direction; this, again, substantially improves the efficiency of the algorithm.

At a transition state, $F(\mathbf{x})$ tends to zero, providing an obvious check on the quality of convergence achieved.

Conceptually, this algorithm has some connection to the ART Nouveau [70] algorithm of Mousseau and collaborators. ART Nouveau depends for efficiency on an

"activation" step, where one displaces the system along some vector in configuration space until a sufficiently-negative eigenvalue is observed (which can be shown by direct calculation of the Hessian and use of (for instance) the Lanczos method, as in their original paper, or by a variational method for partial Hessian evaluation such as that of Munro and Wales [71]; from there, eigenvector-following will lead to a transition state.

It is noteworthy that at no point in the CLM algorithm is the Hessian directly or indirectly evaluated. The new method differs from both the above in that we avoid the second part of their algorithms; instead, CLM searches for directions where activation, constrained by minimization in the tangent space, will lead directly to the transition state without the need for Hessian evaluation. This has one notable advantage; the CLM method is much less sensitive to certain types of noise.

As mentioned, certain (in particular electronic-structure) codes cause an artefact in their potential energy surface called "space-rippling"; this refers to the phenomenon where the energy (and therefore forces) on the system varies with the position of the system with respect to the origin. As we are working under periodic boundary conditions (in the general case), the energy of the system should be invariant under translation.

In the case of the SIESTA code, this space rippling arises from the calculation of some of the Hamiltonian and overlap matrix elements on a real-space grid (the advantage of this being that it enables the "divide and conquer" strategy which enables linear scaling). This is colloquially known as the "eggbox effect".

Consider a hypothetical sinusoidal eggbox field of periodicity 0.1 Åand energetic magnitude of 0.001 eV; these are realistic figures. Therefore, $E(x) = 0.0055 \sin(2\pi x/0.1)$ and, therefore, $d^2E/dx^2 = -1.97E(x)$; as can be seen, this therefore introduces spurious eigenvalues of the exact magnitude one is looking for, polluting the search methods proposed above (which depend, essentially, on being able to make an accurate variational search for the lowest eigenvalue of the Hessian). Furthermore, the larger the eggbox field in magnitude, or the shorter in period, the worse the problem

gets; therefore the problem is worse when one trades off performance for accuracy, making "quick and dirty" calculations entirely impossible.

This noise is an even greater issue in the methods of Henkelman and Jonsson and that of Munro and Wales, which depend on reasonable accuracy in a local variational search for the minimum-eigenvalue mode of the system; our search is on a much longer length-scale, and thus this becomes a non-issue - one simply neglects oscillations in the total energy of the system on the scale of the "eggbox", as we explain later.

To justify this statement, note that the variational method of Munro and Wales depends on minimization of the finite-differences approximation to the second derivative of energy with respect to a trial eigenvector of the Hessian;

$$\frac{\partial^2 E}{\partial \mathbf{x}^2} = \frac{E(\mathbf{x_0} + \Delta\mathbf{x}) + E(\mathbf{x_0} - \Delta\mathbf{x}) - 2E(\mathbf{x_0})}{(\Delta\mathbf{x})^2} \tag{4.39}$$

where $\mathbf{x}$ is the trial eigenvector and $\mathbf{x_0}$ is the position of the system. This is, for constant Δ, essentially equivalent to the Henkelman and Jónsson formulation – the Dimer Method [3] – as will now be demonstrated.

In the Dimer Method, the key quantity to be minimized in the eigenvector search is energy of a "dimer", defined as the sum of the energies of the two images of a system, lying at $+\Delta$ and $-\Delta$ along a trial eigenvector of the Hessian from a given configuration of the system. Let this configuration be x_0; therefore, the quantity minimized with respect to dimer rotation is $E(\mathbf{x_0} + \Delta\mathbf{x}) + E(\mathbf{x_0} - \Delta\mathbf{x})$ with respect to $\mathbf{x}$, the orientation of the trial eigenvector. In principle, upon minimization, this dimer will be oriented along the lowest-eigenvalue eigenvector of the Hessian of the system at that point. The actual minimzation of the energy of the dimer with respect to rotation is performed by an ingenious and extremely aggressive approximation (essentially, a modified Newton method); thus, for each dimer configuration, only

one rotation is performed per dimer-centre, rather than the full determination of the minimum eigenvector performed at each stage in the Munro-Wales method.

Returning to the Munro-Wales formulation, $\mathbf{x_0}$ is constant for any given eigenvector search; therefore, minimization of the partial derivative reduces to minimization of the numerator. This function is therefore identical to the quantity minimized in Henkelman-Jónsson up to multiplicative and additive constants, and both methods should return similar eigenvectors. Therefore, both these methods suffer from the same drawback – tremendous sensitivity to the space-rippling originated in codes which perform integrals on a real-space grid. Their primary differences come in how the step-length along the eigenvector is chosen (the Dimer Method uses a standard first-derivatives force-based minimization whilst reversing the force along the dimer - essentially, driving the system uphill, rather than downhill, along the minimum-curvature mode, whereas the Munro-Wales method uses a finite-differences approximation to an ideal Newton-Raphson step along the eigenvector combined with tangent-space minimization in the remaining degrees of freedom), and the aforementioned variation in approach to minimum eigenvector-finding at each stage. These are significant details in terms of implementation – and performance! – but the underlying ideas of both methods are remarkably similar.

One significant advantage, though, of CLM over these methods is that it is easier to implement – writing a CLM optimizer is, compared to the above methods, trivial. A standard conjugate-gradient minimizer from library code can be used, and all the other routines are acceptably efficient even in naive implementations.

The other method with which CLM has some similarity is Mishin's [72] algorithm, which he used for simulations of grain-boundary diffusion in perovskites. In his work, an ion is displaced along a fixed direction, and then the system is allowed to relax, using the Metropolis Monte-Carlo algorithm, in the tangent space of the atomic displacement. However, the present work has three advantages. Firstly, CLM converges to a genuine transition state. Mishin's method risks the same problem as IRC (ie, CLM without iterative direction update) – that of maximising in a direction

which lies outside the critical region of configuration space, and which will thus miss the transition state. This is a very serious defect; Mishin's method at best provides an approximate upper bound to the true transition state energy. Secondly, CLM uses a direct minimization method in the tangent space of our atomic displacement (as Mishin's studies concentrate on thermal effects, he is not able to do this), which is substantially more efficient. Thirdly, the CLM formalism is defined for arbitrary vectors in configuration space, rather than the displacement of a specific atom, which is much more general. However, Mishin's method does have one advantage over the present work – it includes the effect of finite temperature; therefore, it can be used to map the free energy surface directly, which this method cannot.

There is no guarantee that any given transition state found will be the lowest in energy. This is, in fact, desirable; for this class of methods it is much more important to find as many of the possible transition states from a given basin as possible; these can then be used (in, for example, kinetic Monte Carlo simulations) to determine the long-time-scale statistical thermodynamics of the system. If a specific transition state is being searched for, two-basin search methods are still more efficient.

In my method, the transition state found is highly dependent on the initial direction chosen; therefore, the algorithm parallelizes well, as one can simply select a random distribution of vectors, one (or more) per process, from the active subspace. There is still no guarantee that every (or any particular) transition state will be found, but this is true of any algorithm of this class; only a complete mapping of the potential energy surface can provide complete certainty, and that is only plausible for extremely small systems or exceptionally simple potentials.

4.4 Efficiency of Constrained Linear Maximisation

The efficiency of CLM as a method is primarily determined by three factors; the initial step size used in my method; what CLM terms the bracketing tolerance; and the algorithm used for reorientation of our search direction (and in particular the magnitude of δ, as defined above).

Let us consider each of these in turn.

4.4.1 Step size

The naive assumption to make is that the larger the initial step size one can use, the better; initially, one is attempting to bracket the transition state along our trajectory through the system (in the sense of determining a point which lies directly before it, and a point which lies directly after), and it seems obvious that the larger this initial step, the sooner bracketing will occur.

Unfortunately, this is not the case for two reasons. Firstly, in some systems it is possible to overstep the salient energetic features of the system. This will only occur if the initial step is so grossly large as to be of the same (or larger) length-scale as the distance between key features (transition states and minima) on the reaction coordinate, but this is not always something which can be ruled out. It should be noted that this can also prove to be an advantage if one has a spurious short-range energetic feature which one wishes to ignore, such as the egg-box.

Secondly, and more importantly, each step in our process does not exist in isolation. Consider Figure 4.5:

Firstly, the smaller step provides a tighter initial bracketing of the transition state. Secondly, as can be seen, the smallter the step, the fewer minimization steps required

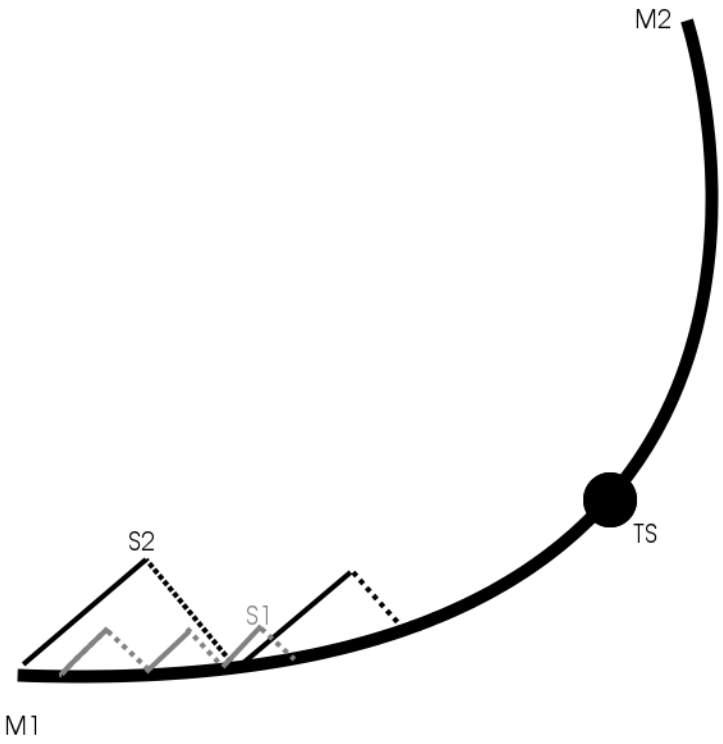


Figure 4.5: Trajectories through the system for two different step sizes; dashed lines represent minimization in the tangent space.

in the tangent space to reach the minimum energy trajectory; therefore, there comes a crossover point where the gain of locating the transition state in the linear space is outweighed by the increased cost of the tangent-space minimization.

Although we reduce the step-size once a transition state has been bracketed, we have not investigated the potential optimization of adapting the step size during the initial bracketing search. It is thought unlikely, given the measured performance data, that this will affect the efficiency of the method by an order of magnitude, but it would probably provide measurable performance gains.

4.4.2 Bracketing tolerance

Consider Figure 4.6. When a transition state is found, we return to before the transition state (by a distance of twice the curent step size to avoid misbracketing of our transition state) and halve the stepsize. When the stepsize is equal to or smaller than the bracketing tolerance (and the energy of the system appears to be continuous in γ), we stop and return the current position. This is the primary control on the accuracy of my method, as well as significantly affecting its efficiency; the looser the bracketing tolerance the less accurate the method is, but the faster it will converge.

In the majority of cases, energy converges faster than distance from the transition state; this is (of course) dependent on the curvature of the energy surface near the transition state. If the imaginary frequency of the mode corresponding to the transition state is small, a large bracket will provide reasonable accuracy in terms of TS energy, but for a high-frequency mode the same bracket will give very poor results.

At present, a fixed bracketing tolerance is used; therefore, for any given system we must set our tolerance to be acceptably precise for all the transition states in the system. Making this task more complex is the fact that one does not, *a priori*, know what these are; therefore, one must be extremely conservative in setting the algorithm's bracketing tolerance, which costs us somewhat in terms of performance.

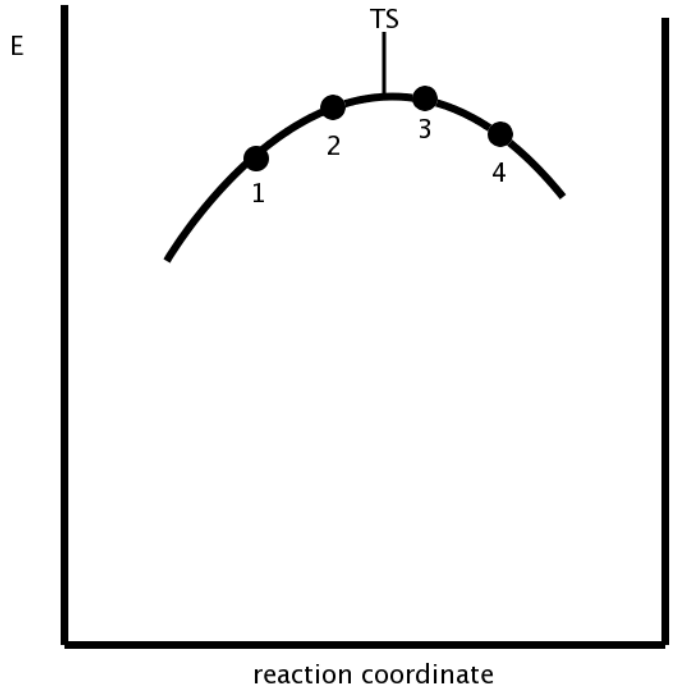


Figure 4.6: As can be seen, the calculated path energy only falls at step 4; by this point, one has already overshot the transition state.

An adaptive bracketing scheme would therefore be ideal; this is a goal for future research. One might, for instance, fit a quadratic to the last three points on the system's trajectory and from that estimate the curvature approaching the transition state; this would allow one to estimate the potential error in energies. An approach along these lines is likely to be rewarding in terms of improving the observed performance of the algorithm.

4.4.3 Reorientation algorithm

As noted above, to a certain extent the reorientation process is self-correcting; assuming that discontinuities are being accurately detected, and that δ (the magnitude of reorientation) is not excessively large, the system will tend towards the correct orientation. However, it is obviously desirable that one does not over-rotate (causing the search direction to undergo what amounts to a damped oscillation about the desired vector), or under-rotate (leading to slow convergence in the other limit); one approach which can be taken here is to fit δ empirically for best performance on a set of test potentials, and this is the approach the CLM algorithm has chosen to take.

It is likely that, like the bracketing tolerance, an adaptive scheme would provide better performance; this has, at present, not been investigated, though it would provide an interesting task for future research.

Additionally, there is substantial room for increased performance here; my algorithm is consciously very simple, and near to a transition state Hessian-based methods are exceptionally efficient at identifying the desired search direction. It is likely that the most efficient approach, in the long term, will be a hybrid - although at what point it would be best to switch from my method to eigenvector-following is not immediately obvious. Additionally, in the "eggbox effect" case, the effect of distortions of the energy space could well render this difficult or impossible; the option remains, in that case, to coarsely locate potential transition states at one, computationally cheap,

level of grid fineness, followed by performing variationally-based search with a much finer grid tolerance (which reduces the scale of the eggbox) for configurations known to be very close to transition states.

4.5 Tests on 2D potentials

My implementation of this algorithm was written using the CAMPOS framework. CAMPOS is an open-source suite of software (in the Python programming language) primarily written by the Centre for Atomic-Scale Materials Physics at the Technical University of Denmark and a worldwide team of volunteers; it is freely available for use and modification. Its primary feature is that algorithms such as ours can be implemented in a totally force-evaluator agnostic way: the same implementation can therefore be reused at all levels of approximation from entirely abstract potentials up to DFT methods (at present, CAMPOS supports the Dacapo and Siesta codes).

Therefore, as a further test for my algorithm, two two-dimensional test potentials were implemented. The quadratic potential has already been described above; the oter is the modified 2D London-Eyring-Polanyi-Sato (LEPS) potential first introduced by Jónsson, Mills and Jacobsen [69], and later modified by Henkelman and Jónsson [3]. The LEPS potential itself, introduced by Sato as a modification of the work of the other three eponymous scientists in 1965 [73], is a model for the interaction of three atoms (A, B and C) confined to motion along a line. Only one bond can be formed, between either of the pairs of adjacent atoms, AB and BC; therefore, there are two degrees of freedom, r_{AB} and r_{BC}, to this problem. By fixing the positions of A and C, this falls to one degree - the position of B on the line. One can now reintroduce a second degree of freedom (producing a more complex, and thus testing, energy surface), by coupling the 1D LEPS system to a harmonic oscillator; the functional forms in question can be found in Henkelman and Jónsson's paper [3], but an isosurface plot of the energy surface can be seen in figure 4.7.

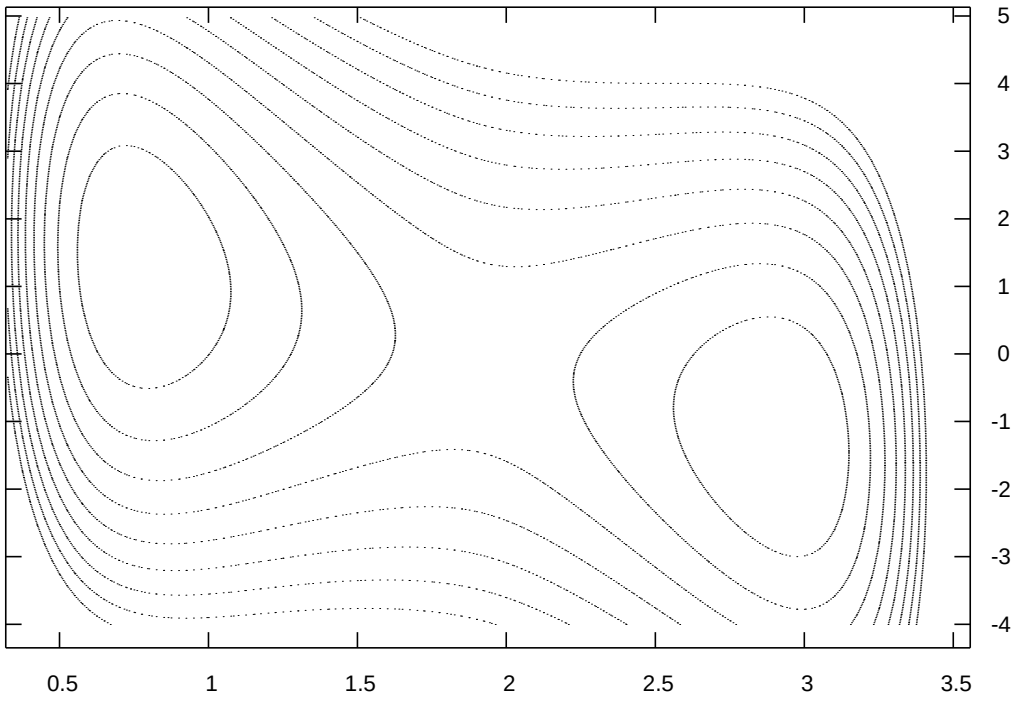


Figure 4.7: 2D LEPS surface

Table 4.1: Parameters for the Gaussian functions added to the modified LEPS potential, after Henkelman and Jónsson [3].

i	1	2
A_i	1.5	6.0
x_{0_i}	2.02083	0.8
y_{0_i}	-0.172881	2.0
σ_{x_i}	0.1	5.0
σ_{y_i}	0.35	0.7

This surface only has one transition state and two minima, but even in this case the na ive Intrinsic Reaction Coordinate method (taking the linear interpolation between the minima the as our postulated reaction coordinate) fails. In order to add extra transition states to this surface, Henkelman and Jónsson added two Gaussian functions of the form

$$G_i(x,y) = A_i e^{-(x-x_{0_i})^2/2\sigma x_i} e^{-(y-y_{0_i})^2/2\sigma y_i} \tag{4.40}$$

where the values taken for the free variables are given in Table 4.1. The resulting energy surface is shown in Figure 4.8.

This surface has three minima and four transition states.

In order to determine the performance of my algorithm, sets of 500 CLM runs were calculated from each of these three basins. Runs which ran straight uphill (determined by their reaching an energy 10 units greater than the initial basin) were restarted on a new random maximization vector and their step-count reset to zero, as in a real system the restriction of our initial maximization vector's degrees of freedom (and the use of periodic boundary conditions) make it extremely unlikely that such a direction will be selected. The initial step-length was varied as denoted below. A convergence tolerance of 0.01 units was used, providing accuracy of substantially

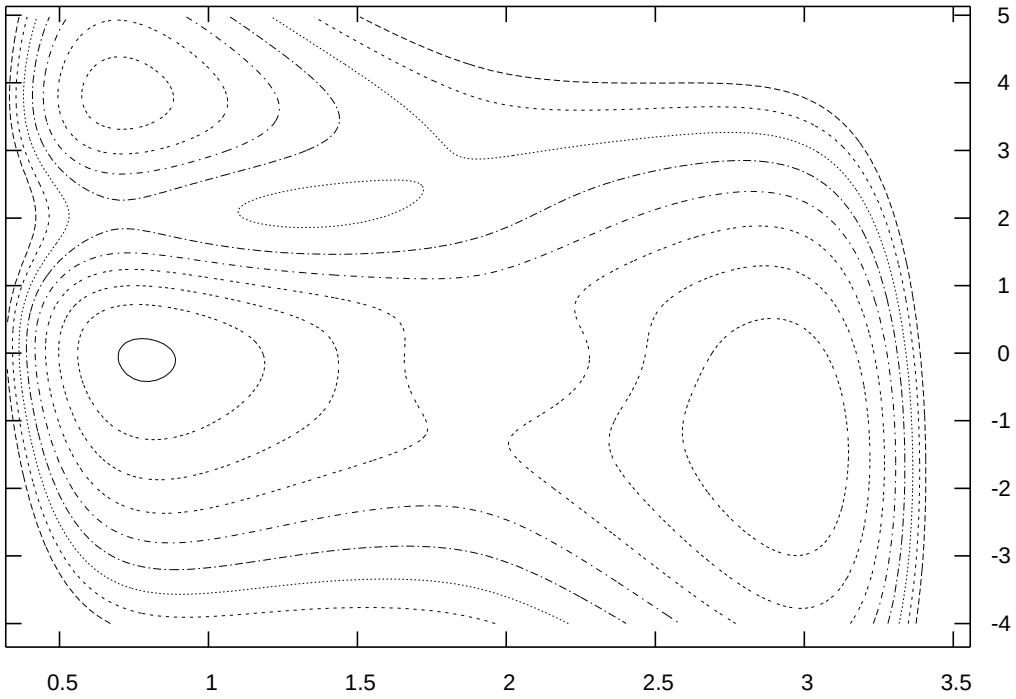


Figure 4.8: 2D LEPS surface with added Gaussians

Start basin	TS1 (1.11)	TS2 (1.55)	TS3 (3.17)	TS4 (3.96)	Mean number of steps
1	93	138	232	37	639.9
2	92	156	218	34	607.4
3	0	0	360	139	452.4

Table 4.2: Results, stepsize=0.05; Note that transition states 1 and 2 are inaccessible from basin 3.

Start basin	TS1 (1.11)	TS2 (1.55)	TS3 (3.17)	TS4 (3.96)	Mean number of steps
1	151	124		37	514.3
2	137	139	222	2	534.3
3	0	0	332	154	481.6

Table 4.3: Results, stepsize=0.1; Note that transition states 1 and 2 are inaccessible from basin 3.

better than 0.01 units in the calculated transition state energy. The results can be seen in tables 4.2 and 4.3, and in figure 4.9.

As can be seen, and (indeed) as expected, the initial stepsize has a relatively small effect on the performance of the algorithm. The impact of the bracketing tolerance on the accuracy and performance of our method was investigated next. Starting from Basin 1, sets of 500 runs were calculated with a stepsize of 0.05 and bracketing tolerances of 0.01, 0.02 and 0.04; this caused a fall in mean-number-of-evaluations from 640 to 513 in the latter case, as is represented in Figure 4.10.

There is a cost in terms of the well-convergedness of our transition states; this can be seen in Figure 4.11, comparing the final converged transition state positions for bracketing tolerances of 0.01 and 0.04 units. However, the latter is far from useless; as can be seen, it still reliably locates the vicinity of the transition states, suggesting our method has some value for a quick, approximate calculation.

One advantage my method possesses over Henkelman and Jónsson's is the simplicity of generating initial search directions; they start by displacing some number of atoms around the minimum configuration and performing an initial dimer rotation, whereas we simply need to select a vector in a subspace around the origin. The strongest

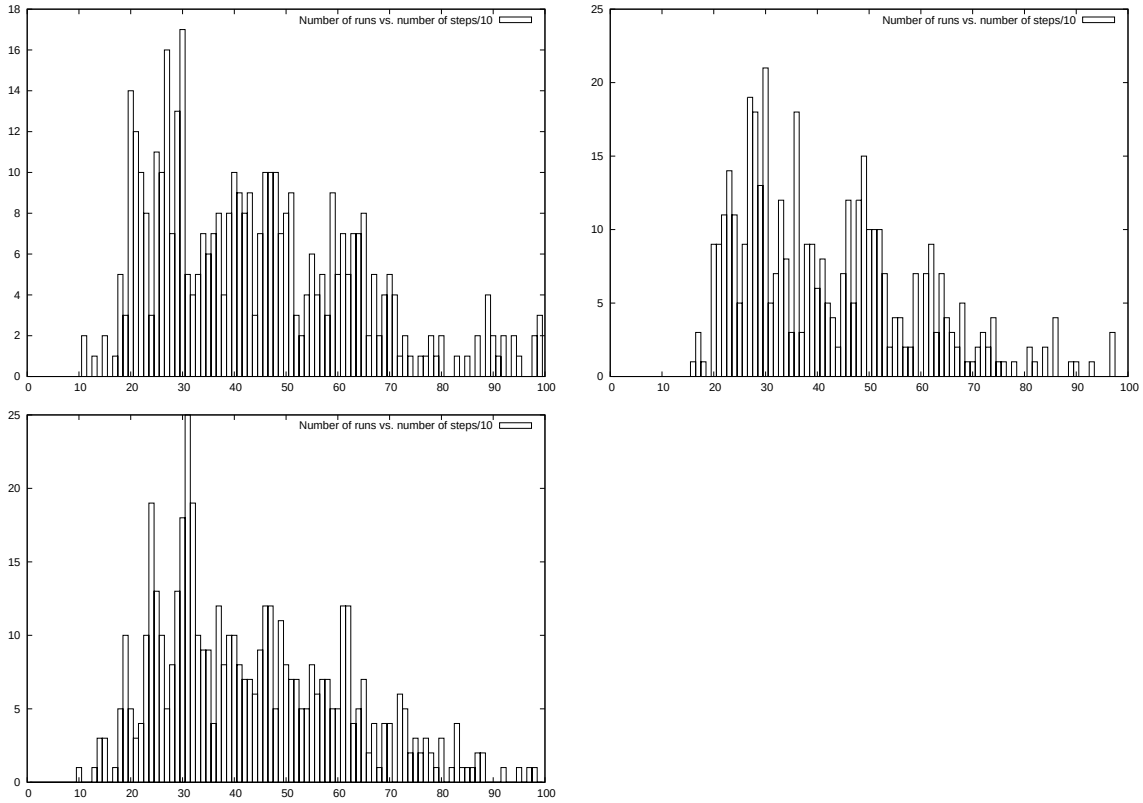


Figure 4.9: Histogram of number-of-step distribution for runs from each basin; step-size = 0.05

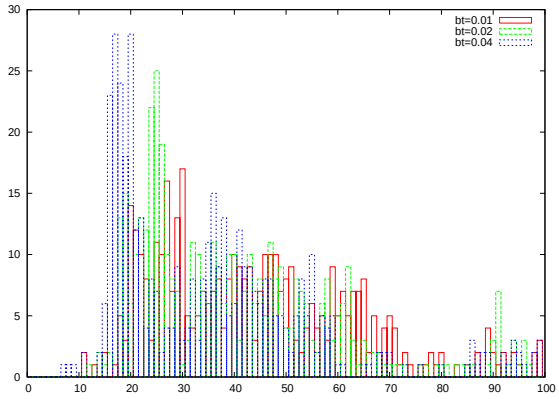


Figure 4.10: Histograms of (Number of force evaluations / 10) versus count for each of three bracketing tolerances; note the general shift of features of the histogram to the left.

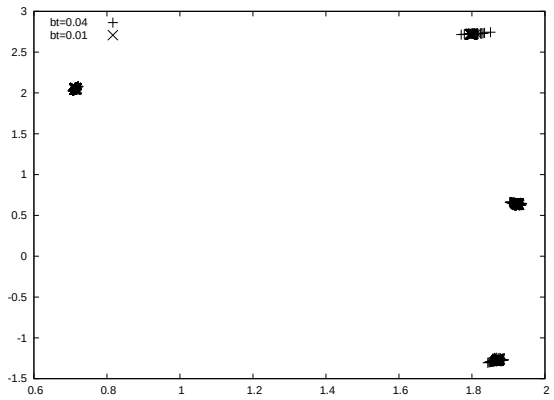


Figure 4.11: Comparative positions of our calculated transition states on the modified 2DLEPS potential for bracketing tolerances of 0.01 and 0.04 units.

feature of my method is its algorithmic and conceptual simplicity.

Henkelman and Jónsson do not give specific numeric performance figures for the efficiency of their method on this potential in their original paper. Assuming, as is implied in their paper, that each dimer calculation takes six evaluations, consideration of their presented trajectories implies that their successful calculations take between 150 and 200 force evaluations on this surface. My method is around two to two and a half times slower, but very little effort has been placed into optimizing the algorithm; with careful choice of tolerances, CLM appears to be competitive, and has the advantages mentioned previously which may make it a better choice for some systems.

A key point to be made here is that one can conceptually divide the search for transition states into two parts; transition state *location* and transition state *optimization*. Most methods excel at one or the other; close to a transition state, Hessian-based methods (Rational Function Optimization (RFO) methods [74] (following from the pioneering eigenvector-following work of Cerjan and Miller [75]), Newton-Raphson, Munro-Wales, Henkelman-Jónsson Dimer, and so forth and so on) are most efficient, because in the general case identifying the correct step direction (the eigenvector associated with complex frequency) is very easy. However, far from the transition

states, these methods are often rather poor. The Dimer method requires very careful selection of step-size and related parameters for best performance, both in terms of number of force evaluations needed and in terms of number of transition states found; transition state discovery can actually be harmed by overconvergence on the minimum eigenvector at early stages in the calculation. This is best illustrated by the review article of Olsen and collaborators published [76] published in 2004. In the terminology introduced here, these methods are very good at transition state optimization but less good at location.

My method has the other problem; it identifies the locales of transition states very rapidly, but optimizing the final transition state is slow, and in systems with highly curved energy spaces (high-energy events - for instance, results (presented later in this work) pertaining to the simulation of self-diffusion in silica glass) can be somewhat inaccurate. In the case of CLM, this is primarily due to the difficulty of setting the free variables which detect the onset of a discontinuity in the search direction; if one is missing small discontinuities close to the transition state one can end up with a systematic overestimate of the transition state energy.

Also, the convergence of the method is poor close to the transition state; my method excels at location, which is desirable when searching a large parameter space for transition state basins, but in some cases needs to be paired with another method to optimize the results within those basins.

Of course, as mentioned earlier, it is possible to switch methods, either by hand or automatically; once a transition state has been bracketed to within a certain tolerance, one could easily switch to either a variational or explicit eigenvector-following method. My work on framework silicates takes this approach; the transition states presented were first identified with CLM and then optimized using the RFO method.

In the present work, the free parameters in CLM have been set throughout as outlined in the work on the LEPS2D potential above; this potential has the right length-scale for interatomic processes (and, as a cheap-to-evaluate model, has allowed the above extensive testing).

4.6 Platinum (001) surface

The (001) surface of face-centred cubic platinum, simulated using effective-medium theory, was used to test the algorithm. It has been known since Kellogg and Feibelman's experimental paper of 1994 [77] that adatom surface diffusion does not progress by simple hopping; rather, a concerted exchange mechanism takes place, where the surface adatom forms a local dimer with another atom, kicking it out of the surface layer and replacing it.

This was extended by Wei Xiao et al. [78], when they considered the effects of strain; they found that extremely extended transition states, "crowdions", existed (by embedded-atom simulation). Therefore, this is an ideal test system – being able to find these multi-atom transition states would demonstrate that the initially-local choice of trial search direction in CLM is not a major weakness.

We used an 1806-atom slab of platinum, with one surface adatom and the lowest of the five 19×19 slabs fixed. The lattice parameter was set to 3.975 Å, and a biaxial strain was applied down x of 2%.

We found multiple transition states, including direct hopping, surface exchange, and crowdion formation; this is a powerful demonstration of the effectiveness of CLM. The energies of the transitions are not meaningful, given the inherent limits of the forcefield used; it was sufficient for my purposes that multi-atom transition states existed, and no more.

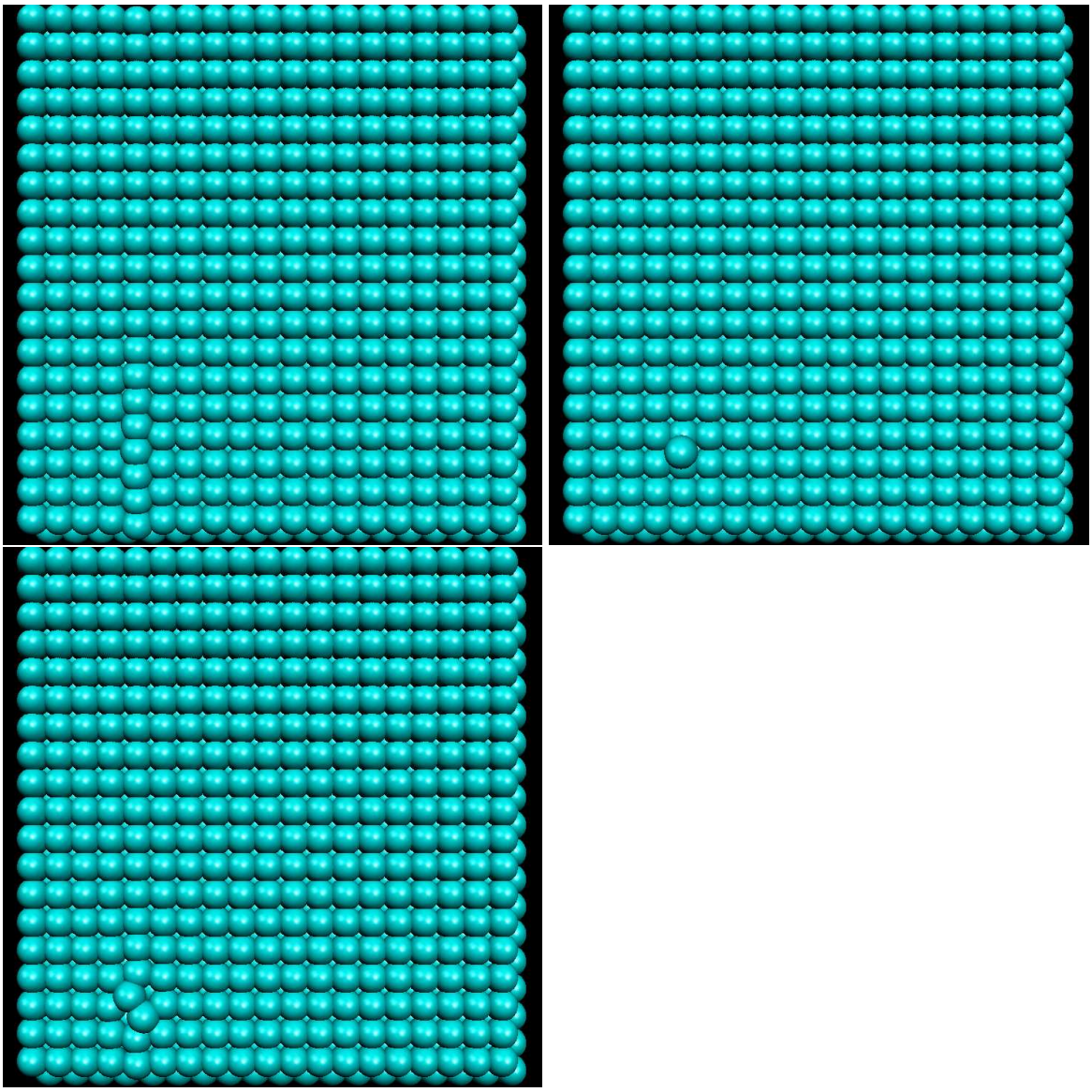


Figure 4.12: Transition states on the Pt(001) surface; from the top, a crowdion, a direct hop and a dimer exchange.

Chapter 5

Diffusion of Li^+ in quartz, cristobalite, and tridymite

5.1 Introduction

Silica (SiO_2) is a key material in many modern engineering applications, whether as a timing crystal (quartz), as the amorphous form in digital electronics, or in another of its many applications. Additionally, the majority of technologically important glasses for engineering applications are silicate-based; in these, doping to adjust certain physical properties (a trivial example being the lowering of the glass transition temperature induced by alkali-ion doping in sodalime window glass) is essential, and the diffusion of dopant species in silicate glasses has, as such, been been a subject of active research for over forty years [79].

In this chapter, the diffusion of Li^+ ions in crystalline SiO_2 is considered. At laboratory conditions, the stable phase of SiO_2 is the low phase, α-quartz [80]. Therefore, it is this to which I pay most attention. I also perform comparative studies, in substantially less detail, of low tridymite and low cristobalite; they have structures derived from rings of SiO_4 tetrahedra rather than the helices evident in quartz, and hence

might be a better archetype for local environments for Li^+ in glass (which I consider in the next chapter). Past simulation studies of these classes of materials have used either molecular-dynamics/forced-diffusion approaches [26] or approximate methods based around rigid-unit relaxation of the tetrahedral Si–O framework [6]. Instead of using these methods, I present results obtained from a direct search of the potential energy surface as modelled by a classical forcefield.

From consideration of the connectivity of our library of identified stable sites and transition states, one firstly identifies the complete mechanism for c-axis diffusion of Li^+ in α-quartz when it is unbound from Al^{3+}, including accurate structures for all the stable and transition states therein. The stable states along this pathway accord well with experimental measurements of stable Li^+ sites by X-ray diffraction. As anticipated, a strong affinity between Li^+ and (impurity) Al species is found; having observed broken symmetry in the diffusion mechanism, we consider the problem of comparatively-long-distance Li–Al interaction. By considering a situation where the charge associated with the Al^{3+} defect is distributed over all the Si/Al sites in the lattice, it is found that induced lattice distortion by Al/Si substitution is a more significant effect than direct Coulombic interaction at Li–Al distances greater than approximately 5Å and is sufficiently substantial to change the order of stability of stable sites within [001] diffusion channels.

Further to this, simulations of lithium diffusion in two other silicate polymorphs, low tridymite and low cristobalite, was performed; a comparative study of diffusion in these structures is presented, as a precursor to considering the problem of Li^+ diffusion in silica glass (in the next chapter).

5.2 Method

5.2.1 Supercell and potential

For all the systems, a classical force-field was used; energies and atomistic forces were calculated using GULP [21]. The potentials were derived by Calleja et al. [26], building on the work of van Beest et al [27]. Van Beest's potentials are known to reproduce the structural and elastic properties of quartz to an acceptable degree of accuracy; Calleja and colleagues then used this forcefield as a basis for theirs, relaxing α-spodumene with their trial field and fitting the Li–O interaction to obtain good correspondence with available experimental data.

The forcefield takes the form of a set of of Buckingham potentials (the first two terms of the following equation), plus a Coulombic term to account for electrostatic interaction:

$$V(r) = A_{ij} \exp\left(-r\rho^{-1}\right) - C_{ij} r^{-6} + q_1 q_2 r^{-1} \tag{5.1}$$

The parameterisation of this forcefield, and the value for formal charges used in the simulation, is given in Table 5.1. The same forcefield was also used by Wells, Sartbaeva et al. in their simulations [6, 17]; I refer to their results in due course. Using this forcefield I constructed, as outlined below, a supercell of each of the systems under study. Each supercell was then relaxed (using a variable cell) in P1 symmetry; having performed this. a single substitution of Al for Si was made, and a well-separated Li^+ ion added to the system – thus retaining charge balance. The generated system was then re-relaxed (this time keeping cell shape and volume constant) before attempting transition state search. Periodic boundary conditions were used for the simulation.

Initial calibratory runs were then performed, consisting of a Constrained Linear Maximization (CLM) run directly followed by an energy minimization, until it was

Species	$A(eV)$	ρ(Å)	$C(eV/Å^6)$
O–Si	18003.757	0.205200	133.54
O–O	1388.7730	0.362320	175.00
Li–O	1149.4126	0.280490	0
O–Al	16008.535	0.208480	130.57
Li–Al	35985.486	0.215695	189.10
Li–Si	18080.333	0.206049	153.87

Species	Charge (e^-)
Li	+1.0
Al	+1.4
Si	+2.4
O	-1.2

Table 5.1: Parameters for the force-field used in these simulations.

clear that the system was well-minimized with the Li+ atom residing in a different channel in the structure to the Al^{3+}. This provided an initial configuration from which to search.

Quartz cell

Quartz, under laboratory conditions, has two commonly encountered polymorphs; a low-temperature trigonal phase (α-quartz) and a high temperature hexagonal form, first identified by Le Chatelier in 1889 [81]. These are separated by a displacive phase transition at 846K. This transition is driven by a soft rigid unit mode [82]; it preserves the Si-O-Si bond angle, and hence is low-energy, minimizes distortion of the SiO_4 polyhedra, and requires no change in the topology of the structure. As this is a zone-centre transition [83], the unit cell contains the same atoms in both the high and low phases; there are three formula units per unit cell, and hence nine atoms.

Both these structures are composed of helices of SiO_4 tetahedra, with channels down the *c*-direction; this is to be contrasted with tridymite and cristobalite, both of which are made up of rings of similar corner-linked SiO_4 units.

	Simulated	Undoped α-quartz at 20 K [80]
a ($\AA$)	14.829	14.7057
b ($\AA$)	14.832	14.7057
c ($\AA$)	10.867	10.7976
α	89.69°	90°
β	90.21°	90°
γ	120.16°	120°

Table 5.2: Parameters for the α-quartz supercell, compared against experiment. As can be seen, c is 0.6% larger, and a 0.84% larger, in this doped cell than in the experimental undoped values at 20 K measured by Tucker et al; this is regarded as acceptable, certainly within the spirit of empirical-potential calculations.

The present study uses a 3×3×2 supercell of low quartz, which allows both the Li^+ ion not to self-interact in the z-direction – the predicted primary direction of motion – and also for there to be one full channel between Li^+ and Al^{3+} in the xy-plane, minimizing the effect of Li–Al interactions. There were, therefore, 163 atoms in the supercell used in the calculations; the geometry of the supercell is given in Table 5.2.1.

The structures of tridymite and cristobalite

It is worth considering the structures of tridymite and cristobalite together. These, similar to quartz, undergo displacive phase transitions driven by rigid unit modes [84, 85]; these lower the symmetry from hexagonal (high tridymite) or cubic (high cristobalite) to orthorhombic (although more on this later) and tetragonal respectively.

Both of these phases are metastable with respect to quartz at low temperature. They are composed of sheets of hexagonal rings of SiO_4, corner-sharing, with adjacent tetrahedra oriented alternately up and down [83]. The structures of both these phases can be derived, topologically, from polymorphs of solid carbon; by replacing

	Simulated	High-temperature aristotype, experimental [4]
a (Å)	18.693	20.20
b (Å)	19.393	20.20
c (Å)	16.074	16.48
α	90.11°	90°
β	89.87°	90°
γ	118.71°	120°

Table 5.3: Parameters for the tridymite supercell; experimental data from Kihara et al. [4]

C with Si, placing oxygen atoms between each Si–Si pair (expanding the structure so that this is energetically reasonable), and allowing the structure to relax. In particular, tridymite can be derived from the lonsdaleite (hexagonal diamond) structure, with adjacent layers stacked in an AB'AB'AB' pattern; the B' layers are laterally translated and inverted A layers. Cristobalite can be derived from the silicon/diamond structure; it has an ABCABC stacking arrangement, and layers parallel to the (111) plane.

Tridymite cell

A larger supercell of tridymite, containing 385 atoms, is used; this corresponds to a 4×4×2 supercell of the high-temperature aristotype. The structure of tridymite at low temperature is complex and heavily history-dependent [86], so in the semi-quantitative spirit of these simulations it was decided to use a large supercell of the high-temperature phase and simply relax it at low temperature rather than attempting to recreate the incommensurate geometry of the low-temperature phase(s). The geometry of the supercell is given in Table 5.2.1, with reference to the high-temperature phase.

	Simulated	Experimental [5]
a ($\mathring{A}$)	14.765	14.871
b ($\mathring{A}$)	14.765	14.871
c ($\mathring{A}$)	13.202	13.7806
α	90.00°	90.00°
β	90.00°	90.00°
γ	90.00°	90.00°

Table 5.4: Parameters for the α-cristobalite supercell; experimental data from Pluth et al. [5]

Cristobalite cell

The cristobalite supercell used contains 217 atoms (a 3×3×2 supercell once again); a substantial (around 4%) shortening was observed in the c lattice parameter, but the topology of the structure remains correct; therefore, although our results are approximate (which, given the potential, was always the case), this is acceptable within the spirit of our efforts here.

5.2.2 Event discovery

The method of Constrained Linear Maximization, as derived in the last chapter, was used to build a library of transition state events in each supercell. As the present study is only, in the first instance, interested in events which result in a displacement of Li^+ from its original location within the lattice, initial trial displacement vectors were constructed by choosing random unit displacements on the Li^+ atom; however, no constraint was placed on the ultimate value of the search vector, other than to remove the possibility of choosing a vector corresponding to bulk translation of the system by fixing the position of a single atom within each of the systems we consider.

This removes exactly three degrees of freedom from the system, corresponding to bulk translations along x, y, and z.

This is only half of the problem, though; although it is highly likely, it is not certain that the transition states thus found are directly connected to the original minima from which the searches were performed. Therefore, a three-stage protocol was adopted throughout. Firstly, a set of candidate transition states were identified with CLM. Secondly, to ensure extremely accurate convergence to the transition state, we converged our candidate states using Rational Function Optimization [74] (in the form of the RFO optimizer in GULP) to a saddle-point of order 1 (i.e., our transition state); displacements and changes in energy were small (of the order of 0.01 eV), as one would expect. It was then possible to take these refined structures and perform an RFO optimization from the transition state to a minimum. The resulting set of minima are connected to the transition state (which of the two minima connected to a given transition state is selected by this procedure is essentially determined by noise); it will not follow the steepest descent path during the optimization, and hence will not follow the exact reaction pathway, but for present purposes – one is solely seeking extrema – this is not an issue.

Thus, two sets of special points on the energy surface are obtained, minima and transition states – and, also, a set of connections between members of these two sets. These can therefore be represented as vertices on a graph, each transition state being connected to at most two minima and each minimum being connected solely to transition states. This presents the nature of any activated processes which may have been found in a particularly clear way – any path between two minima on this graph corresponds to a physically-possible diffusion pathway within the system.

This kind of graph layout can be automated; in the present work, we use the *dot* language and the `graphviz` graph-layout suite [87] to produce our energy surface diagrams. For small systems, it would be plausible to draw graphs by hand, but for extended or disordered systems it would rapidly become time-consuming and error-prone; instead, it is near-trivial to write a small program which outputs *dot*–format

lists of vertices (both minima and transition states) and edges.

5.3 Quartz

5.3.1 Prior results

In quartz, it is well-known that diffusion is markedly more rapid along the c-axis when compared to [110]-type directions, due to the presence of large channels in the structure down the c-axis; these arise from the highly directional helical structure of the material. Therefore, in looking at Li^+ diffusion in this phase, our primary goal was to attempt to identify the stable sites for Li^+ in the primary channel in the α-quartz structure (in other words, down the c channel in the structure), and the transition states connecting these stable states.

Experimentally, Campone et al. [88] found an activation barrier of 0.25 eV in β-quartz for c-axis diffusion, whereas Sartbaeva found 0.29 eV in β-quartz and 0.68 eV in the α phase. This process has also been simulated before; two approximate methods have been used, in both cases using exactly the same forcefield as being used here. Firstly, there is an applied-electric-field forced-diffusion molecular dynamics simulation by Calleja [26] et al.; secondly, a hybrid geometric-algebra approach taken by Sartbaeva et al [6, 17]. Calleja found than an unpinning energy of 0.11eV was required to initiate Na^+ diffusion (in the [001] direction) in α-quartz at 300 K; Sartbaeva and Wells found that the minimum activation energy for Li^+ diffusion, using the same potentials as the present study, was approximately 0.25 eV. The key insight of Sartbaeva et al.'s simulations [6, 17], however, is the role of channel flexibility in the diffusion proces; with their geometric-algebra approach, they suggest that oscillations in the position of oxygen atoms of up to 0.5 Å can be accommodated at a cost of no more than approximately 0.5 eV. It is particularly notable that the simulated values are substantially lower than the values identified by experiment.

5.3.2 Calculated activated pathways

The topology of the potential energy surface for α-quartz, in the presence of a single Al^{3+} substitution, is presented in Figure 5.1. Looking at this, there is one clear highly extended process; configurations on this on this are labelled M1 to M5 (and T1 to T5. accordingly – M for minima, T for transition states). These correspond to diffusion up the [001] direction, as presented in Figures 5.2 to 5.10. In all these figures, oxygen atoms are red; silicon atoms are blue; lithium and aluminium atoms (which should be unambiguous by context) are both purple. Each of these figures consists of a pair of images of the same structure; one oriented down the [001] direction (on the right), and one oriented so that [001] runs up the page (on the left).

Therefore, the mechanism for diffusion up the c-axis, using the labels in figures 5.2 to 5.10, can be seen to be M5 (fourfold) to T5 (activation energy 0.053 eV) to M2 (fourfold) to T3 (activation energy 0.194 eV) to M3 (fourfold) to T4 (activation energy 0.036 eV) to M4 (twofold); as we can see, that takes us through one full unit cell in the $c-$direction, and diffusion will continue on upwards through a transition state similar to T5.

It follows that an activation energy for this process ($c-$axis diffusion) of approximately 0.19 eV is therefore found, lying between the values found by Calleja and Sartbaeva; a first move for [110] diffusion was also obtained, and thus one finds an activation barrier for that process of 0.75 eV. This is substantially lower than the transition-state energy found by Calleja, by molecular dynamics at 300 K under applied field, of 1.55 eV; regardless, [110] is unambiguously the “slow” direction for diffusion in this structure.

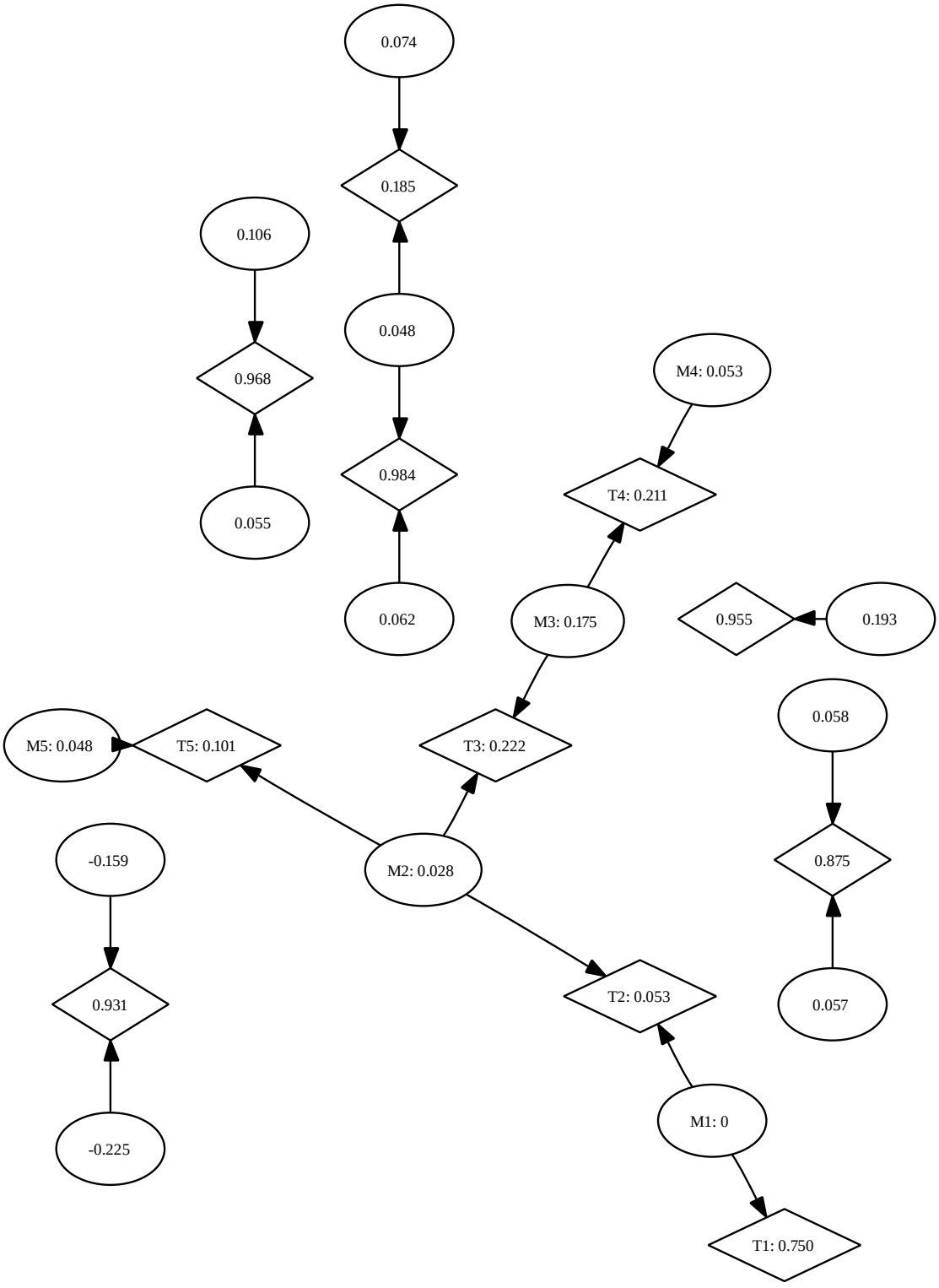


Figure 5.1: Connectivity of minima and transition states found in singly Li^+-doped α-quartz; diffusion up the c-axis is shown in the labelled vertices M1–M5 and T1–T5. Ellipses represent stable states, diamonds transition states; all energies are presented in eV.

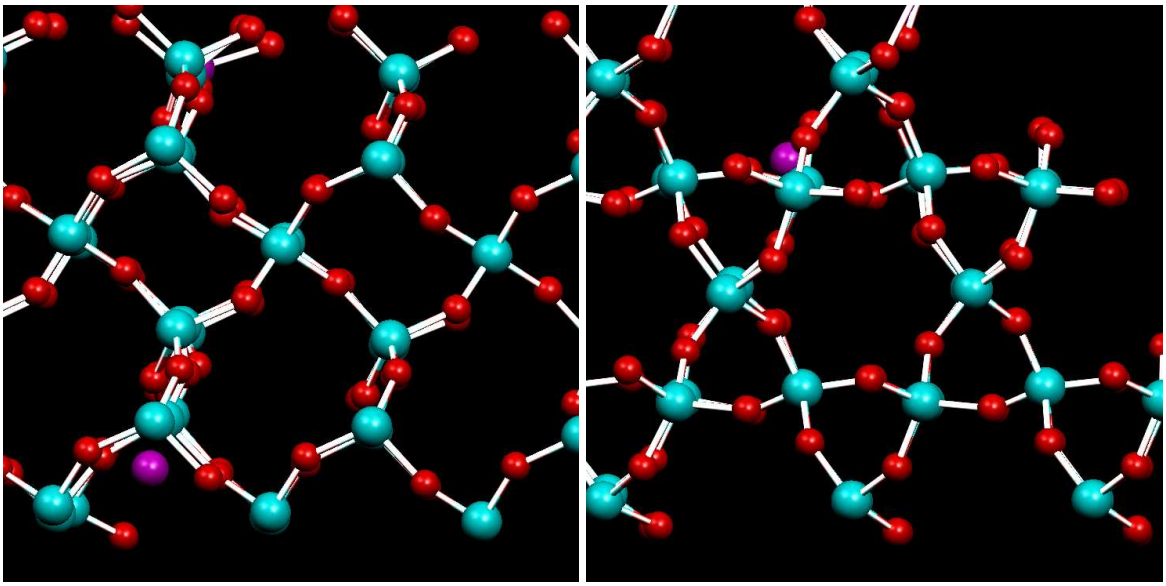


Figure 5.2: Transition state T1 (0.750 eV), connected to M1 (the "wall" site, Figure 5.3). This is the activation barrier for diffusion through channel walls, and hence in [110]-type directions in this structure.

5.3.3 Stable states

One point of interest arising from these calculations is that we identify a stable state of the system which was not picked up by Sartbaeva et al [6]; whereas they identify a twofold site (our Figure 5.8) and a fourfold site (our Figure 5.6), they neglect our "wall" site. This may explain their overestimate, compared to present simulations, of the activation energy for diffusion: under the assumption that the system progresses directly from the wall site to the transition state leading to the fourfold site, one would obtain an activation energy, from our calculations, of 0.22 eV (compared to Sartbaeva's value of 0.19 eV). It is difficult to be certain, but given the approximate nature of the method used in their work, it seems plausible that this is one of the causes of the difference between our results and theirs.

The two low-energy stable sites, M1 (Figure 5.3, the wall site), and M2 (Figure 5.6, the fourfold site), are extremely close in energy; the latter is metastable by only

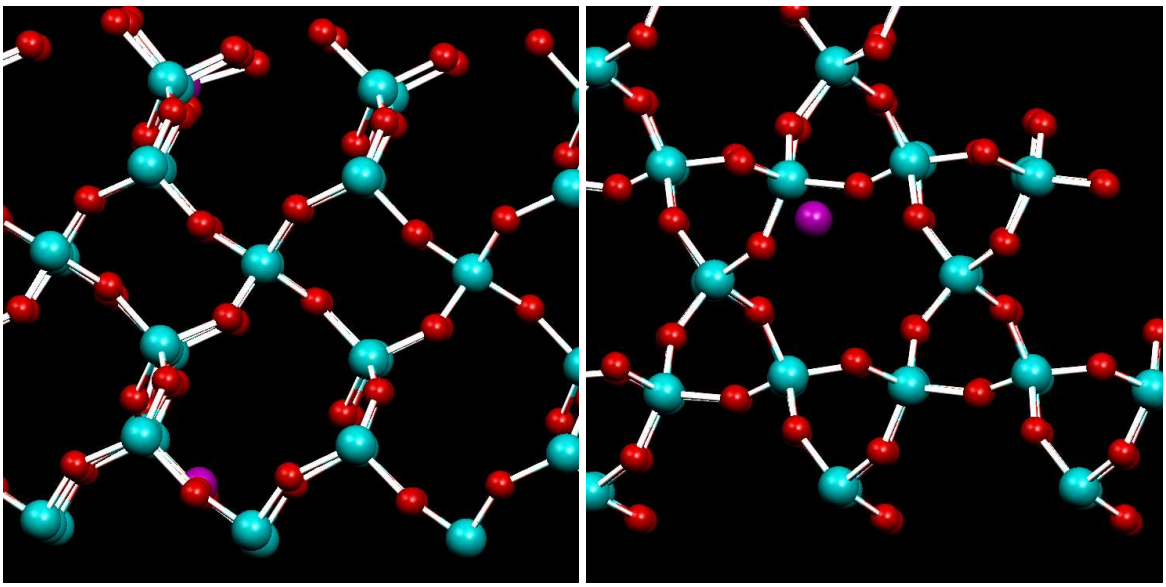


Figure 5.3: Minimum M1: the lowest minimum found on the *c*-axis mechanisml this is what has been dubbed a "wall" site, residing (as it does) in a [100]-channel wall. Connected to T1 (Figure 5.2) and T2 (Figure 5.4, the twofold site using Sartbaeva [6]'s notation.)

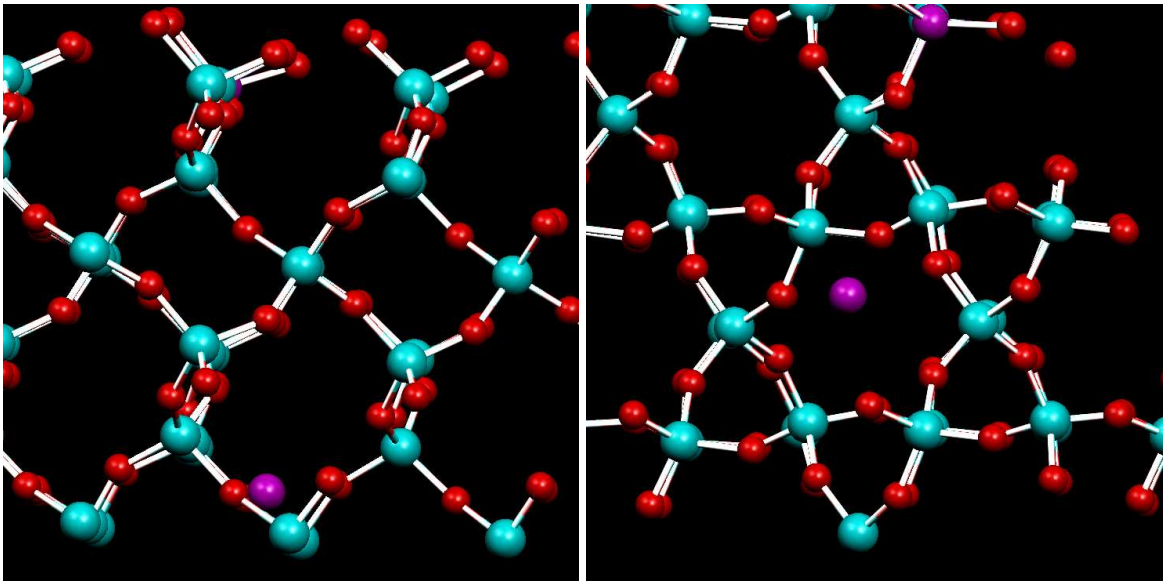


Figure 5.4: T2 (0.053 eV): transition state connecting M1 (the wall site, Figure 5.3) and M2, the twofold site (Figure 5.6.)

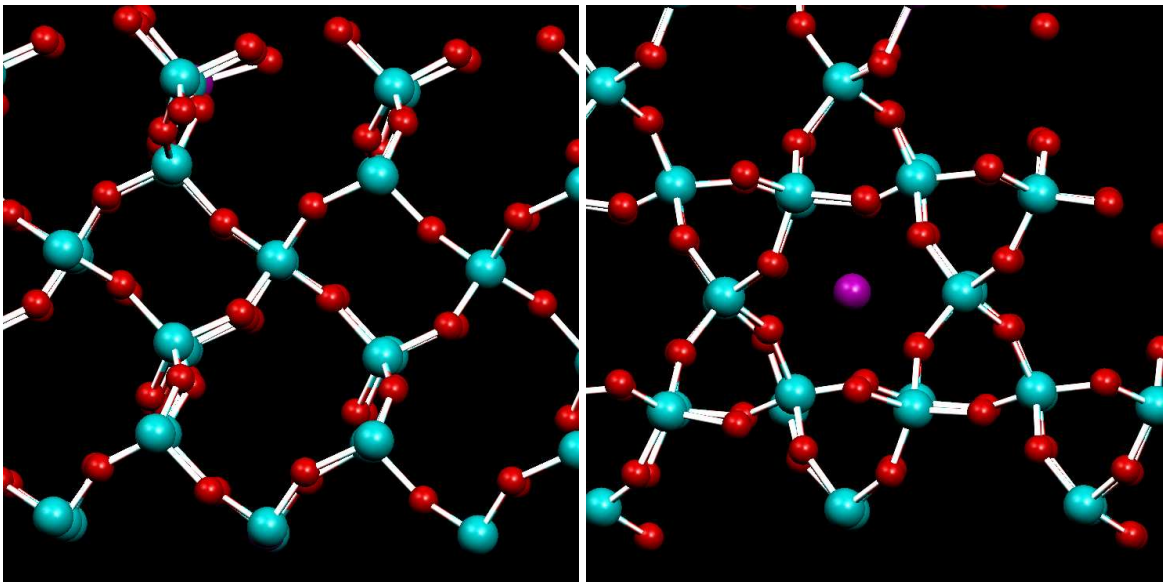


Figure 5.5: T5 (0.101 eV): this transition state, showing twofold linear coordination of Li^+ by Si^{4+}, connects M2 (Figure 5.6 with the equivalent twofold site directly below it (M5, not shown).

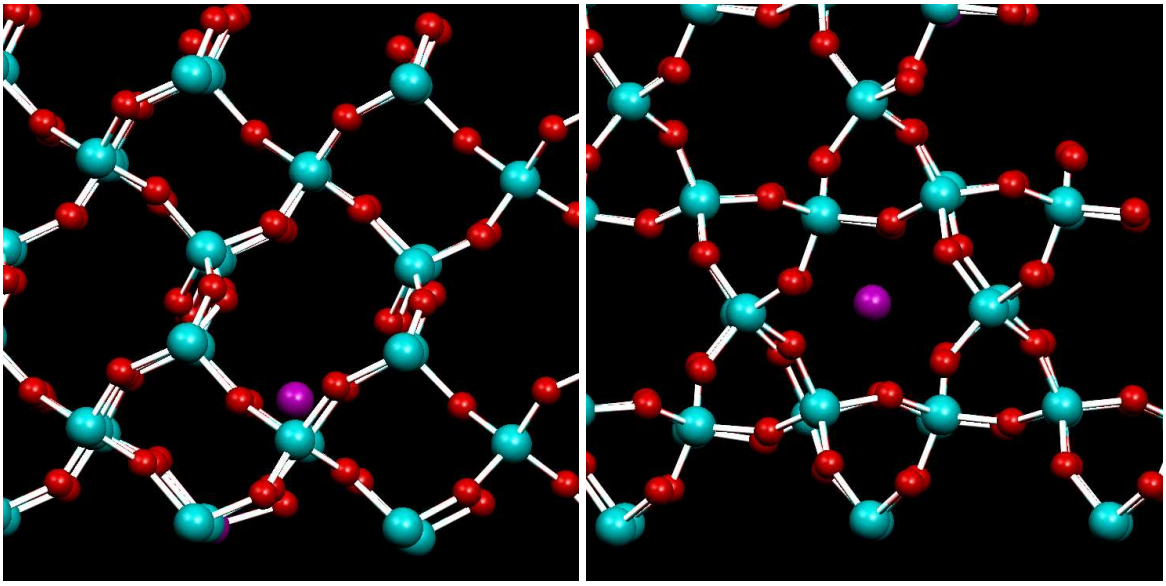


Figure 5.6: M2 (0.028 eV): connected to M1, the wall site, by T2 (Figure 5.4). This is Sartbaeva and Wells' "fourfold" site, as can be seen on the left.

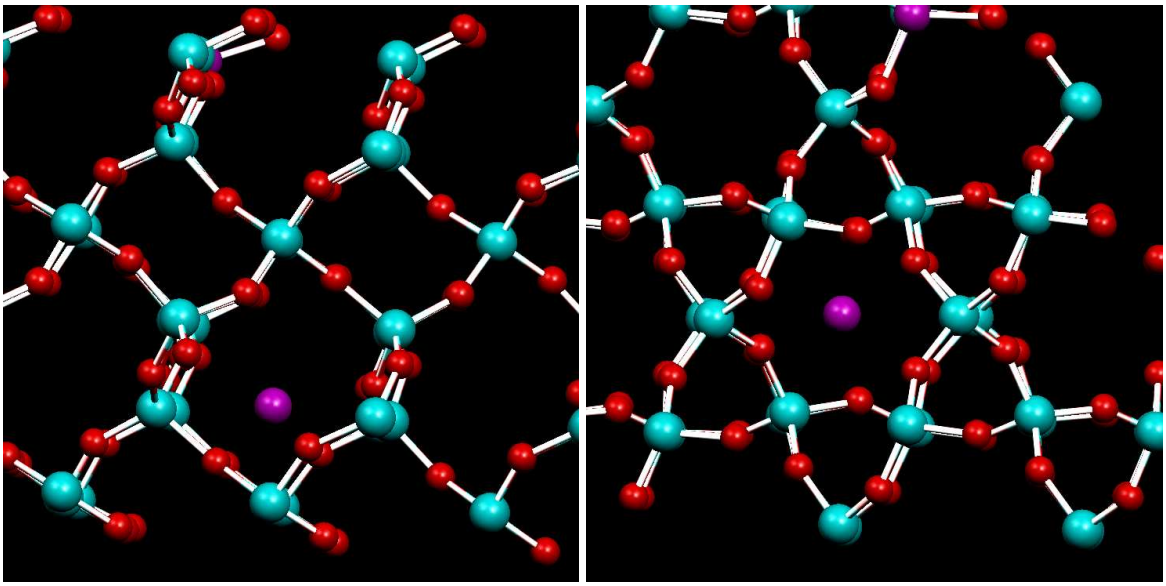


Figure 5.7: T3 (0.222 eV): connects M2 (Figure 5.6, the fourfold site, with M3, the twofold site (Figure 5.8). This is the rate-determining step for [001] diffusion, with an activation barrier of 0.194 eV.

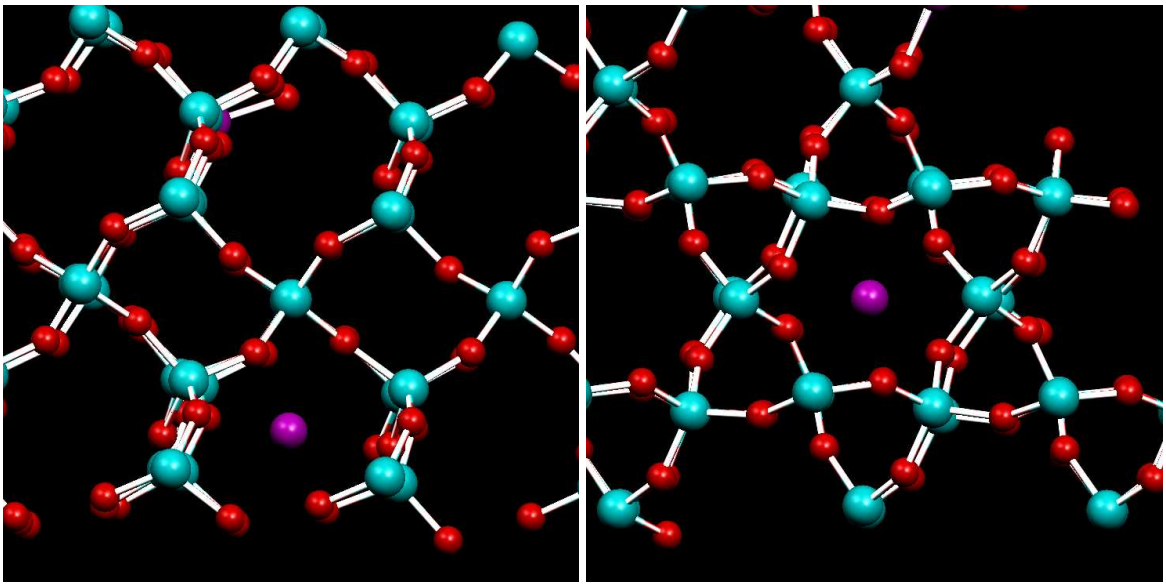


Figure 5.8: M3 (0.175 eV); this is Sartbaeva and Wells' "twofold" site, with the Li^+ ion directly between two oxygens, and is connected to the upper twofold site (M4, Figure 5.10) and the lower twofold site (M2, Figure 5.6) through T4 (Figure 5.9) and T3 respectively.

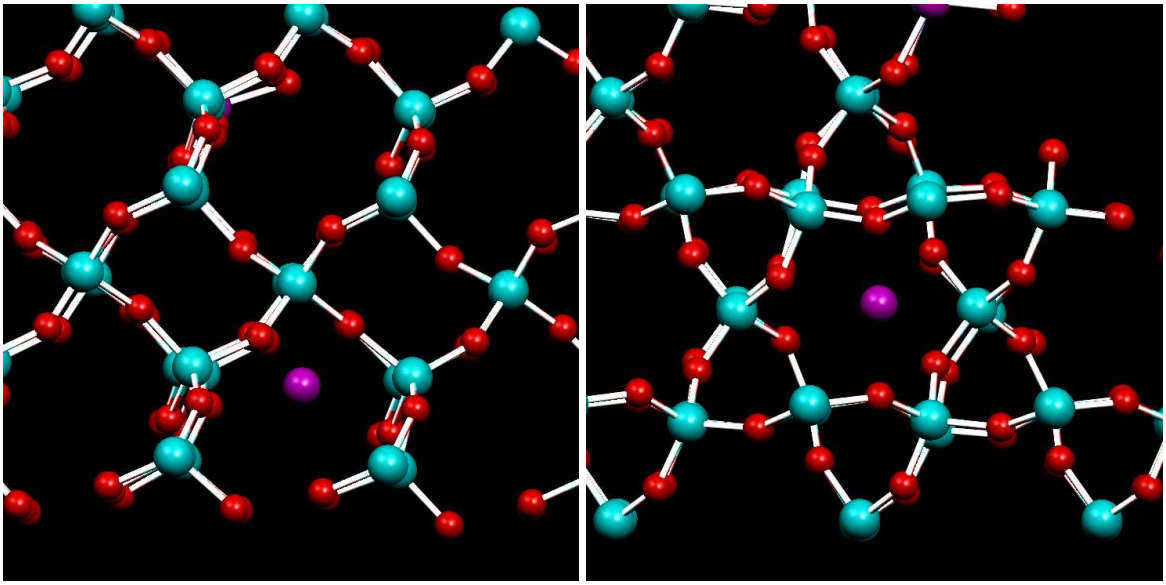


Figure 5.9: T4 (0.211 eV): this is the transition between twofold M3 (Figure 5.8) and and upper fourfold M4 (Figure 5.10).

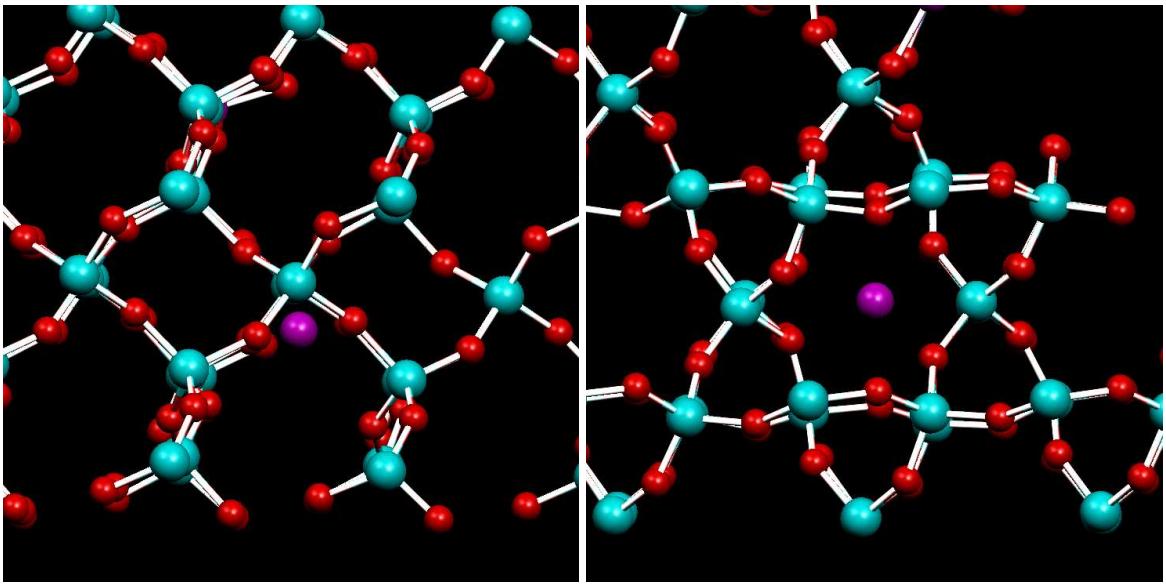


Figure 5.10: M4 (0.053 eV): this is a fourfold site like M2 (Figure 5.6), connected to twofold M3 by T4 (Figure 5.9).

0.028 eV. They are separated by activation barriers of 0.053 eV (on the more stable side) and 0.025 eV (on the less stable one). Therefore, calling these two activated processes P_1 (motion from M2 to M1 via T2) and P_2 (M1 to M2 via T2), less stable side (M2 side) first, and their corresponding reaction rate constants k_{P_1} and k_{P_2};

$$\begin{aligned} k_{P_1} &= A_{P_1} e^{-0.025/kT} \\ k_{P_2} &= A_{P_2} e^{-0.053/kT} \\ \Rightarrow K_{P_1,P_2} &= \frac{A_{P_1} e^{(0.025-0.053)/kT}}{A_{P_2}} \\ \Rightarrow K_{P_1,P_2} &= \frac{A_{P_1}}{A_{P_2}} e^{-0.028/kT} \end{aligned}$$

If one then assumes that $A_{P_1} \approx A_{P_2}$, and therefore that their ratio is approximately 1, one can plot the behaviour of the equilibrium constant – and therefore, the ratio of the populations of M2 and M1 for this process against temperature (Figure 5.11). This is, of course, just the Maxwell-Boltzman distribution.

This, of course, neglects the effect that changes in the Si–O lattice may have on the relative activation energy of these two processes; still, it is clear that in lab conditions (say, 300 K), the ratio Occupancy(M2)/Occupancy(M1) is around 0.3 – i.e, the occupancy of M1 is around three times higher than that of M2, but both will be clearly occupied in, for instance, diffaction patterns.

Therefore, this suggests that at room temperature there is substantial occupancy of both these sites. One would expect to find evidence of this, and indeed, in the PhD thesis of Asel Sartbaeva she found substantial evidence for disorder in the position of the fourfold site via X-ray diffraction, and I believe that this provides a mechanism by which that disorder can occur.

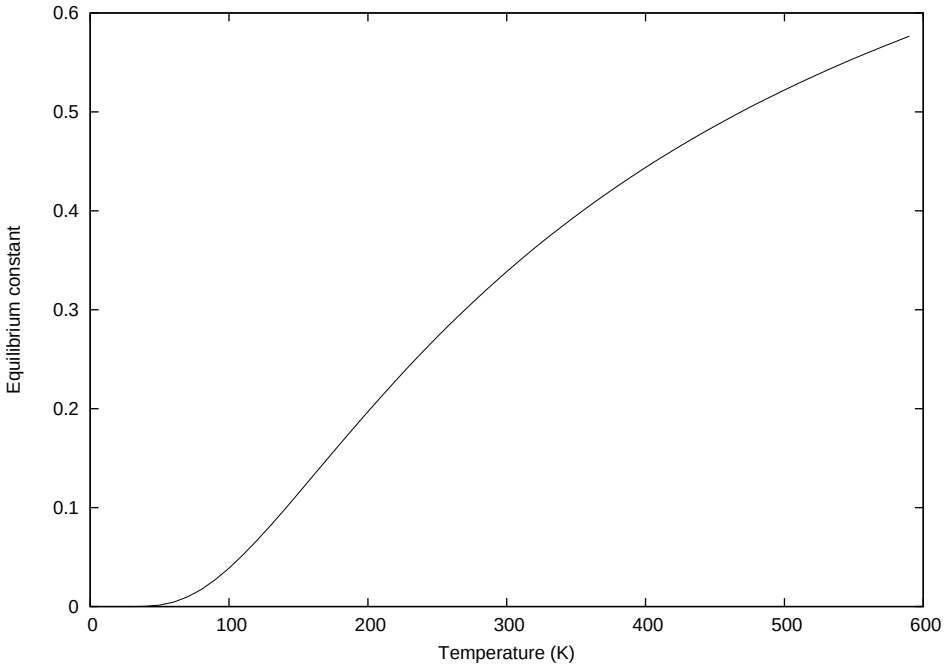


Figure 5.11: Approximate equilibrium constant between the two most stable Li^+–Si sites in α-quartz

5.3.4 Variations in channel width throughout the diffusion process

In analysing the diffusion mechanism in this structure, it has been suggested [6, 17] that the great flexibility of the quartz structure – arising from rigid unit modes [89, 90, 80] – is key. These low-energy vibrational modes control the width of channels in the structure, what might be termed its "openness", without requiring polyhedral distortion; in the limiting model of rigid tetrahedra with force-free joints at the corner, these whole-structure vibrational modes have zero frequency.

The central question is one of a chicken-and-egg nature, to an extent; what role does channel distortion play in the motion of Li^+ atoms up the [001] axis? One can consider this from two directions; either flexion of the channel permits diffusion, or motion of the diffusing species *causes* the distortion, with the barrier height being controlled by the energetic cost of the induced strain.

Therefore, firstly, it is necessary to characterise the structural distortions occurring during motion of Li^+. This can be achieved by appropriate measurements of Si–Si and O–O distances, as metrics for the channel width of the structure at each of the critical points (minima and transition states) throughout the structure.

I consider these in three channels: the primary channel, containing the Li^+ ion; the secondary/adjacent channel, next to the primary channel; and the tertiary channel, which is two channels away from the primary channel. The pairs of atoms considered can be seen in Figures 5.12 and 5.14; the variations of the distance between them are in Figures 5.13 and 5.15. The states are numbered in these graphs in terms of the names assigned in the captions to Figures 5.2-5.10; states 2 (M2, Figure 5.6), 4 (M3, Figure 5.8), and 6 (M4, Figure 5.10) are stable states, with 1, 3, and 5 being the transition states (T5, T3 and T4 respectively, as seen in Figures 5.5, 5.7 and 5.9) linking them.

It is worth noting at this juncture that the structure is, neglecting Al^{3+}, approx-

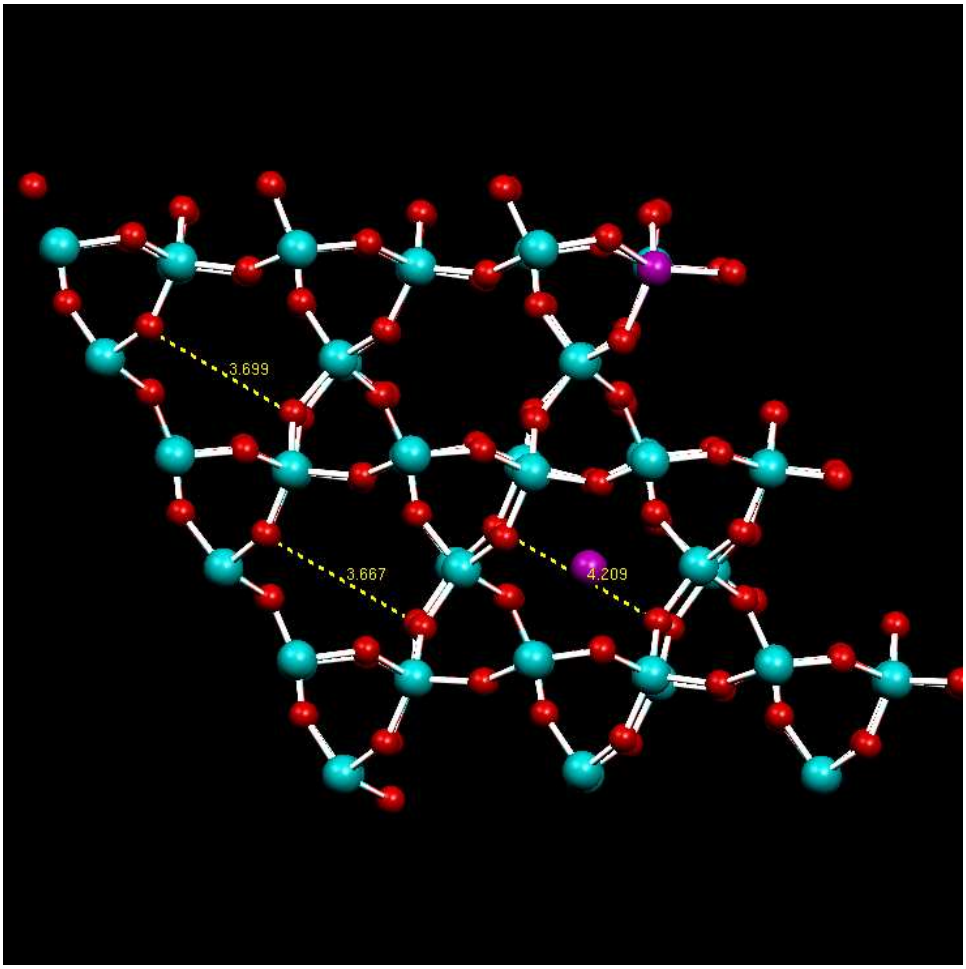


Figure 5.12: Three equivalent O–O pairs in three channels of α-quartz - primary (containing Li), adjacent, and tertiary (two channels removed from Li). The two oxygens selected are the primary oxygens in the "twofold" coordination site described above.

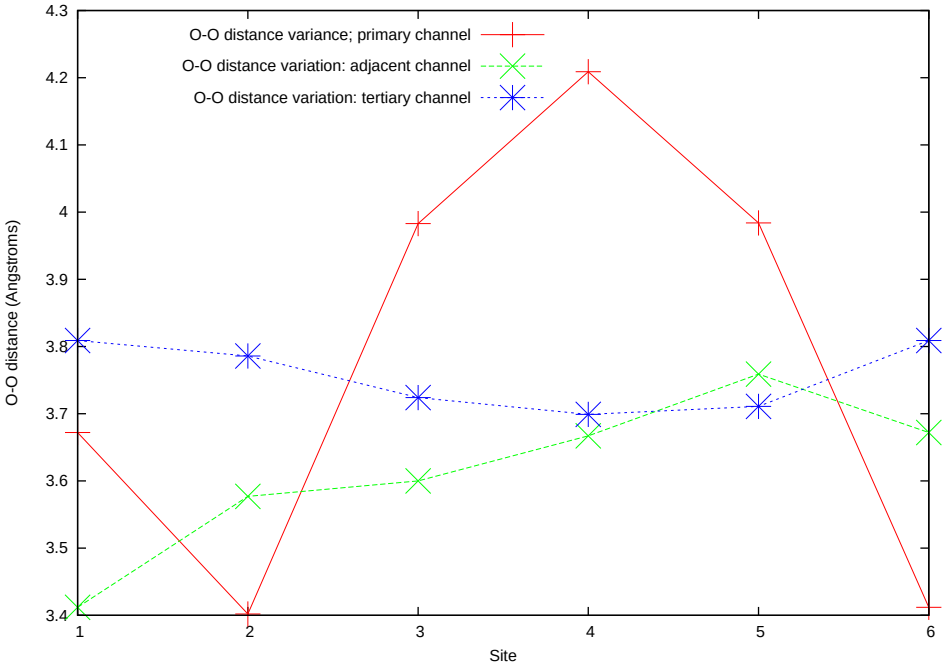


Figure 5.13: Distance between three equivalent O–O pairs in three channels of α-quartz, for each of the six sites in our diffusion mechanism above; in order, sites T5, M2, T3, M3, T4 and M4.

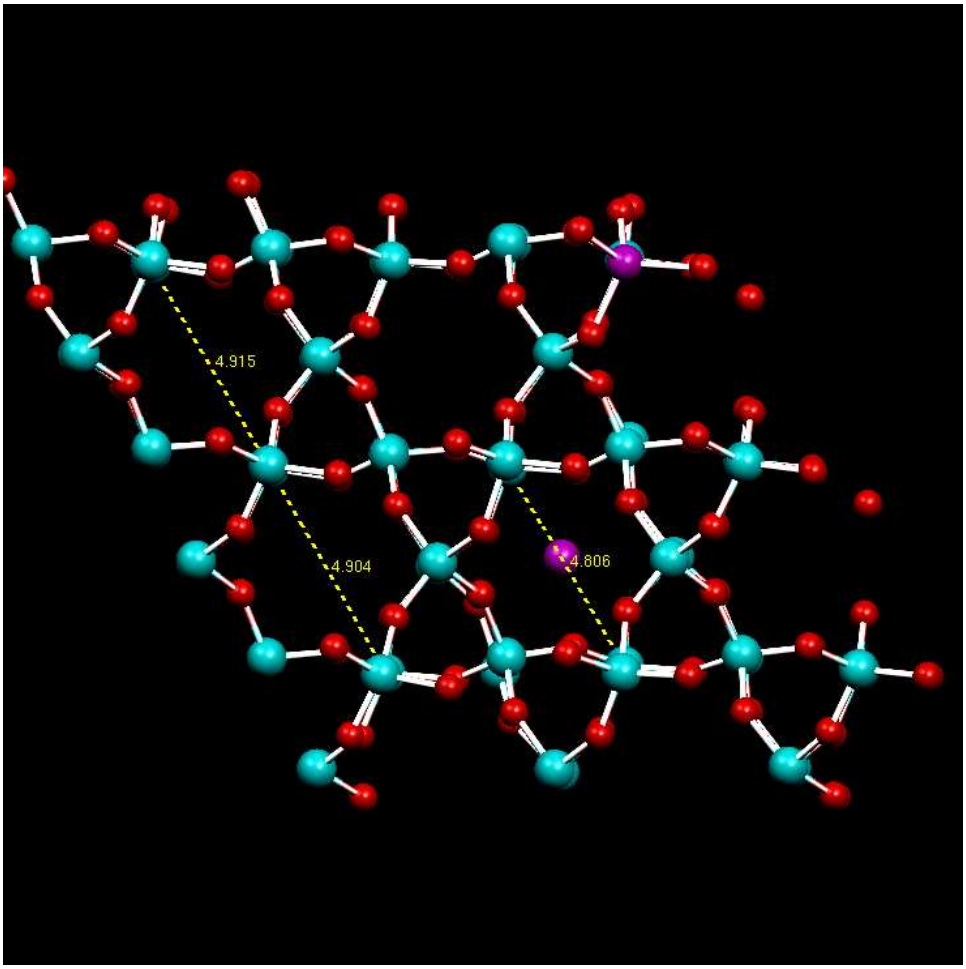


Figure 5.14: Three equivalent Si–Si pairs in three channels of α-quartz - primary (containing Li^+), adjacent, and tertiary (two channels removed from Li^+).

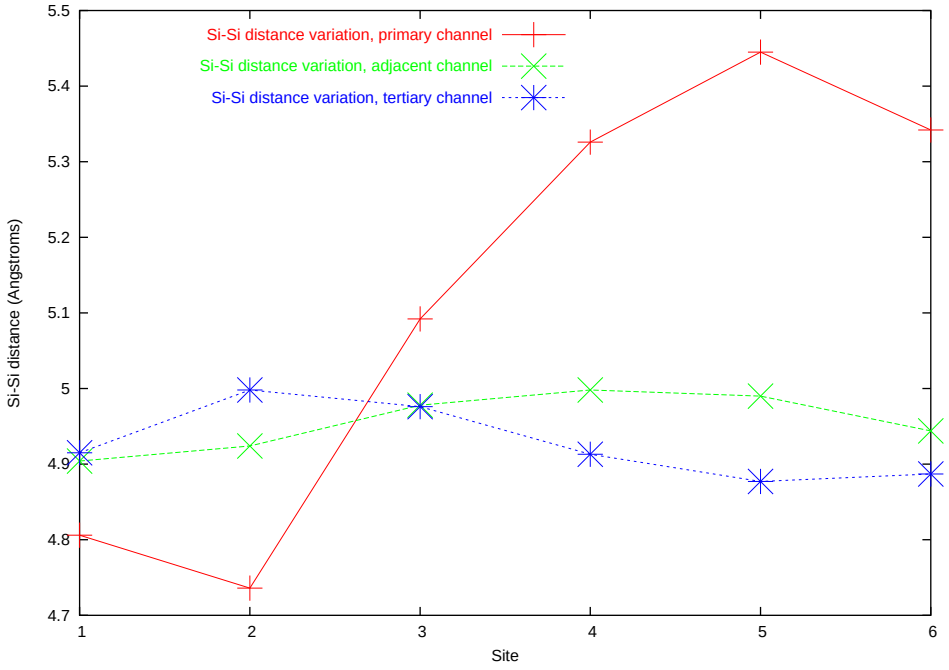


Figure 5.15: Distance between three equivalent Si–Si pairs in three channels of α-quartz, for each of the six sites in our diffusion mechanism above; in order, sites T5, M2, T3, M3, T4 and M4.

imately twofold rotationally symmetric around an axis perpendicular to [001] and passing through the Li^+ position in site M3 (Figure 5.8); therefore, transition states T4 and T3, and stable states M2 and M4, are similar both in geometry and in energy. However, they are not identical; the symmetry of the syestem is broken by the existence of the Al atom. This is apparent in Figures 5.12 and 5.14; in each of these, the interatomic distances are approximately symmetric around site M3.

In all of these figures, variations in channel width are much less marked in non-primary than primary channels, and fall in magnitude the further the channel is from the primary channel; in the case of O–O distance, the channel varies by approximately 0.8 Å in the primary channel, which is around 25% of the mean channel width, but by only 0.1 Å in the tertiary channel. This is much smaller than the observed variation between pairs in the same primary channel, even when they are some distance from the diffusing species; therefore, the distortion induced by the Li^+ atom is strongly directional, favouring [001]. This may help explain the variation in activation energy – it falls, dramatically, to the point where these systems become fast-ion conductors – noted in heavily-doped derivative systems such as β-eucryptite.

One point immediately apparent from Figure 5.12 is that the O–O distance in the primary channel is alternately maximised or minimized for the three transition states observed; observation of the structure for state M3 in Figure 5.8 finds that the distance is maximised when the Li^+ ion lies linearly in between the two oxygen species. This is held for all of the three stable minima identified in the structure, as is shown in figure 5.16. Similarly, one should consider figures 5.17 and 5.18; these demonstrate that the converse relationship holds for Si–Li–Si interactions, in that these are linear at transition states and away from linear otherwise.

This is not in itself so surprising, given that one would anticipate that any Si–Li or O–Li interactions would largely be electrostatic in nature; Si and Li repel each other, whereas Li and O attract. What it does reveal, however, is quite subtle; given the remarkably similar local environment experienced by Li^+ across all three transition states and all three minima any variation in transition state or stable state energies

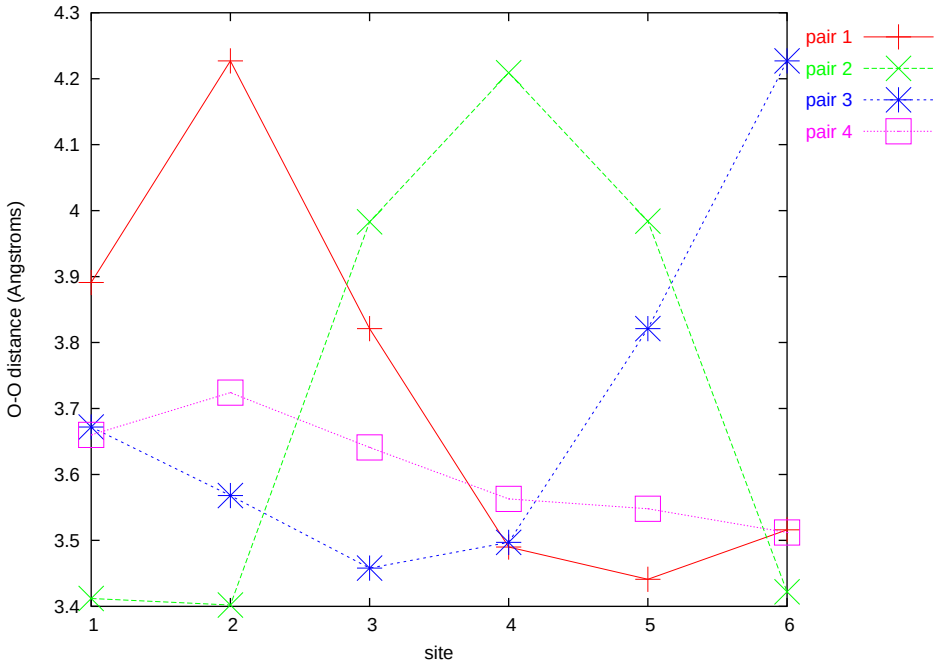


Figure 5.16: Distance between O–O pairs across our six sites in the diffusion process, numbered 1-4, bottom-up (with the first stable state, M2 (Figure 5.6), lying between Pair 1.)

for sites on the main [001] diffusion path is largely, or entirely, dictated by the elastic deformation – and rigid-unit-type relaxation – of the Si–O framework.

This is an interesting, if mainly semantic, reversal of the position taken by Sartbaeva and Wells [17]; in their work, they proceed from the assumption that thermally-induced vibrations in the Si–O framework, which will be large in amplitude given the presence of low-frequency rigid unit modes, permitting enhanced motion of the Li^+ ion along [001] channels when the channel width is favourable. Our results are not so much in disagreement with this as lending themselves to the same interpretation backwards; diffusion is easy along [001] because the motion of Li^+ can *cause* extremely large – as one has seen, of the order of 25% – variations in channel width without excess elastic strain and hence energy cost. The flexibility of the structure along [100] and [001] directions due to the influence of the rigid-unit modes permits it.

5.3.5 Li–Al interactions

One major factor, however, is still to be considered – the Li–Al interaction. Given that our observed reaction path breaks substantially from the symmetry one would expect if this interaction were negligible, as can be seen by the asymmetry in the O–O and Si–Si distances, it is necessary to test our initial hypothesis that Al^{3+} and Li^+ were sufficiently separated within our supercell so as not to strongly interact.

One way of achieving this is by taking our set of transition states, replacing Al with Si, and delocalizing the dopant charge previously concentrated on the Al – one electron – over all the Si atoms in the system equally (thus setting z(Si) $\approx$ 2.3815, while keeping the remainder of our forcefield the same; the remainder of the parameters are in Table 5.1). Having done that, one can then reoptimise (using RFO) the transition states and associated minima as above, under the assumption that the transition state configurations in this new, modified forcefield will be close to the transition states for the original forcefield. This turns out to be the case.

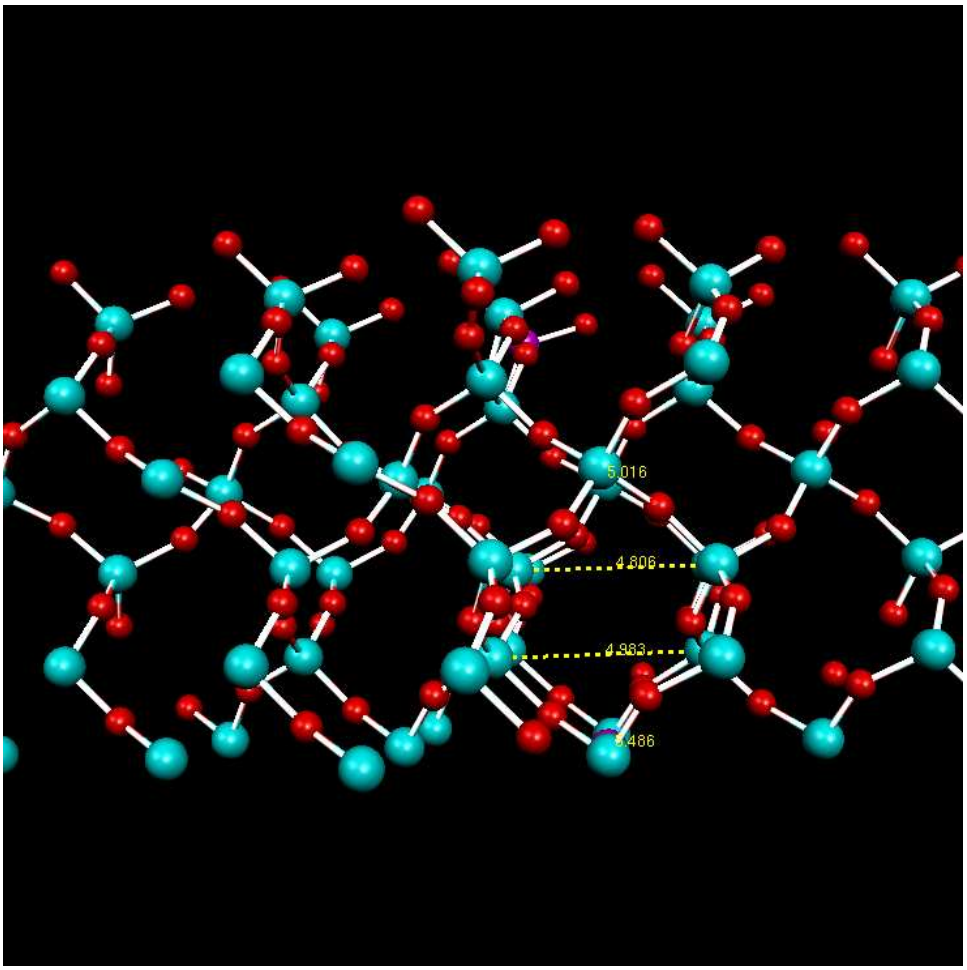


Figure 5.17: The four Si–Si atom pairs relevant to our diffusion process; these are numbered 1–4 from the bottom upwards.

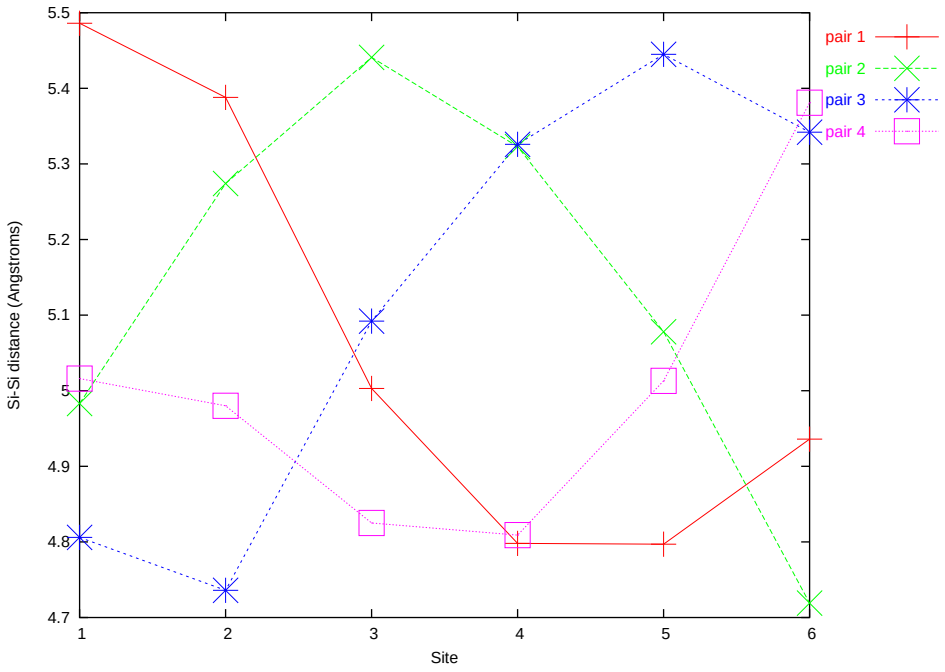


Figure 5.18: Distance between Si–Si pairs across our six sites in the diffusion process, numbered as in Figure 5.17.

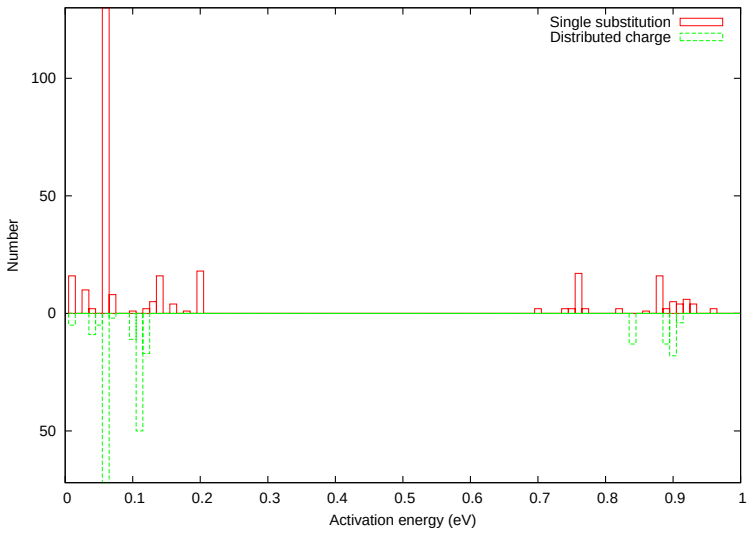


Figure 5.19: Histograms of number-of-states versus activation energy for our sample quartz configurations, with and without defect charge delocalization.

A histogram of transition state energies can therefore be plotted, as can (after identifying two states with the same local configuration) a plot of stable-state energy versus activation energy. These appear in figures 5.19 and 5.20; looking at these, it is firstly apparent that while the overall picture is not radically changed by the inclusion of Al^{3+}, there are variations in the fine detail of the reaction mechanism (Figure 5.20 making this clearest.)

The only way to identify this fine detail is to analyse the transition states found as above. Therefore, I replot the $c-$axis diffusion network first shown in Figure 5.1, using the same labels for transition states and minima as before (and therefore, up to optimization, the same structures as in Figures 5.2–5.10). This can be seen in Figure 5.21. In short, there are two major results. Firstly, the energetic order of the fourfold and wall sites is reversed; the twofold site remains the least stable, with the fourfold site becoming most favourable (although the effect is small). Secondly, the activation barrier for [001] diffusion is substantially lowered, falling to 0.09 eV (approximately half the prior value, and close to Calleja's values for Na^+ diffusion [26]); a large part of the diffusion barrier must therefore be unambiguously due to the

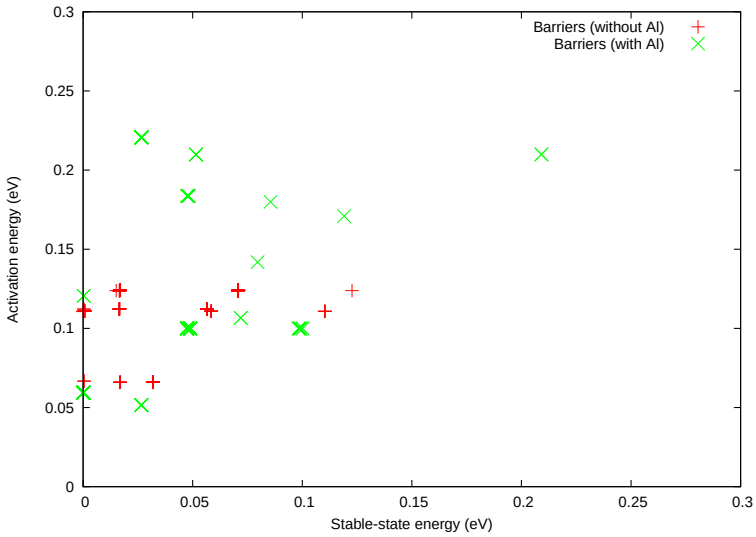


Figure 5.20: Absolute energy versus activation energy for our sample quartz configurations, with and without defect charge delocalization.

effect of Al^{3+}.

The key question is "can the alteration in transition state energies between our Li^+ configurations be explained adequately purely by the change in electrostatic environment caused by substitution of Al for Si, even if Al^{3+} and Li^+ are well-separated?" To answer this, one can calculate the Li–Al distances – allowing for periodic boundary conditions – in each of our stable configurations, and then plot energy against distance. If there is a clear electrostatic effect, this graph should be approximately linear; any deviations from linearity should be of the order of diffusion barriers in our delocalised-Al-defect system, and hence small for sites involved in [001]-diffusion.

Firstly, postulate that the differences between Li^+–Al and Li^+–Si interactions are primarily Coulombic in character. Therefore, one can fit $f(x) = a + c/x$ to our set of stable configurations, and use this to set a baseline for both our stable and transition states. The result can be seen in Figure 5.22; the asymptotic standard errors in a and c are both under 3.5%.

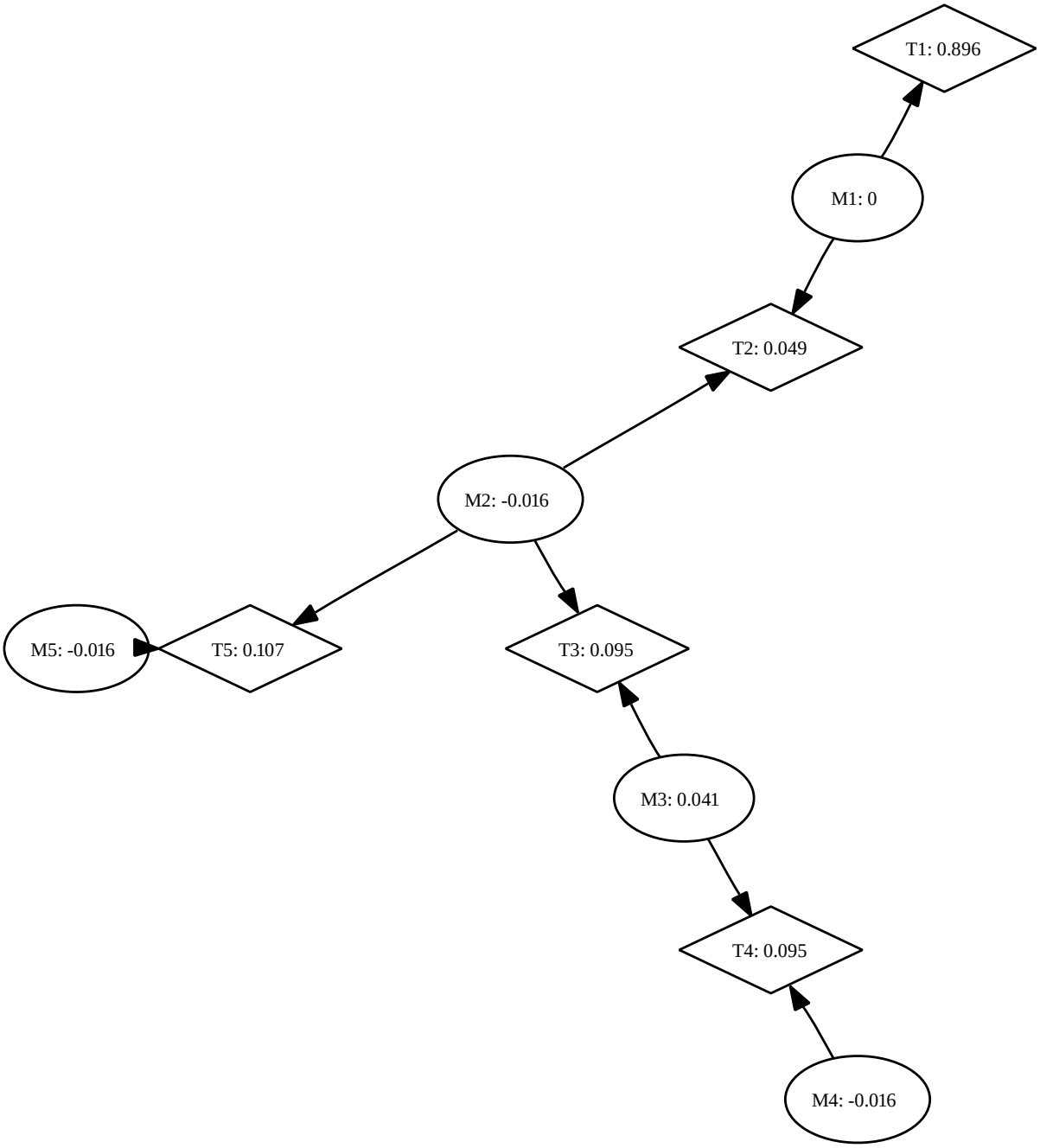


Figure 5.21: Connectivity of minima and transition states found in singly Li^+-doped, Al^{3+}-defect-delocalised α-quartz; diffusion up the c-axis is shown in the labelled vertices M1–M5 and T1–T5. Ellipses represent stable states, diamonds transition states; all energies are presented in eV.

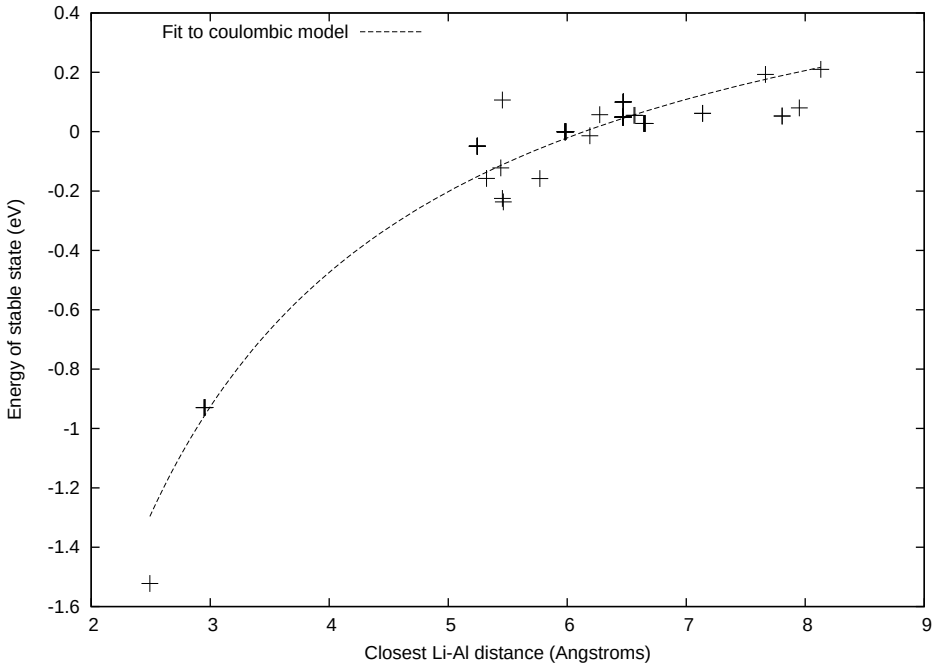


Figure 5.22: Absolute energy versus Li–Al distance; the curve is a functions of the form $a + c/x$. Our fitted value for a is 0.886 eV; for c, -5.44 eV/Å.

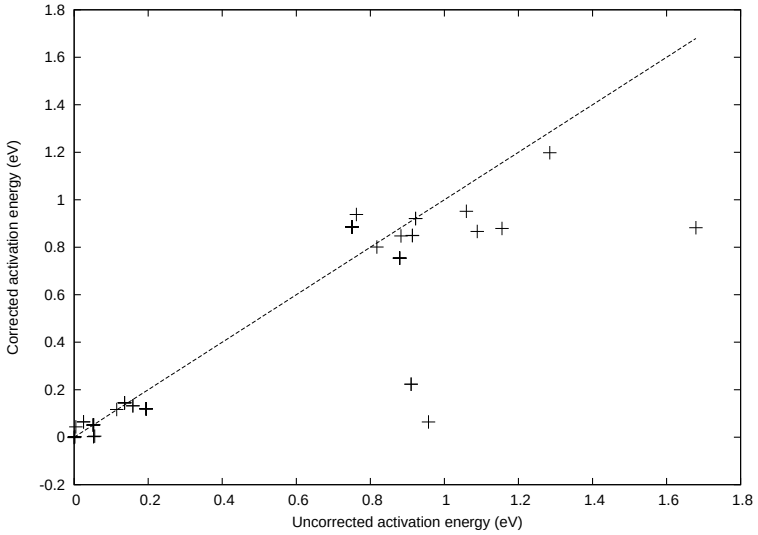


Figure 5.23: Electrostatic-energy corrected activation energy versus uncorrected transition energy; note the lack of deviation from x=1, showing that (apart from three outliers) electrostatic Al–Li interactions have a negligible effect on the activation energies observed.

Having set this baseline, the activation energy for each event which has been identified is calculated. Plotting the corrected transition energies against the uncorrected values, no significant deviation (at low activation energy) from linear behaviour is found. (figure 5.23.)

After this, one can plot the magnitude of the corrections imposed by our model against the (stable) Li–Al distance in Figure 5.24. The same three outliers seen in Figure 5.23 are seen here; they are three of the four closest Li–Al pairings, and furthermore suggests that the dissociation energy for an Li–Al pair will be of the order of 1 eV; this accords well with Sartbaeva's [6] experimental results. This leads to the conclusion that Al affects the dynamics of our system in two ways; close interactions, where Li–Al defect pairs form, and more distant situations – say, Li–Al interactions greater than approximately 5Å – where distortion of the Si–O network caused by substitution of Al for Si, rather than direct electrostatic interaction, is the dominant effect in controlling diffusion dynamics.

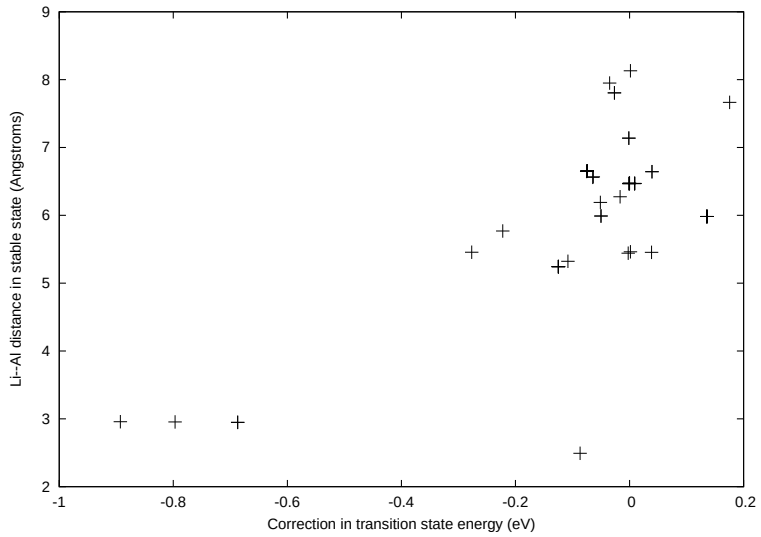


Figure 5.24: Transition-state electrostatic-energy correction (in eV) versus stable-state Li–Al distance (in Å). Note the three outliers.

5.4 Cristobalite and Tridymite

As mentioned earlier, cristobalite and tridymite have similar local geometry; an interesting question is therefore whether they exhibit similar stable and transition states to each other. Searches found three unique barriers in our tridymite cell, and five in cristobalite; undoubtedly, many more than this exist, but my searches in these structures have been vastly less extensive than those performed on quartz. The aim of this section was purely to collect a small, hopefully representative, set of possible events, rather than to carry out an extensive mapping of the structure.

The activation barriers identified range from 0.30 eV to 1.25 eV in tridymite, and from 0.48 eV to 1.19 eV in cristobalite; there is no low-energy diffusion direction of the form seen in quartz, which is to be expected given the absence of channels in these structures. Indeed, this is closer to the behaviour which one might expect to see in glassy structures (which are also unlikely to have channels); that is considered in the next chapter.

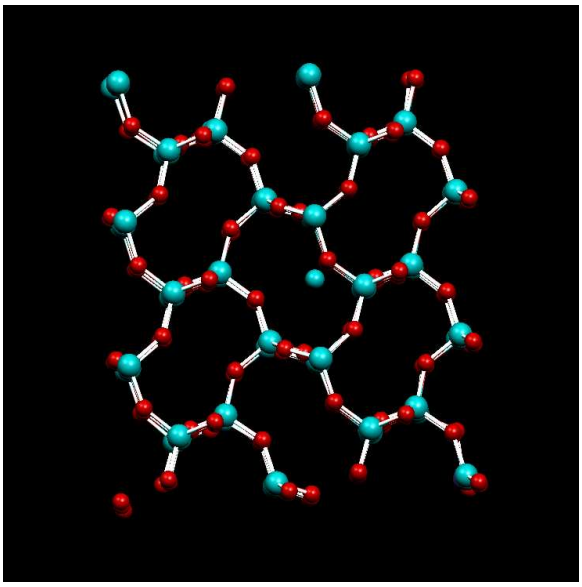


Figure 5.25: CM1 (0 eV): this is the lowest cristobalite minimum found.

We present the events found in the following figures; one point immediately worth drawing attention to is that, considering the stable states alone, the two highest-energy states in cristobalite include a substantial reorientation of an entire chain of polyhedra, right across the supercell; this comes at a substantial energetic cost, as the two states where this is observed are over 0.8 eV higher in energy than the lowest minimum found. No similar states were found in tridymite, however; there, the two minima we identify are only around 0.2 eV apart.

5.4.1 Cristobalite

Four distinct minima and five distinct transition states were found. These can be seen in Figures 5.25 to 5.33. The large blue atoms are silicon; the smaller blue atom lithium; the red atoms oxygen; and the purple atom aluminium.

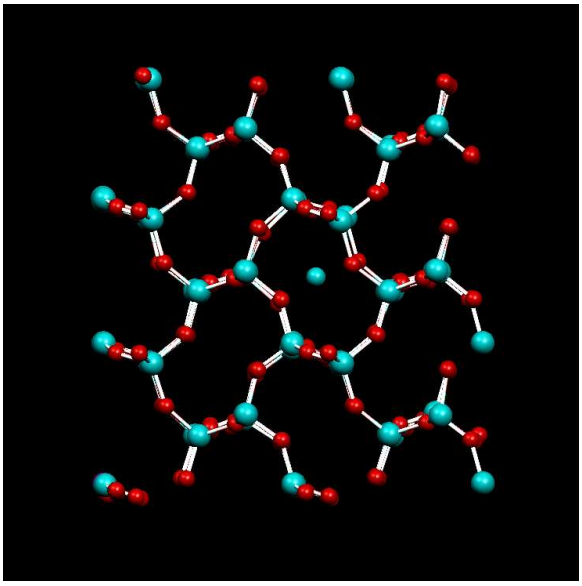


Figure 5.26: CT1 (0.665 eV): this transition state is connected to minimum CM1 in Figure 5.25 with an activation barrier of 0.665 eV.

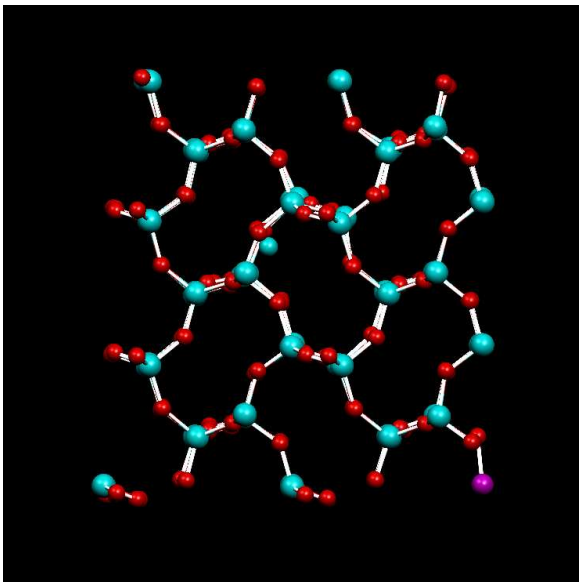


Figure 5.27: CT2 (0.692 eV): this transition state is, like CT1 (Figure 5.26), connected to minimum CM1 (in Figure 5.25). The activation barrier for this process would be 0.692 eV.

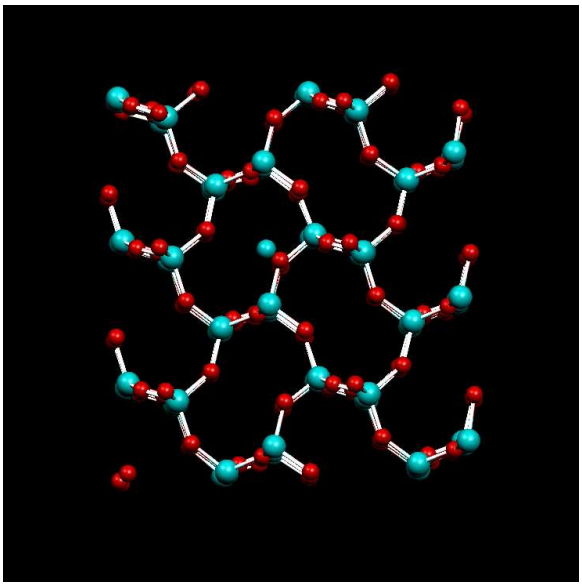


Figure 5.28: CM2 (0.0271 eV): The second-lowest minimum found by my search in cristobalite.

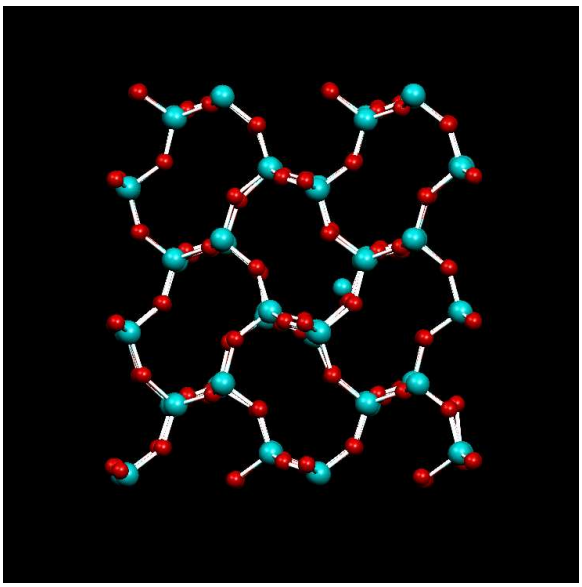


Figure 5.29: CT3 (0.584 eV): This transition state is connected to minimum CM2 in Figure 5.28 by an activation barrier of 0.557 eV.

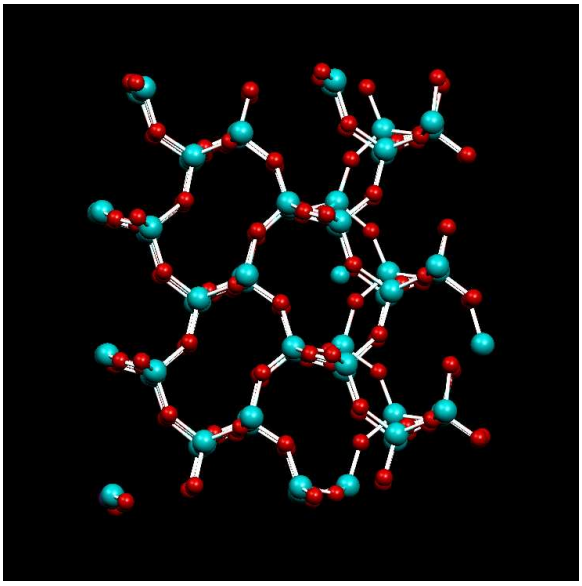


Figure 5.30: CM3 (0.794 eV): Note that in this minimum, and minimum CM4 (Figure 5.32, there is a reoriented chain of tetrahedra.

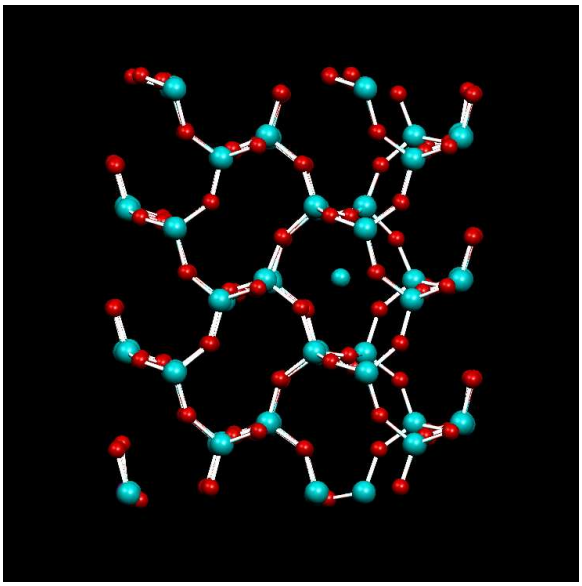


Figure 5.31: CT4 (1.360 eV): transition state connected to minimum CM3 in Figure 5.30. This move is similar to the CM1–CT1 (Figures 5.25 and 5.26) move above, apart from the reoriented chain; the activation energy is also similar, at 0.566 eV.

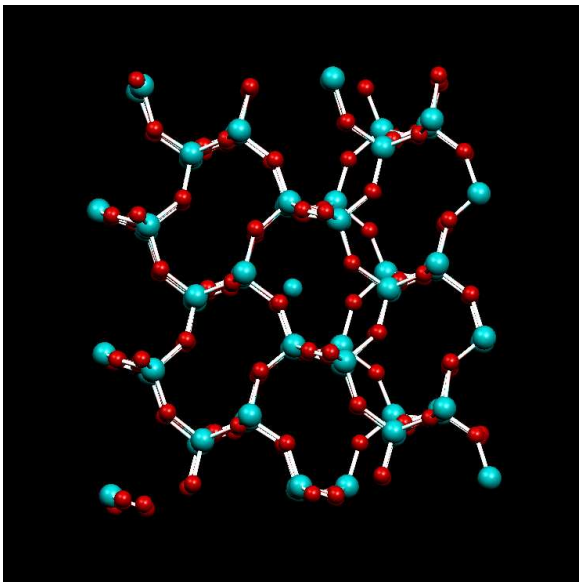


Figure 5.32: CM4 (1.024 eV): as in CM3 (Figure 5.30), a chain of tetrahedra is reoriented with respect to the minimum-energy structure.

5.4.2 Tridymite

In tridymite, the search found two distinct minima and three transition states; these are shown in Figures 5.34 to 5.38. As can be seen, the diffusion processes we have found are predominantly in the xy-plane; some similarity, however, between the moves found in tridymite and cristobalite is parent, purely by visual inspection of the structures shown here. It is clear that Li^+ diffusion is markedly more hindered in these structures than it is in quartz; the steps made are more like those in [110] directions in quartz (Figure 5.3, for example).

5.5 Conclusions

In this chapter, I have investigated Li^+ motion in quartz in great detail, and also made preliminary comparative studies of cristobalite and tridymite. Quartz, as

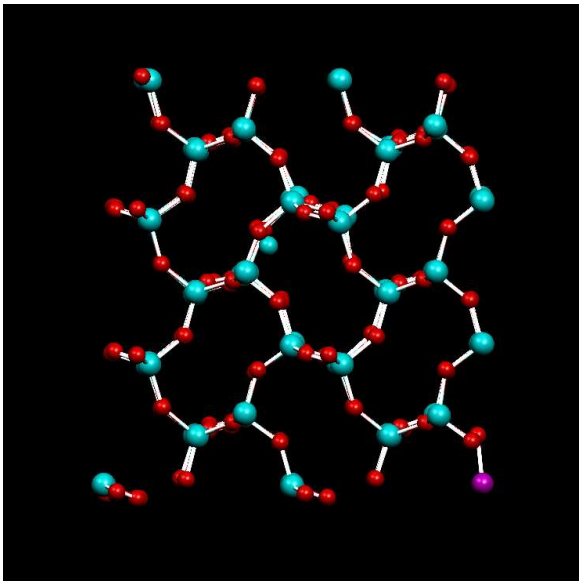


Figure 5.33: CT5 (1.245 eV): transition state connected to minimum CM4 in Figure 5.32. This move, with a low activation energy, is particularly interesting - as it causes the reoriented chain of polyhedra to return to their normal position with an activation barrier of only 0.221 eV.

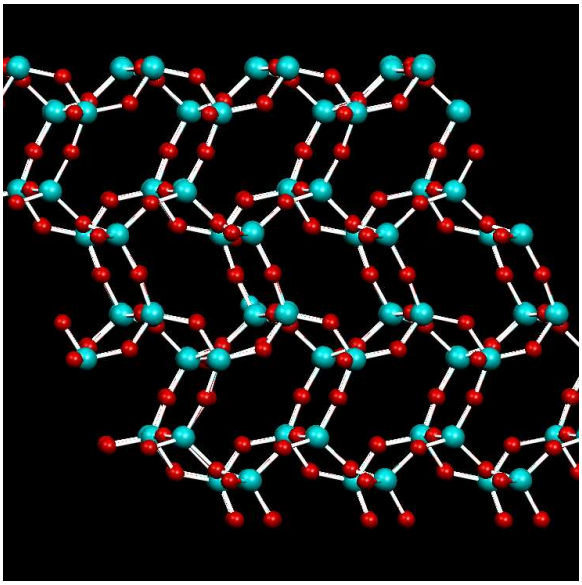


Figure 5.34: TM1 (0 eV): Lowest tridymite minimum found.

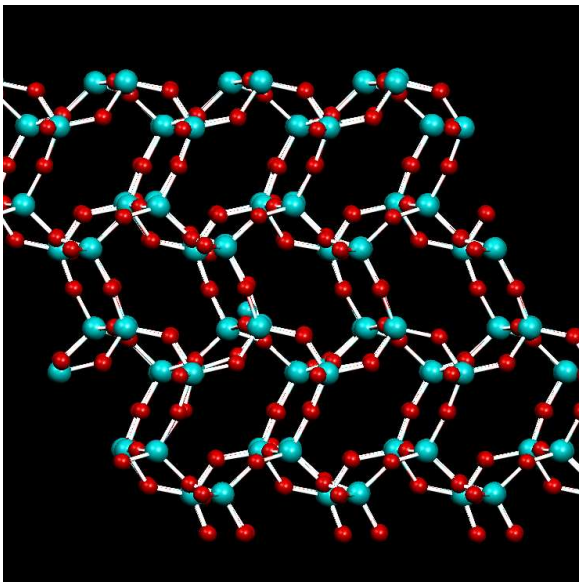


Figure 5.35: TT1 (0.295 eV): transition state connected to TM1 (Figure 5.34), with an activation energy of 0.295 eV. It is difficult to see the lithium ion in this picture; it can be found just to the left of the centre of the image.

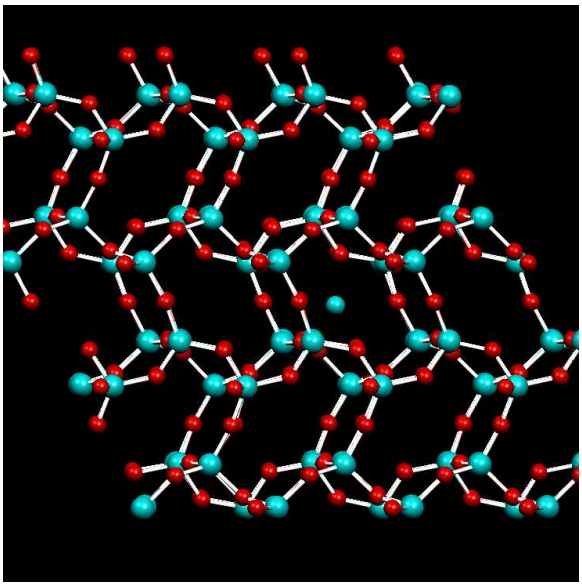


Figure 5.36: TM2 (0.206 eV): Of the two tridymite minima found, the higher in energy.

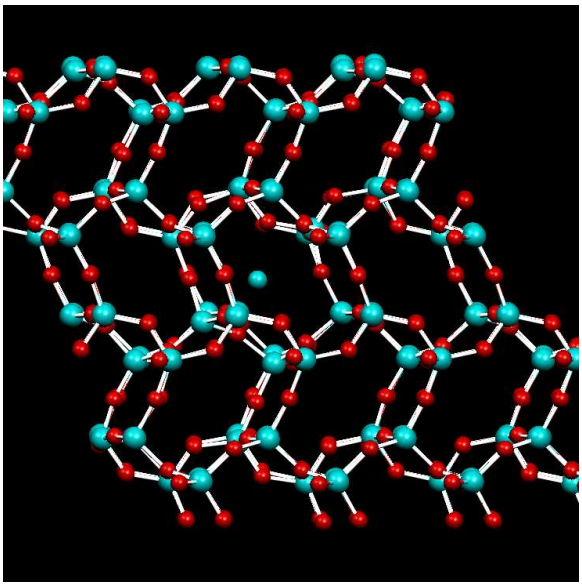


Figure 5.37: TT2 (1.283 eV): Transition state connected to TM2 (Figure 5.36), with an activation energy of 1.077 eV.

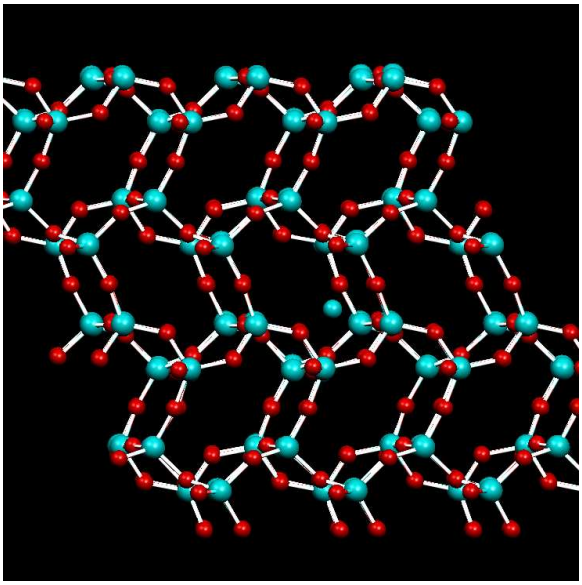


Figure 5.38: TT3 (0.628 eV): like TT2, connected to TM3 (Figure 5.36), but with a much lower activation energy of 0.657 eV.

expected, has a fast direction for diffusion ([001]), due to the channels in its structure, whereas cristobalite and tridymite do not demonstrate this; therefore, the low-energy behaviour of quartz is radically more complex than in the other silica polymorphs considered here.

Methods have also been introduced for identifying extended networks of events from a library of minima and transition states; both the scientific insights gained into diffusion behaviour in silica networks, and these techniques, will be used heavily in the next chapter of this thesis.

Chapter 6

Li^+ diffusion in SiO_2 glass

6.1 Introduction

SiO_2 glass typifies an important class of inorganic amorphous materials. Related systems are, obviously, of exceptional technological importance – whereas, often, a writer would advise his readers to "just look out of the window" to assess the value of some material, in this case one needs only to ask you to look at the window itself!

Simulating extended amorphous systems has only recently become a tractable problem, and in this chapter I outline the problems faced in undertaking simulations of, and further to of that transition-state search in, these kinds of materials. Following this, mechanisms and energy barriers are presented for Li^+ diffusion in silica glass, in the spirit of the simulations of the crystalline phases of SiO_2 in the previous chapter.

In simulating amorphous solids, one faces two major problems. Firstly, one must consider size effects – conventionally, we simulate under periodic boundary conditions, and if our system is not large enough, it will self-interact giving rise to spurious long-range order. Secondly, the periodic boundary conditions are themselves a problem; this constrains the positions of atoms at the edges of the unit cell heavily. This

is not *in itself*, from the perspective of local order, a problem – once the unit cell is constructed, you can consider any point in it to be on the edge – but it does make *constructing* the sample simulation cell difficult.

6.2 Method

The trial glass systems used here were generated by Dr. Kostya Trachenko [91]. The method used is, in essence, fairly simple. Bulk Si has the same network coordination as the glasses we wish to study; thus, an amorphous sample of Si is generated using the method of Wooten, Winer and Weaire [92]. This structure is then, geometrically, expanded, and an oxygen atom placed between each pair of Si atoms; this creates a fully-coordinated, if slightly aphysical, silica network under periodic boundary conditions, which can then be equilibrated by molecular dynamics and/or energy minimization under variable-cell conditions. This results in the creation of a Zachariasen–type [93] continuous random $Q(4)$ SiO_2 network.

In construction of the glass supercell, and his research on this material, Dr. Trachenko used Tsuneyuki's [94] potentials; for consistency with the last chapter I, instead, use Calleja's potentials [26], which are derived from the potentials designed for amorphous SiO_2 by van Beest et al. [27]. Therefore, although this results in a change in energy, I believe it has no significant effect on either the topology or the behaviour of the glass under simulation. I used the 648-atom supercells generated by Dr. Trachenko; the results of the previous chapter indicated that this would be a good balance between performance and accuracy.

This glass was then modified by addition of a Li^+ ion and replacement of one silicon atom with aluminium; the resulting trial glass is, of course, not minimized in energy - particularly as the present study uses a different potential. One needs to freeze the system into a local minimum before performing simulations, and this was achieved by BFGS energy minimization. This had no effect on the topology of the structure.

Following this process, one can be confident that there is a Li^+ dopant in the system, and that it is in a physically reasonable position, but not that it is necessarily in a reasonably likely one; therefore, it is necessary to perform some exploratory transition state searches to "pre-condition" the system. One can achieve this by the simple expedient of running transition-state searches to approximate convergence, then minimizing the total energy of the system from the resulting transition states; one can then choose the lowest energy minimum found, and iterate until stability is achieved. This approach to global optimization is not new; it is similar to that proposed by many previous authors, including both Wales [71] and the ART [70] method of Mousseau and collaborators.

In calculating transition states, the Constrained Linear Maximization method was used, which has described extensively throughout this thesis, with any further checking and refinement of transition states by the Rational Function Optimization [75] method within the GULP program. Indeed, as in the previous chapter, simulations were performed using the GULP [21] code as an energy evaluator, and the CLM algorithm implemented in the CAMPOS-ASE [39] architecture.

6.3 Calculated transition states

Following the same protocol as taken in the previous chapter, we obtain a set of 250 pairs of connected transition states and minima; this is a large, complex dataset, and as such is essentially impossible to analyse by hand. One must resort to statistical and computational methods to make headway in this forest of data. Therefore, initially, I take the same graph-theoretic approach to representing the energy surface as was taken in the previous chapter; again, the calculated library of minima and transition states is transformed into a set of vertices (corresponding to said stationary points on the energy surface) and edges (connecting pairs of transition states and minima), and then plotted using an automated graph-drawing algorithm – again, `graphviz` [87] was used.

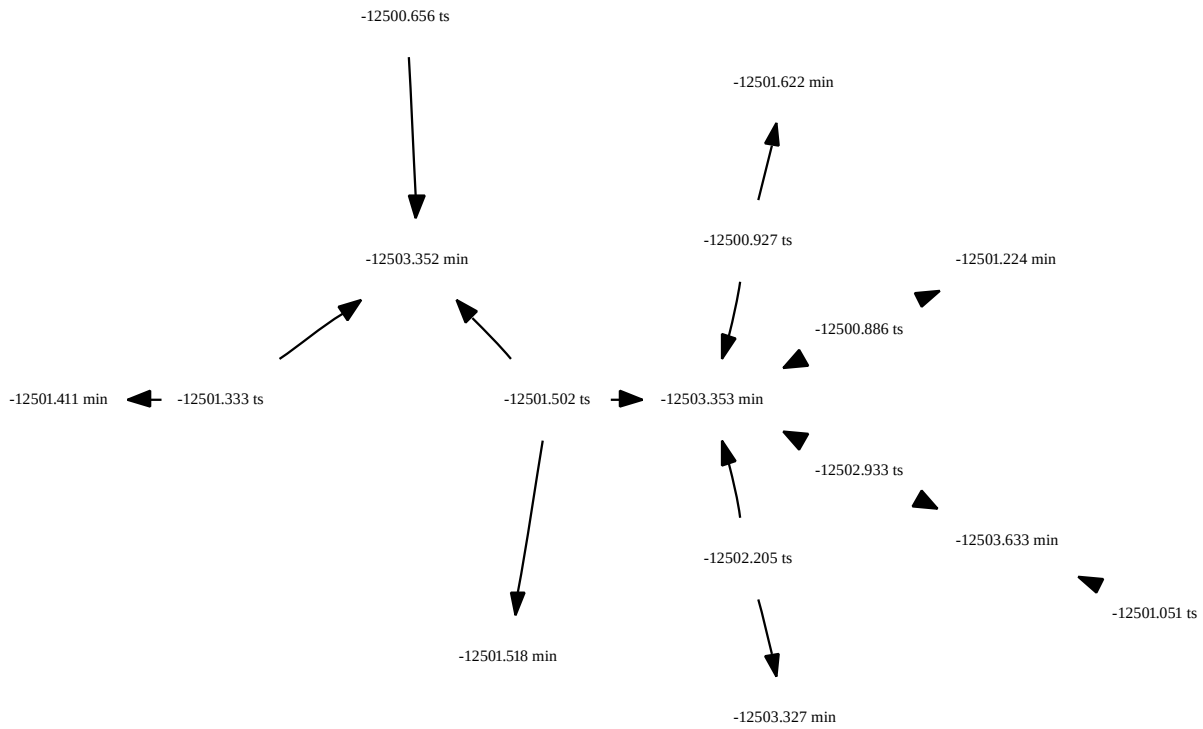


Figure 6.1: Extended event network in glass (i).

This identifies a number of connected subsets of transition states on the graph. Two in particular are of great interest, ranging (as they do) over a fully 4 eV range; extracting extended event networks of this range and scale from the simulation is a result in itself. It is in extended systems like this where the value of this kind of data analysis becomes self-evident; by comparison, α-quartz is a relatively limited problem, given its smaller size and constrained symmetry (and thus lower dimensionality). These are shown in Figures 6.1 and 6.2.

Two major questions present themselves; what can be learned about the number and nature of the stable states in the glass, and about the number and nature of the transition state events therein. In the remainder of this chapter, I present a number of approaches to analysing this, and draw both some comparisons with the

Figure 6.2: Extended event network in glass (ii).

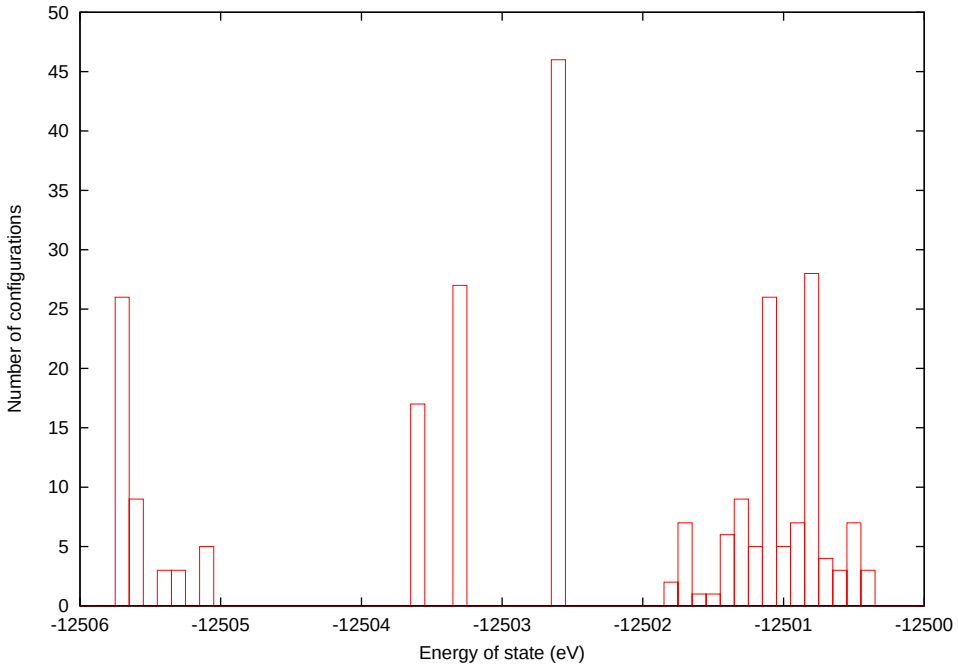


Figure 6.3: Histogram of stable states identified in our Li^+-doped glass sample.

crystalline silica phases and some preliminary conclusions about diffusion behaviour in this system.

6.3.1 Stable states; distribution and comparison with crystalline phases

In Figure 6.3, a histogram (with a bin-width of 0.1 eV) of the energies of the 250 stable configurations calculated is shown. There are two distinct energy bands whith substantial occupancy - approximately -12500 to -12502 eV, and -12505 to -12506 eV, with three states dividing the gap between these two bands at around -12503 eV.

It is also worth plotting the stable states against the activation energies associated with their connected transition states; this is shown in figure 6.4. Looking at this, it

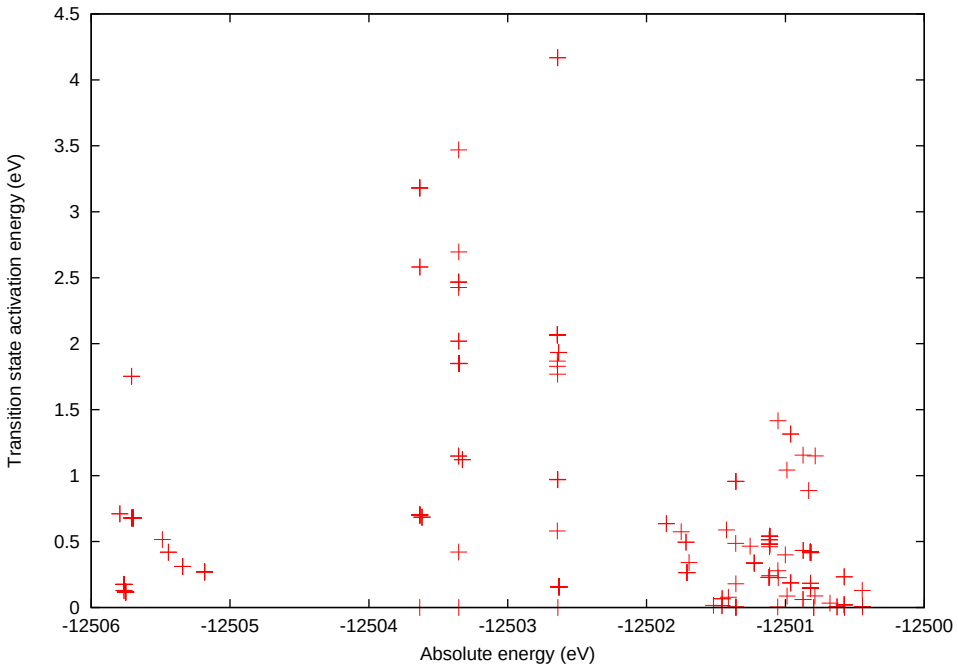


Figure 6.4: Spectrum of stable states and associated activation energies in our glass sample.

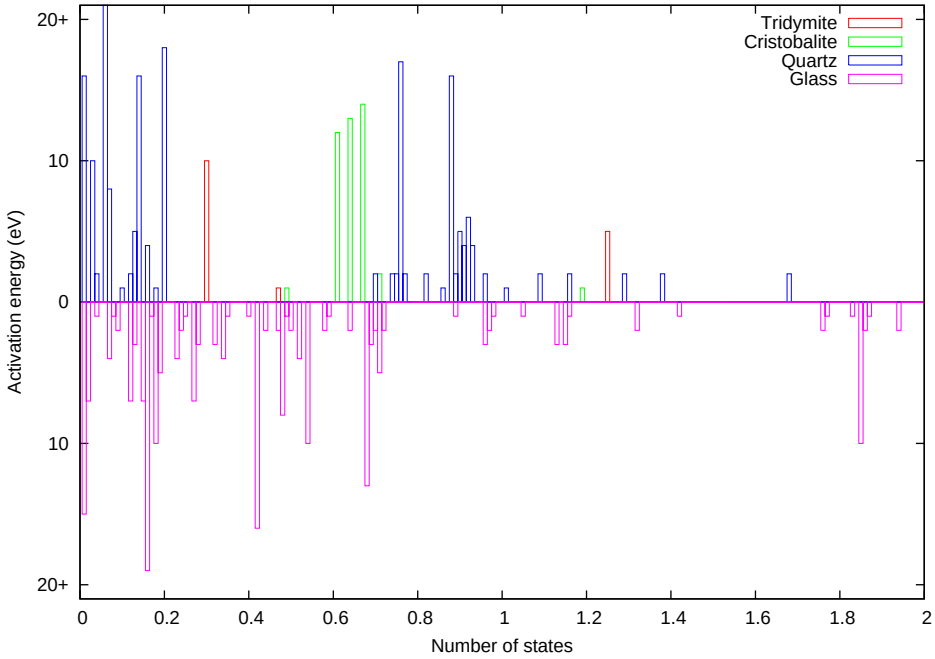


Figure 6.5: Distribution of activation barriers in SiO_2 glass, quartz, tridymite and cristobalite.

is notable that lower-energy transition states exist in the -12500 to -12502 eV band than in the rest of the space, and that the transition state energies around -12503 eV are particularly high; this is a subject which will be returned to later in this chapter.

It is worth comparing the distribution of activation energies in the glass with the three crystalline phases earlier. Some authors – notably Keen et al. [95] – have suggested some similarity between the immediate local structures of silica glass and β-cristobalite; simulations in the last chapter were on the α (low temperature) phase, rather than the β phase, but the spectrum should not be markedly different. It is clear that the energy distribution of transition state events is much broader in the glass than in any of the silica phases; furthermore, the large number of low-energy events in the glass opens the question whether there are regions of glass which are locally quartz-like (open, helical, channel-based) rather than ring-based, as in

cristobalite and tridymite.

To approach this, it is necessary to consider some statistical measures of the local environment around Li, and I turn to this next.

6.3.2 Transition state events; nature, and correlation with stable states

Unlike in crystalline systems, where in general the coordination of one species by another is well-defined, in a glass the system can be so profoundly disordered that it is difficult to speak in terms of coordination spheres. The typical approach to then take is the plotting of radial distribution functions; this is a sensible approach when one has a lot of potential pairs to plot, as would be the case for Si–O in this system, but this is not true with respect to Li.

Furthermore, defining a coordination is complicated by potential discontinuities in many of the measures which could be used - counting all atoms less than a specific cutoff distance, for example, would suffer discontinuities with change of cutoff, whereas counting all atoms of one species (in this case, O) closer to the Li atom than the atoms of another (say, Si) would excessively penalise cases where an Si atom happens to be relatively close to Li.

The measure thus invented is given in Equation 6.1. To explain, it is a sum over all atoms of a given species; the Gaussian term evaluates to 1 when an atom is at 2.1 Å from Li, and falls rapidly after that point. From this, one can get a good idea of the local oxygen (and silicon) environment around our Li dopant. Here, r_x is the position vector of species x.

$$\eta = \sum_{1 \ldots n_{ELEMENT}}^{N} e^{-(2.1-|\mathbf{r_n}-\mathbf{r_{Li}}|)^2} \tag{6.1}$$

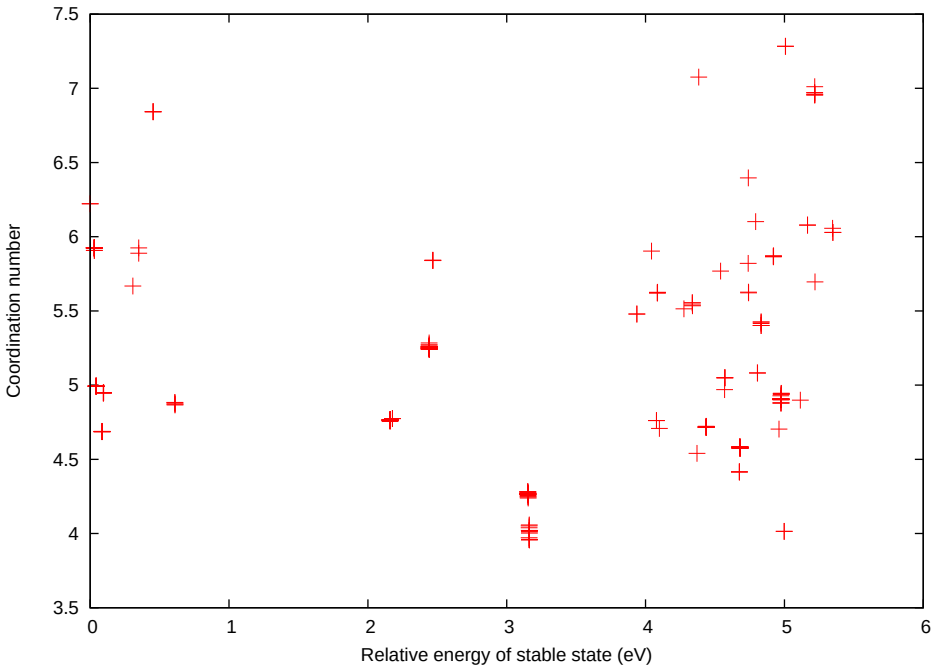


Figure 6.6: Coordination of Li by O (as defined in the text) in each of our stable states.

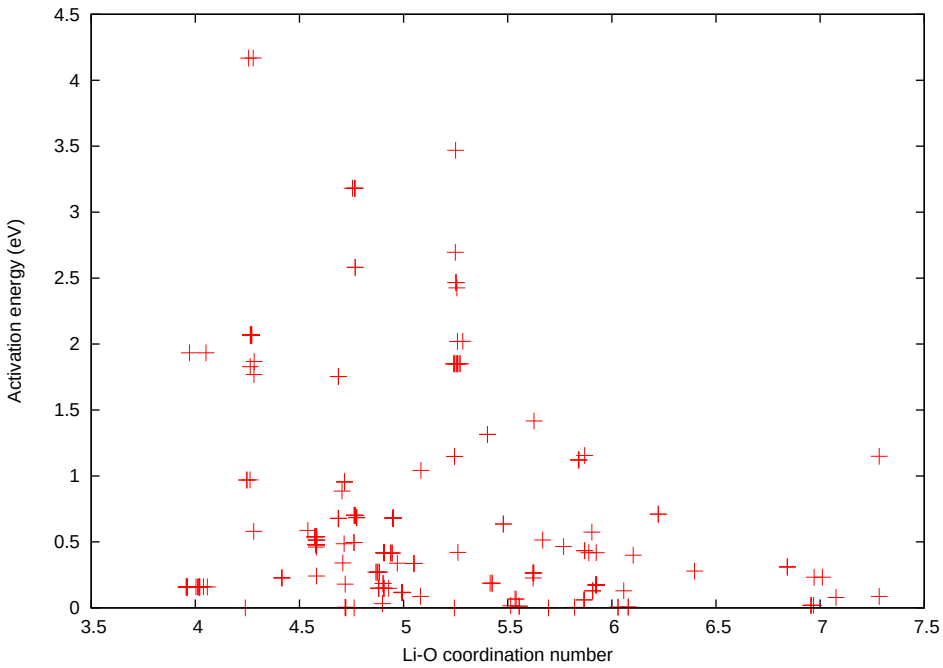


Figure 6.7: Coordination of Li by O in the stable state (as defined in the text) versus activation energy for each process

Using this, the coordination of Li by O is plotted in Figure 6.6; it is notable that the coordination is markedly lower in the -12503 eV states than in the other two bands. This is of some importance, and will be returned to later. Plotting coordination against activation energy, it is also evident that the higher the coordination, the smaller (and lower) the spectrum of energies of transition state events is; this is shown in Figure 6.7.

Plotting the relative Li-O and Li-Si coordination against each other, as in Figure 6.8, we see that these are linearly correlated; this is what we would naively anticipate, given that Si and O positions are strongly correlated by their being bonded to each other, but it is worth checking. There is some scatter, as one would expect; some local sites are slightly closer to Si or O, as in the quartz, and these display marginally enhanced or diminished coordination as appropriate.

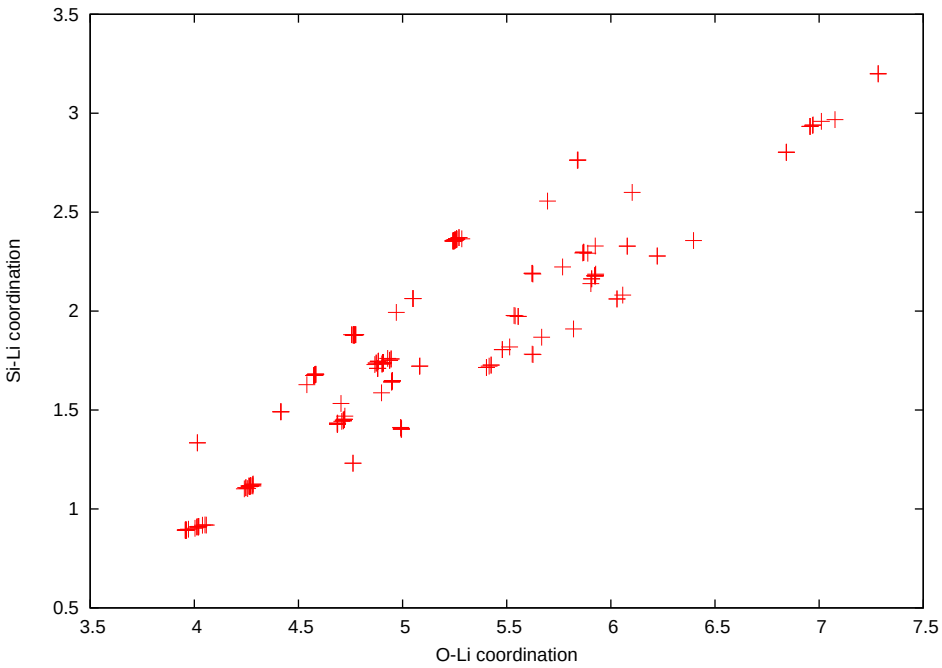


Figure 6.8: Relative Li-O and Li-Si coordination in stable states in our glass

In figure 6.9, we show the mean Li–O distances for the ten closest oxygens to Li in three energy ranges of stable states - around -12501 eV, -12503 eV, and -12505 eV. From this, one can see some of the reasons for the variation in transition-state spectra across the range; Li is markedly closer to its four nearest neighbouring oxygens in the -12503 eV case than the others. The -12505 eV case gains energetically over the -12501 eV case in the longer-range oxygens; the immediate, local environments are broadly similar, but from oxygens 5–10 the lower-energy state consistently displays shorter distances.

The next issue to consider is how coordination changes, relative to the stable state, when the system moves to a transition state; this is plotted in Figure 6.10. Diffusion events below 1 eV have a slight skew towards a decrease in the coordination of Li by O; above 1 eV, the coordination invariably increases. In terms of the above, it can be argued that in the latter case, the system is moving – relatively speaking –from a short-bond, close-and-low-coordination-number configuration (like those around

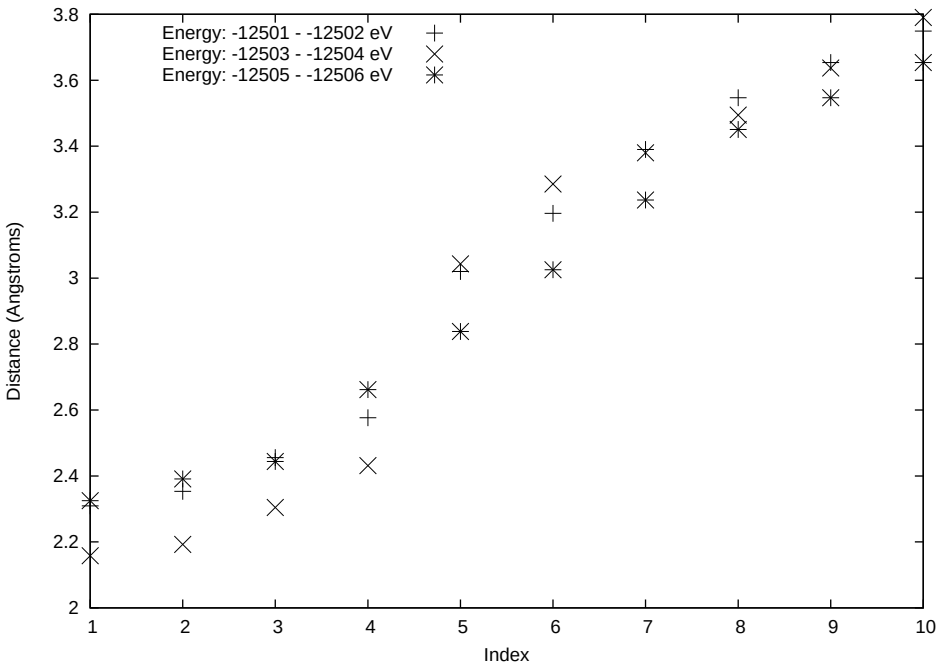


Figure 6.9: Mean Li-O(1) ... Li-O(10) distances for three energy ranges

-12503 eV) to a larger, more diffuse locale (like those at -12501 or -12505 eV). It would be interesting to study this in substantially more detail.

Another measure which can be defined is a local density (Equation 6.2); here, γ controls the size of the environment summed over.

$$\eta = \sum_{1...n_{ELEMENT}}^{N} e^{d_{Li-Element}/\gamma^2} \tag{6.2}$$

Using $\gamma = 4$Å, we can calculate a change in density for a given transition state for each element; plotting the change in density for Si versus that for O for each transition in our sample-set, we obtain Figure 6.11, normalizing the data by halving the value of the O density (in light of there being twice as many O atoms as Si in the system). From this, it can be concluded that oxygen atoms move more than

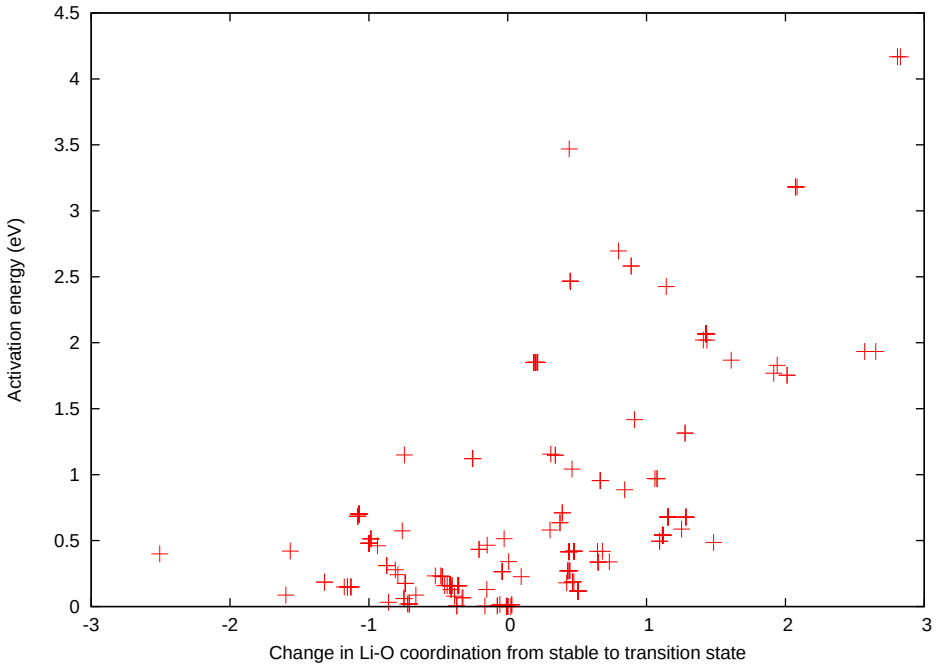


Figure 6.10: Change in coordination of Li by O between the stable and transition state (as defined in the text) versus activation energy, in eV

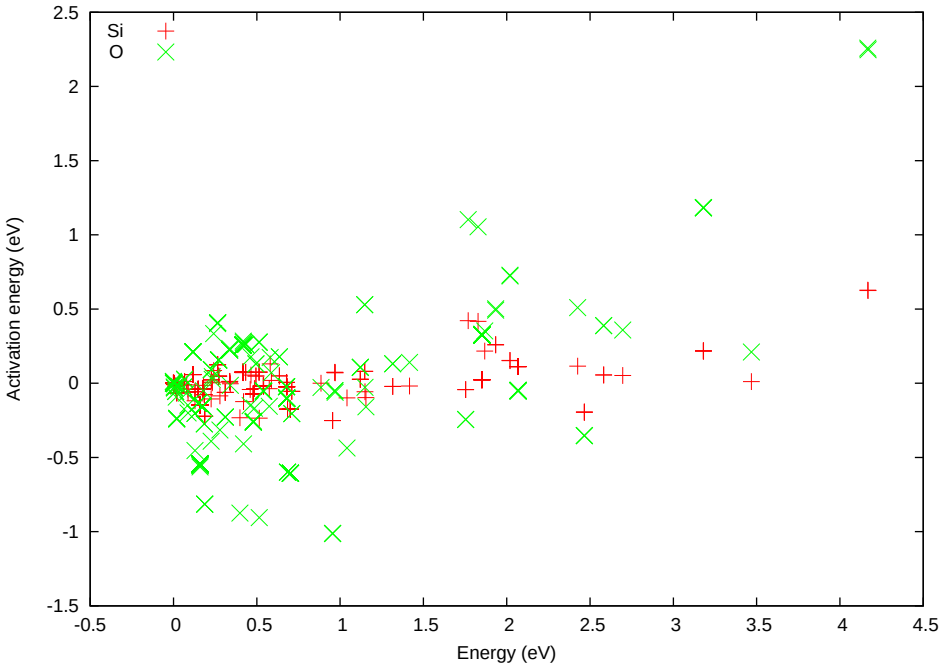


Figure 6.11: Normalized change in density of Si and O around Li between stable and transition states in glass

silicon atoms do in our transition states; this is suggestive of the presence of rigid-unit type relaxation (as mentioned in the previous chapter in quartz). Imagine that the glass is composed of rigid SiO_4 tetrahedra, joined at the corners; many elastic processes can therefore occur by pivoting of the tetrahedra around these corners, moving the oxygen atoms substantially but not necessarily disturbing the relative Si–Si positions. Therefore, there is value in seeking further evidence as to whether this is, in fact, the case.

6.3.3 Role of network distortion in transition-state events

In order to consider the role of network distortion, it is firstly incumbent upon us to work out over what distance the distortion of the Si–O network associated with transition states extends. Consider each structure as a vector of length $3N$,

composed of each atom's coordinates in Cartesian space, and construct three unit vectors in this space corresponding to translation of the entire system along each of x, y and z. I call these p_x, p_y and p_z. Taking our system and a reference system as vectors S and S_R;

$$\begin{aligned}
\delta &= S - S_R \\
\delta_1 &= \delta - p_x(p_x.\delta) \\
\delta_2 &= \delta_1 - p_y(p_y.\delta_1) \\
\delta_3 &= \delta_2 - p_z(p_z.\delta_2) \\
S_1 &= S_R + \delta_3
\end{aligned}$$

Here, S_1 is the closest setting of system S to S_R; that is to say, we have chosen the unit cells for each of our systems for which the vectorial distance between S and S_R is minimized. If one performs this resetting *neglecting* the position of Li, what is obtained is a vector representing the distortion of the Si–O network during a transition event.

Consider figure 6.12; there is clearly some positive correlation – looking at the data, arguably multiple overlaid correlations – between the magnitude of this vector, and hence of the distortion of the Si–O network, and the energy of transition. The nature of this correlation is unclear – to the naked eye, it appears to be supralinear, but one cannot afford to simply trust one's intuition. Therefore, one must investigate further. If one plots the magnitude of Si–O distortion against the energy of the stable site, one obtains Figure 6.13. Considering this figure, it is interesting to note that in the more stable structures, a substantial distortion of the Si–O network – at least 1Å – is required to initiate a transition process. This suggests that the enhanced stability of the more stable sites comes from permitted elastic relaxation of the Si–O network, not possible at the higher-energy metastable sites.

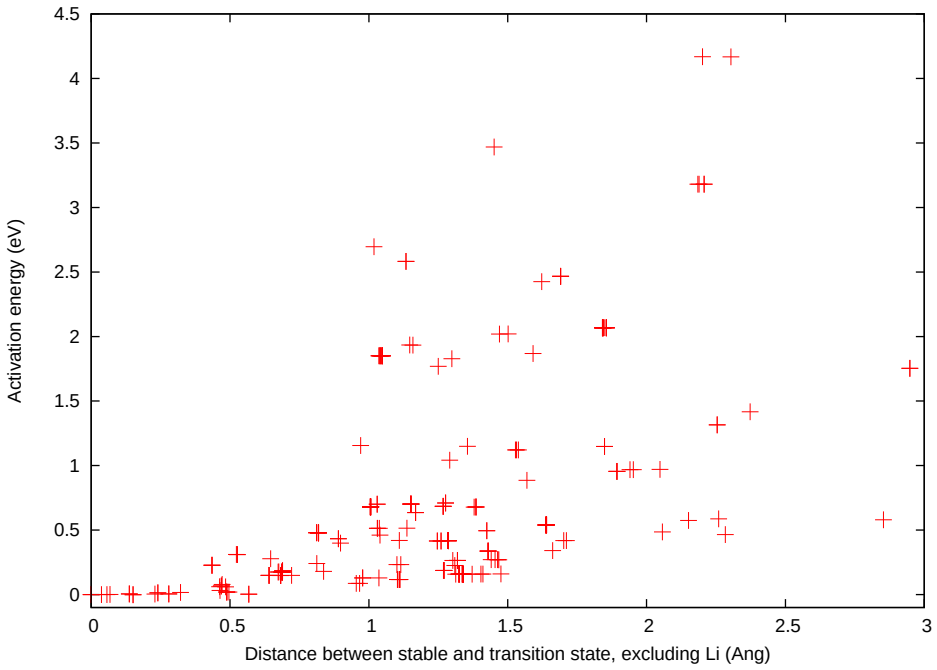


Figure 6.12: Distance between stable and transition state, excluding Li, for events in our sample, in Å, versus activation energy in eV

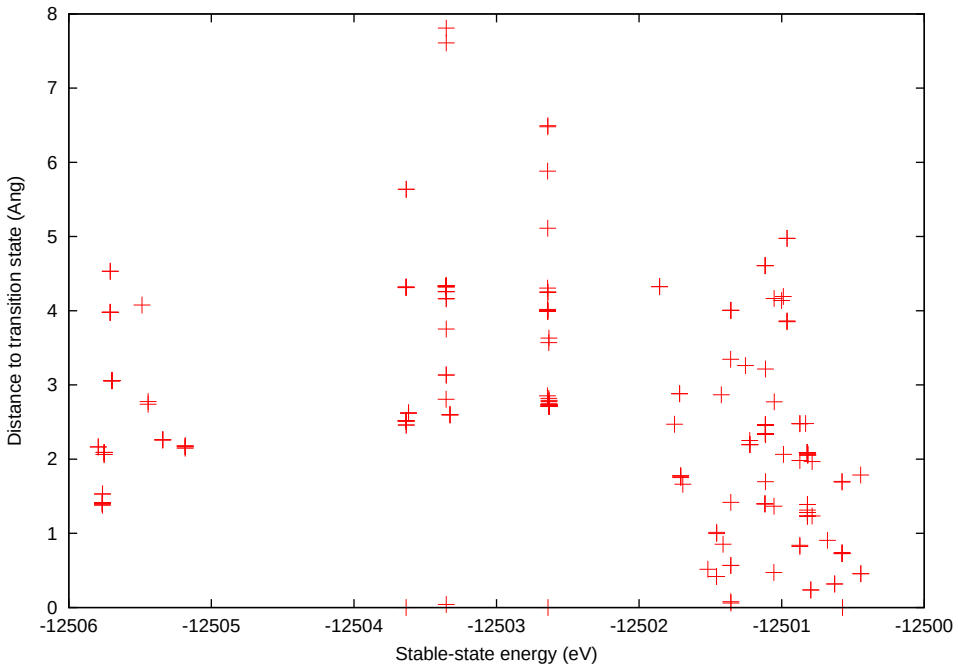


Figure 6.13: Magnitude of Si–O network distortion in the transition state (in Ang) versus the energy of the stable site (in eV) to which the transition state is connected, and from which the distortion is measured.

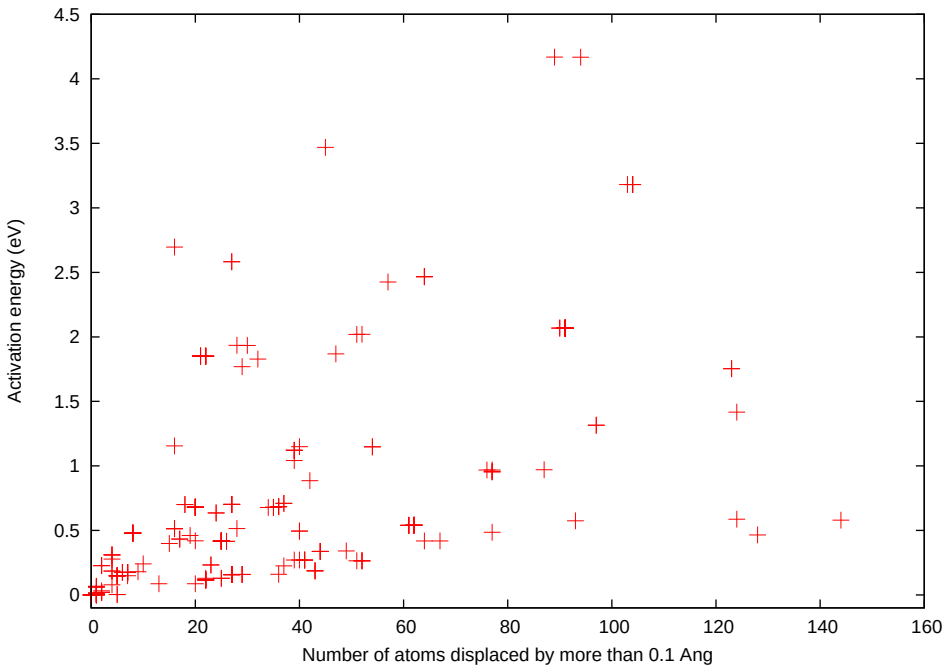


Figure 6.14: Number of atoms displaced by more than 0.1 Å in a transition-state event in our sample, versus activation energy in eV

An approach to take from here is to consider the number of atoms displaced in any given transition event. Breaking the distortion vector obtained above down into its atomic components, one can set some distance as a tolerance – say, 0.05 Å – and count the number of atoms which are displaced by at least this distance. In the case of 0.1 Å, and plotting number of atoms displaced versus activation energy, this gives rise to Figure 6.14. What is immediately apparent is that the displacement of a very large number of atoms is not necessarily a bar to a given transition event being relatively low-energy; there appear to be several overlaid curves in this figure, but considering the lowest, there is an event which causes the displacement of more than 140 atoms – around 25 percent of the entire cell! – yet is only 0.5eV. The simplest explanation for this is the effect of long-range collective distortions – i.e., rigid unit modes – in our system.

This opens the question of the size – in terms of unit cell volume – of low-energy

diffusion events in our sample.

By varying the cutoff tolerance, the number of atoms displaced by more than some given distance can be calculated for each of our transition state/stable state pairs. Considering only the pairs of stable and transition states with an activation energy under 0.5 eV, a plot of number of atoms versus cutoff (as in figure 6.15) is calculated for each cutoff; it is found, as can be seen in this figure, that the number of atoms placed by events of a given magnitude varies linearly with the cutoff. Therefore, varying the tolerance and fitting the resulting data, by linear regression, to $F(x) = cx$, a succession of gradients are obtained; each gradient therefore links directly the number of atoms displaced by a given distance in an even with the magnitude of that event.

These gradients can then be plotted against the cutoffs themselves, and this is shown in figure 6.16.

It is apparent that this conforms to a function of the form $F(x) = Ae^{-bx}$; the magnitude of displacement of Si and O atoms within a given transition state event conforms to a power law distribution (equation 6.3, where γ is the cutoff, in Angstroms.)

$$N = 331.5 \pm 25.3e^{(-29.4\pm1.5)\gamma} \tag{6.3}$$

Using this, it is possible to estimate of the size of a typical transition event. Using the same set of transitions (those under 0.5 eV), it is found that there are 139 events with a mean displacement of 1.09 Å; we then take the critical region to be all those atoms displaced by more than 0.1 Å in a given event. These values can then be substituted into Equation 6.3, which tells us that, on average, 19.1 Si and/or O atoms are displaced in one of these events. The supercell used in our simulations has volume 10146 $Å^3$, and contains 648 relevant atoms; as such, a typical diffusion event in the range of 0.5 eV in this sample has a strain field of volume 299 $Å^3$ – which, if the assumption is made that the strain field is spherical, corresponds to a radius of 4.14 Å. Some events, of course, are enormously larger, as we have seen, but

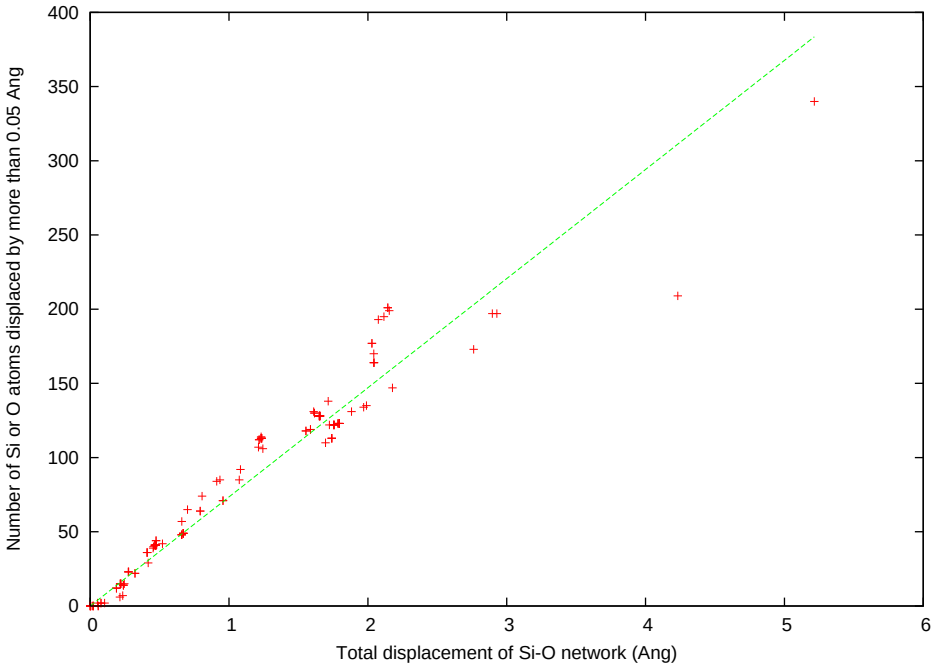


Figure 6.15: Number of atoms displaced by more than 0.1 Å in a transition-state event in our sample, versus magnitude of Si–O lattice distortion (in Ang)

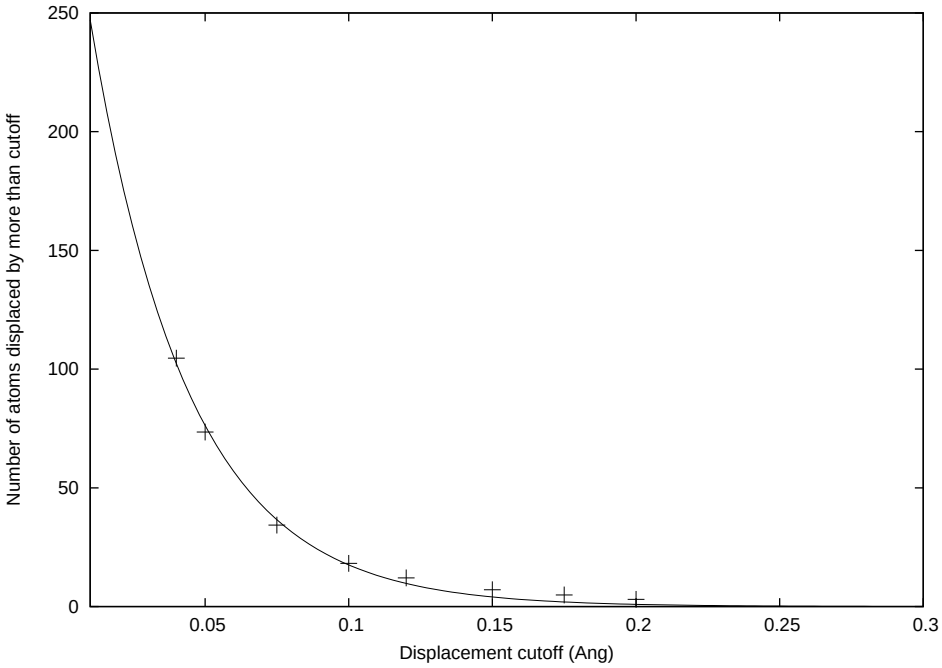


Figure 6.16: Number of atoms displaced by more than cutoff x (in Angstroms) versus cutoff; the curve shown has the equation .

it is unarguable that long-range strain effects will play a major part in determining diffusion behaviour.

Considering this, the remarkable effect of even a single Al substitution on the transition states in quartz, observed in the previous chapter, becomes more explicable. Even if a substitution has a minimal effect on the geometry of the system, it will have some effect on the local elastic relaxation of the environment. Indeed, if nothing else, the present work demonstrates how even such a seemingly local process as motion of Li can be affected by the extended environment – at the very least, one has to consider a sphere at least eight Å across in diameter around the stable state, which is of the same order of magnitude, volume-wise, as the unit cell in α-quartz!

Chapter 7

Conclusions

If there is a recurring theme throughout this thesis, it is the interplay between the local and the global; between short and long timescales and lengthscales. In each of the systems I have studied, events or modifications on the atomic scale – doping in WO_3, crowdion formation on platinum surfaces, the interplay between structure and (particularly) structural flexibility and light-atom diffusion in framework silicates and silica glass – have had measurable, quantifiable long-range effects, whether mediated through the effects of electron delocalization in WO_3 or long-range strain effects carried through rigid unit modes.

For scientists active in the area of solid-state simulation, the lengthscale and timescale problems are among the most challenging we face on a daily basis. These are particularly pressing when one considers, as I have, disordered systems; one cannot directly simulate the dynamics, because they happen too slowly, and one must always be wary of simulating too small a system and thus artifically imposing spurious symmetry.

In the lengthscale case, the primary technique I have used to obviate the problem (that is to say, to avoid having to use an extremely large supercell) is to treat the doping process as symmetrically as one can get away with – in other words,

(partial) delocalisation of the disorder (as in quartz and WO_3). In the WO_3 case, I disentangled the effect of electron doping from chemical disorder and size effects; this made the root cause of the complicated series of phase transitions observed on doping explicit in a way which could not otherwise be achieved. Of course, great care needs to be taken to ensure that this is a valid approximation for any given system, but the results presented are very positive; in particular, the reinterpretation of the $NaWO_3$ experimental record in light of our results is of interest. Since the publication of the work in Chapter 3, I was pleased to note other authors [96, 97, 98, 99, 100, 101] have taken up the method for study of related defects in WO_3 and, in the case of Ferrari, other perovskites; I believe it would be worthwhile to extend the study in the thesis to further materials in that class in the future. Also, there is more to be said about WO_3, as the work of Lambert-Mauriat and others indicates; the work in this thesis opens the debate, rather than closing it.

The lengthscale problem, of course, is still an issue when we move onto the work presented on Li^+ diffusion in silicates; the results in Chapter 5 show just how difficult it is to separate one effect (the Li–lattice interaction) from another (the Li–Al interaction), and Chapter 5 and 6 both demonstrate just how large – spatially – the strain field arising from even seemingly local events like a Li^+ hop can be! The main thrust of this work, however, has been overcoming the timescale problem and through that saying something about the dynamics of diffusion in this material; of course, this is a fiercely difficult problem which could occupy several PhD students for far longer than any one PhD could take, but during the present PhD I have made progress on several fronts.

Firstly, I have developed a new technique for transition state detection in general. It has a few key differences from other techniques, such as the Dimer method [3], as elucidated in Chapter 4; in some sense, the remainder of the thesis acts as an extended field-trial of the method, and it is at least a viable option as of now for real research. Additionally, as Chapter 4 demonstrates, for certain potentials – particularly certain ab-initio codes – it has very real advantages over the methods currently available. As of now, it is not numerically as efficient as competing methods, but

very little effort has been placed into performance tuning; further development of the CLM technique is one possible path in the future.

Secondly, our results on silicates and glasses (Chapters 5 and 6) are, in and of themselves, interesting insights into the structure and behaviour of these systems; regardless of the technique used, the work here stands by itself. The quartz results accord well with the prior results of Calleja [26] and Sartbaeva et al. [6, 17], whilst expanding substantially on their conclusions; the situation in glass is somewhat more complex, but the approach taken has gone some distance in explaining the role which cooperative distortion of the Si–O network (and by extension, rigid unit modes) play in permitting diffusion in these types of materials. Once again, my work on glass is a continuation of a scientific debate rather than the end of one, but it points to approaches that may not previously have been seriously considered; it would be worthwhile, for instance, to use a method like on-the-fly kinetic Monte Carlo, similar to that proposed recently by Henkelman and Jonsson recently [102], to obtain realistic predictions of diffusion coefficients.

In any case, it has been unambiguously established that direct simulations of glass dynamics by using transition-state search are certainly reasonable: this opens up many classes of disordered system to simulation. Indeed, one aspect which Chapter 6 makes very clear is that the primary challenge, in many ways, is not simulation itself – we can easily generate gigabytes of data – but the analysis of the results. The approaches taken in this thesis are effective, but another which may be fruitful in the future for these kinds of network structures is the melding of geometric-algebra methods, like that of Wells [103], with direct atomistic search of the kind I use. This holds promise both for accelerated search and for analysing the resulting trajectories.

In conclusion, by using novel techniques and careful approximations, I hope I have demonstrated some new approaches which can be used in the study of disordered systems; the methods developed in this thesis are a contribution to the long process of developing the appropriate computational and analytical techniques, and point to new vistas of materials which we can now examine. There is much still to be done in

these areas of work, whether by the present author or by other scientists; however, through the work presented here, I hope that I have gone some way to clearing a path ahead.

Bibliography

[1] S. G. Louie, S. Froyen, and M. L. Cohen. Nonlinear ionic pseudopotentials in spin-density-functional calculations. *Phys. Rev. B.*, B26, 1982.

[2] G. A. de Wijs, P. K. de Boer, R. A. de Groot, and G. Kresse. *Phys. Rev. B.*, 59(4):2684, 1999.

[3] G. Henkelman and H. Jonsson. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.*, 111(15):7010, 1999.

[4] K. Kihara. Thermal change in unit cell dimensions, and a hexagonal structure of tridymite. *Zeitung f ur Kristallographie*, 148:237, 1978.

[5] J. J. Pluth, J. V. Smith, and Jr. J. Faber. Crystal structure of low cristobalite at 10, 293, and 473 K: Variation of framework geometry with temperature. *J. Appl. Phys.*, 57(4):1045, 1985.

[6] A. Sartbaeva, S. A. Wells, and S. A. T. Redfern. Li^+ ion motion in quartz and β-eucryptite studied by dielectric spectroscopy and atomistic simulations. *J. Phys: Cond. Mat.*, 16(46):8173, 2004.

[7] Tetsu N., Atsushi K., and Takahiro K. An 8.61 Tflop/s molecular dynamics simulation for NaCl with a special-purpose computer: MDM. *Supercomp.*, page 11, 2001.

[8] W. Kohn and L. J. Sham. *Phys. Rev. A*, 140:1133, 1965.

[9] P. Hohenberg and W. Kohn. *Phys. Rev. B.*, 136:864, 1964.

[10] W. L. McMillan. Ground state of liquid ^{4}He. *Phys. Rev.*, 138(2A):A442, 1965.

[11] C. G. Granqvist, editor. *Handbook of Inorganic Electrochromic Materials*. Elsevier, Amsterdam, 1995.

[12] P. M. S. Monk, R. J. Mortimer, and D. R. Rosseinsky. *Electrochromism: Fundamentals and Applications*. VCH Verlagsgesellschaft, Weinheim, 1995.

[13] M. C. Gallardo, J. Jiménez, J. del Cerro, and E. K. H. Salje. The structural phase transition in img under uniaxial stress in the [110] direction: a calorimetric study. *J. Phys: Cond. Mat.*, 8(1):83, 1996.

[14] M. Dawber and J. F. Scott. A model for fatigue in ferroelectric perovskite thin films. *Appl. Phys. Lett.*, 76(8):1060, 2000.

[15] A. M. Glazer. *Acta Crystallogr., Sect B*, B28:3384, 1972.

[16] S. Reich, G. Leitus, Y. Tssaba, Y. Levi, A. Sharoni, and O. Millo. Localised high-T_c superconductivity on the surface of Na-doped WO_3. *J. Superconductivity*, page 855, 2000.

[17] A. Sartbaeva, S. A. Wells, S. A. T. Redfern, R. W. Hinton, and S. J. B. Reed. Ionic diffusion in quartz studied by transport measurements, SIMS and atomistic simulations. *J. Phys: Cond. Mat.*, 17(7):1099, 2005.

[18] W. Press, B. Renker, H. Schulz, and H. Böhm. Neutron scattering study of the one-dimensional ionic conductor β-eucryptite. *Phys. Rev. B*, 21(3):1250, 1980.

[19] J. Habasaki and Y. Hiwatari. Characteristics of slow and fast ion dynamics in a lithium metasilicate glass. *Phys. Rev. E*, 59(6):6962, 1999.

[20] J. M. Soler, E. Artacho, J. D. Gale, A. García, Javier Junquera, Pablo Ordejón, and Daniel Sánchez-Portal. *J. Phys: Cond. Mat.*, 14:2745, 2002.

[21] J. D. Gale. GULP: A computer program for the symmetry-adapted simulation of solids. *J Chem. Soc. Faraday Trans.*, 93:629, 1997.

[22] K. W. Jacobsen, J. K. Norskov, and M. J. Puska. Interatomic interactions in the effective-medium theory. *Phys. Rev. B*, 35(14):7423, 1987.

[23] P. Ewald. Die berechnung optischer und elektrostatischer gitterpotentiale. *Ann. Phys.*, 64:253, 1921.

[24] M. T. Dove. *Introduction to Lattice Dynamics*. Cambridge University Press, 1993.

[25] J. D. Gale. *General Utility Lattice Program, version 3.0*, 2006.

[26] M Calleja, M T Dove, and E K H Salje. Anisotropic ionic transport in quartz: the effect of twin boundaries. *J. Phys: Cond. Mat.*, 13(42):9445, 2001.

[27] G. J. Kramer, N. P. Farragher, B. W. H. van Beest, and R. A. van Santen. Interatomic force fields for silicas, aluminophosphates, and zeolites: Derivation based on ab initio calculations. *Phys. Rev. B*, 43(6):5068, 1991.

[28] M. Born and J. R. Oppenheimer. *Ann. Phys. (Leipzig)*, 484:457, 1927.

[29] E. Schrödinger. *Ann. Phys. (Leipzig)*, 489:79, 1926.

[30] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal. Quantum monte carlo simulations of solids. *Rev. Mod. Phys.*, 73(1):33, 2001.

[31] A. D. Becke. *J. Chem. Phys.*, 98:1372, 1992.

[32] C Lee, W Yang, and R G Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B.*, 37(2):785, 1988.

[33] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B*, 33(12):8822, 1986.

[34] N. Trouiller and J. L. Martins. *Phys. Rev. B.*, 43, 1993.

[35] P Ordejón, D A Drabold, M P Grumbach, and R M Martin. Unconstrained minimization approach for electronic computations that scales linearly with system size. *Phys. Rev. B*, 48(19):14646, 1993.

[36] O F Sankey and D J Niklewski. Ab initio multicenter tight-binding model for molecular-dynamics simulations and other applications in covalent systems. *Phys. Rev. B*, 40(6):3979, 1989.

[37] G. van Rossum. Python programming language, version 2.3.

[38] Travis E. Oliphant and Paul F. Dubois. SciPy Extension Library.

[39] K. W Jacobsen. Centre for Atomic-scale Materials Physics Open Software Project. 2006.

[40] P. M. Murray-Rust and H. S. Rzepa. Chemical markup, XML, and the World-wide Web. 1. Basic Principles. *J. Chem. Inf. Comput. Sci.*, 39:928, 1999.

[41] T Bray, J Paoli, C M Sperberg-McQueen, and E Maler. Extensible Markup Language (XML) 1.0 (Second Edition). 2000.

[42] W3 Consortium XHTML 1.0 Working Group. XHTML 1.0 - The Extensible Hypertext Markup Language (Second Edition). 2002.

[43] A. S. Yu, N. Kumagai, Z. L. Liu, and J. Y. Lee. Electrochemical lithium intercalation into WO_3 and lithium tungstates $Li_xWO_{\frac{x}{2}}$ of various structures. *J. Sol. Stat. Electrochem*, 2(6):394, 1998.

[44] J. J. Berzelius. *J. Chem. Phys (Berlin)*, 16:476, 1816.

[45] B. W. Brown and E. Banks. *J. Am. Chem. Soc*, 76:963, 1954.

[46] G. Kohlstrung. Über WO_3-umwandlungen bei der thermischen zersetzung der ammoniumparawolframate. *Phys. Stat. Sol*, 2:85, 1962.

[47] R. Clarke. *Phys. Rev. Lett.*, 39(24):1550, 1977.

[48] F. Corà, M. G. Stachiotti, C. R. A. Catlow, and C. O. Rodriguez. *J. Phys. Chem. B*, 101:3945, 1997.

[49] M. G. Stachiotti, F. Corà, C. R. A. Catlow, and C. O. Rodriguez. First-principles investigation of ReO_3 and related oxides. *Phys. Rev. B*, 55(12):7508, 1997.

[50] N. A. Hill. Why are there so few magnetic ferroelectrics? *J. Phys. Chem. B*, 104(29):6694, 2000.

[51] A. Hjelm, C. G. Granqvist, and J. M. Wills. *Phys. Rev. B.*, 54(4):2436, 1996.

[52] F. Detraux, P. Ghosez, and X. Gonze. *Phys. Rev. B.*, 56(3):983, 1997.

[53] P. Ordejón, E. Artacho, and J. M. Soler. *Phys. Rev. B.*, 53:10441, 1996.

[54] J. P. Perdew and A. Zunger. *Phys. Rev. B.*, 23:5048, 1981.

[55] L. Kleinman and D. M. Bylander. *Phys. Rev. Lett.*, 48:1425, 1982.

[56] E. Anglada, J. M. Soler, J. Junquera, and E. Artacho. Systematic generation of finite-range atomic basis sets for linear-scaling calculations. *Phys. Rev. B.*, 66:205101, 2002.

[57] J. Moreno and J. M. Soler. *Phys. Rev. B.*, 45:13891, 1992.

[58] H.J. Monkhorst and J. D. Pack. *Phys. Rev. B.*, 13:5188, 1976.

[59] E. K. H. Salje, S. Rehmann, F. Pobell, D. Morris, K. S. Knight, T. Hermmansdörfer, and M. T. Dove. *J. Phys: Cond. Mat.*, 9:6564, 1997.

[60] O. Loopstra and H. M. Rietveld. *Acta Crystallogr., Sect. B*, B25, 1975.

[61] E. K. H. Salje and K. Viswanathan. *Acta Crystallogr., Sect. A*, A31, 1975.

[62] S. Tanisaki. *J. Phys. Soc. Jpn*, 14:566, 1960.

[63] M. B. Robin and P. Day. *Adv. Inorg. Chem. Radiochem.*, 10:247, 1967.

[64] I. B. Bersuker. *Electronic Structure and Properties of Transition Metal Compounds: Introduction to the Theory*. Wiley and Sons, New York, 1996.

[65] P. W. Atkins. *Physical Chemistry (Sixth Edition)*. Oxford University Press, Oxford, 1998.

[66] A. F. Voter. Hyperdynamics: Accelerated molecular dynamics of infrequent events. *Phys. Rev. Lett.*, 78(20):3908, 1997.

[67] A. Laio and M. Parrinello. Escaping free-energy minima. *PNAS*, 99(20):12562, 2002.

[68] K. Fukui. Formulation of the reaction coordinate. *J. Phys. Chem*, page 4161, 1970.

[69] H. Jónsson, G. Mills, and K. W. Jacobsen. *Classical and Quantum Dynamics in Condensed Phase Simulations*. World Scientific, 1998. Nudged Elastic Band Method for Finding Minimum Energy Paths of Transitions.

[70] G. T. Barkema and N. Mousseau. The activation-relaxation technique: an efficient algorithm for sampling energy landscapes. *Comp. Mat. Sci.*, 20:285, 2000.

[71] L. J. Munro and D. J. Wales. Defect migration in crystalline silicon. *Phys. Rev. B*, 59(6):3969, 1999.

[72] Y. Y. Mishin. *Defect and Diffusion Forum*, 143-147:1357, 1997.

[73] S. Sato. Potential energy surface of the system of three atoms. *J. Chem. Phys.*, 23(12):2465, 1955.

[74] A. Banerjee, N. Adams, J. Simons, and R. Shepard. Search for stationary points on surfaces. *J. Phys. Chem.*, 89:52, 1985.

[75] C. J. Cerjan and W. H. Miller. *J. Chem. Phys.*, 75:2800, 1981.

[76] R. A. Olsen, G. J. Kroes, G. Henkelman, A. Arnaldsson, and H. Jonsson. Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys.*, 121(20):9776, 2004.

[77] G. L. Kellogg and Peter J. Feibelman. Surface self-diffusion on Pt(001) by an atomic exchange mechanism. *Phys. Rev. Lett.*, 64(26):3143, 1990.

[78] W. Xiao, P. A. Greaney, and D. C. Chrzan. Pt adatom diffusion on strained Pt(001). *Phys. Rev. B*, 70(3):033402, 2004.

[79] R. H. Doremus. Exchange and diffusion of ions in glass. *J. Phys. Chem.*, 68:2212, 1964.

[80] M. G. Tucker, D. A. Keen, and M. T. Dove. A detailed structural characterization of quartz on heating through the α–β phase transition. *Min. Mag.*, page 489, 2001.

[81] P. J. Heaney and D. R. Veblen. Observations of the α-β phase transition in quartz: A review of imaging and diffraction studies and some new results. *Am. Min.*, 76:1018, 1991.

[82] K. D. Hammonds, M. T. Dove, A. P. Giddy, V. Heine, and B. Winkler. Rigid-unit phonon modes and structural phase transitions in framework silicates. *Am. Min.*, 81:1057, 1996.

[83] A. Putnis. *Introduction to Mineral Sciences*. Cambridge University Press, 1992.

[84] A. P. Giddy, M. T. Dove, G. S. Pawley, and V. Heine. The determination of rigid-unit modes as potential soft modes for displacive phase transitions in framework crystal structures. *Acta Crystallogr., Sect. A*, 49(5):697, 1993.

[85] M. T. Dove, A. K. A. Pryde, and D. A. Keen. Phase transitions in tridymite studied using "rigid unit mode" theory, reverse Monte Carlo methods and molecular dynamics simulations. *Phys. Chem. Miner.*, 64 (2):267, 2000.

[86] A. K. A. Pryde and M. T. Dove. On the sequence of phase transitions in tridymite. *Phys. Chem. Miner.*, 26:171, 1998.

[87] J. Ellson, E. Gansner, Y. Koren, E. Koutsofios, J. Mocenigo, S. North, and G. Woodhull. `graphviz` *manual*, 2006.

[88] P. Campone, M. Magliocco, G. Spinolo, and A. Vedda. Ionic transport in crystalline SiO_2: The role of alkali-metal ions and hydrogen impurities. *Phys. Rev. B*, 52(22):15903, 1995.

[89] M. Gambhir, M. T. Dove, and V. Heine. Rigid unit modes and dynamic disorder: SiO_2 cristobalite and quartz. *Phys. Chem. Miner.*, 26 (6):484, 1999.

[90] S. A. Wells, M. T. Dove, M. G. Tucker, and K. Trachenko. Real-space rigid-unit-mode analysis of dynamic disorder in quartz, cristobalite and amorphous silica. *J. Phys: Cond. Mat.*, 14(18):4645, 2002.

[91] K. Trachenko and M. T. Dove. Floppy modes in silica glass under pressure. *J. Phys: Cond. Mat.*, 14(6):1143, 2002.

[92] F. Wooten, K. Winer, and D. Weaire. Computer generation of structural models of amorphous Si and Ge. *Phys. Rev. Lett.*, 54(13):1392, 1985.

[93] W. H. Zachariasen. The atomic arrangement in glass. *J. Am. Chem. Soc.*, 54:3841, 1932.

[94] S. Tsuneyuki, M. Tsukada, H. Aoki, and Y. Matsui. First-principles interatomic potential of silica applied to molecular dynamics. *Phys. Rev. Lett.*, 61(7):869, 1988.

[95] D. A. Keen and M. T. Dove. Total scattering of silica polymorphs: similarities in glass and disordered crystalline local structure. *Min. Mag.*, 64:447, 2000.

[96] V. Ferrari, J. M. Pruneda, and E. Artacho. Density functionals and half-metallicity in $La_{2/3}Sr_{1/3}MnO_3$. *Phys. Stat. Sol. (a)*, 203:1437, 2006.

[97] C. Lambert-Mauriat and V. Oison. Density-functional study of oxygen vacancies in monoclinic tungsten oxide. *J. Phys: Cond. Mat.*, 18(31):7361, 2006.

[98] B. Ingham, S. C. Hendy, S. V. Chong, and J. L. Tallon. Density-functional studies of tungsten trioxide, tungsten bronzes, and related systems. *Phys. Rev. B*, 72(7):075109, 2005.

[99] C. V. Ramana, S. Utsunomiya, and R. C. Ewing. Electron microscopy investigation of structural transformations in tungsten oxide (WO_3) thin films. *Phys. stat. sol. (a)*, 202 (10):108, 2005.

[100] J. F. Scott. Response to "Comment on 'Infrared and second harmonic generation in barium strontium titanate and lead zirconate titanate thin films: 'Polaron' artifacts' " [J. Appl. Phys. [96], 2409 (2004)]. *J. Appl. Phys*, 96(4):2410, 2004.

[101] Q. Wang, Z. Wen, Y. Jeong, J. Choi, K. Lee, and J. Li. Li-driven electrochemical properties of WO_3 nanorods. *Nanotech.*, 17:3116, 2006.

[102] G. Henkelman and H. Jonsson. Long time scale kinetic monte carlo simulations without lattice approximation and predefined event table. *J. Chem. Phys.*, 115(21):9657, 2001.

[103] S. A. Wells, M. T. Dove, and M. G. Tucker. Finding best-fit polyhedral rotations with geometric algebra. *J. Phys: Cond. Mat.*, 14(17):4567, 2002.