

KING'S COLLEGE LONDON

UPGRADE REPORT

**Discovering atomic scale mechanisms
of stress corrosion cracking in
aerospace titanium alloys.**

Author:

Tigany Noor Abubaker Tigany Zarrouk

Supervisor:

Prof. Tony Paxton

Theory & Simulation of Condensed Matter
Physics Department
King's College London

Contents

1	Introduction	3
2	Dislocations and titanium	5
2.1	Metallurgy of titanium	5
2.2	Dislocations	6
2.2.1	Classical models for dislocation simulation	7
2.3	Titanium Slip Systems and Plasticity	8
2.3.1	Dislocation Cores and Effect of Impurities	9
2.4	Computational Simulations of titanium	11
2.4.1	Dislocations	11
2.4.2	Stacking Faults	12
2.4.3	Embedding Methods	12
2.5	Numerical Optimisation	13
3	Quantum Mechanics and Numerical Optimisation	15
3.1	Density Functional Theory	15
3.2	The Hohenberg-Kohn-Sham Functional	15
3.3	Non-self-consistent Tight Binding	16
3.3.1	Distance Dependence of Bond Integrals	17
3.3.2	Tight Binding Band Model	18
3.3.3	Tight Binding Bond Model	18
3.3.4	Local Charge Neutrality	19
3.4	Self-consistent Tight Binding	19
3.4.1	Self-consistent Charge Transfer Model	19
3.4.2	Polarisable-ion tight binding	20
3.5	Bond Order Potentials	20
3.5.1	Moment Theorem	21
3.5.2	Lanczos Algorithm	21
3.5.3	The Green's Function	21
3.6	Embedding: tight binding and BOP	23
3.6.1	Block-matrix Solution	24
3.7	Elastic constants	25
3.8	Bayesian Optimisation	27
3.8.1	Bayesian Linear Regression	27
3.8.2	Gaussian Processes for Regression	30
3.8.3	Expected Improvement	32
4	Results and Discussion	33
4.1	Tight Binding Band Model: oxygen	33

4.2	Tight Binding Band Model: nitrogen	37
4.3	Tight Binding Band Model: titanium	37
4.4	Fitting of Parameters for Pure titanium	38
4.5	Numerical Optimisation: Tests	41
4.6	Numerical Optimisation: titanium	44
5	Future Work	49
5.1	Outline of Strategy	49
5.2	Gantt Chart	50
6	Conclusion	53
7	Notes	55
7.1	Notes on BZ integration	55
7.2	Why are electron states in the band gap	56
7.3	What is a polaron	56
7.4	Electron-hole Interaction	56
7.5	Stoner Magnetism	57
7.6	Friedel Oscillations	57
7.7	Surface States	57
7.8	Mott -Hubbard Insulator	57
7.9	Notes on reversible integrators	58
7.10	Polarisable-ion Tight Binding	58
7.11	Universal Tight Binding III	59
7.12	Electronic structure of reduced titanium dioxide	59
7.13	Calculation of TiO ₂ Surface and Subsurface Oxygen Vacancy by the Screened.	60
7.14	Surface and subsurface oxygen vacancies in anatase TiO ₂ and differences with rutile.	61
7.15	A model for time-dependent grain boundary diffusion of ions and electrons through a film or scale, with an application to alumina	61
7.16	Bayesian Optimisation Spiel	61
7.17	Additional BOP Notes	62
8	TBE to BOP Ti	63
8.1	Ti Dimer	63
8.2	Elastic and lattice constants in BOP	65
8.2.1	Four recursion levels	65
8.3	Five recursion levels	68
8.4	Six Recursion Levels	71
8.5	Anisotropic Elasticity Solutions for Hexagonal material	75
8.6	Checks for other codes	75

Chapter 1

Introduction

Titanium is an essential material in aerospace engineering. This is due to its corrosion resistance, strength-to-weight ratio and resilience at high operating temperatures. For this reason, it finds applications in turbojets for current and next-generation civilian aviation, as well as more general military and space-based applications. However, for all the incredible properties titanium possesses, it is not perfect.

Serious failures have been noted recently by the cracking and corrosion of titanium turbine fan blades. In Rolls-Royce Trent 1000 engines, this caused mass grounding of planes with an estimated repair cost of 340 million [?]. As has also been the case this year, failure of fan blades during flight can cause unfortunate fatalities [?]. These cases could have been avoided by thorough scientific investigation into the causes of corrosion and failure.

The corrosion resistance of titanium is from its strong, but brittle, oxide layer—titanium dioxide—which quickly forms in a gaseous oxygen environment. It has been thought, as fan blades are situated downstream from the the combustion chamber of jet engines—with an environment of elevated temperature—there is an enhanced susceptibility of corrosion by an increased rate of oxygen diffusion, causing a thicker, brittle oxide layer, greatly increasing vulnerability to cracking and failure.

Cracking can occur within a material by the build-up and propagation of defects, the most important being dislocations. The ability for materials to undergo plastic deformation without cracking is determined by how mobile dislocations are on their slip planes. The structure of dislocation cores governs the active slip planes available for deformation.

Oxygen-induced changes to dislocation core structure have been implied by significant hardening and modifications to the onset of strain with rate of increasing stress (strain-rate sensibility of stress), with oxygen content. The mechanism for how oxygen interacts with these cores, consequently modifying such macroscopic properties of titanium, has not been fully elucidated. If such an explanation can be achieved, it would result in a more complete picture of plasticity and fracture in titanium, from which improvements to its materials can be made.

Investigation into oxide layer growth, from oxygen diffusion in a pressured environment, would provide crucial insights into the onset of brittleness in these materials. In conjunction with dislocation simulations at the oxide layer/titanium interface, scrutiny of Stress Corrosion Cracking (SCC) and its proposed exacerbation at areas of high oxygen concentration, can be achieved.

Extending the investigation into alloys and more solutes greatly increases the complexity of interaction with dislocation cores. Impurities such as nitrogen, always prevalent due to its high solubility, occupies interstitial sites, like oxygen, suggesting similar interaction with cores. Alloying additions, such as Al and V, as used in the ubiquitous alloy Ti-6Al-4V, are however substitutional solutes. How such impurities and alloying additions interact and affect the core structure of dislocations in the material, thereby modifying plasticity, has not been explored at all [?].

Simulating such systems requires a large scale such that the dislocation core and its strain fields can be accurately replicated. Long range strain fields produced from dislocations are particularly sensitive to boundary conditions. The stalwart of computation electronic structure, Density Functional Theory (DFT), is not suitable for this reason. Small system sizes (on the order of a few hundred atoms) result from its computational expense, such that external boundaries can hinder both core and long range field development [? ?]. Furthermore, DFT fails when estimating the band gaps of systems making it unreliable for simulation of transition metal oxides [?]. So a different model has to be adopted, which: has sufficient scaling properties to accurately model the long range strain fields of dislocations; does not hinder the relaxation of atoms in the dislocation core; retains the quantum accuracy necessary to model non-linear stresses in the core and faithfully models the transition metal oxide polymorphs of TiO_2 .

Accurate modelling of TiO_2 and its surface under influence of a partial pressure environment can be done within tight binding [?]. The Bond Order Potential (BOP) method has been used to great effect in simulation of titanium and titanium aluminides [? ?]. Combining these two approaches would enable very large system sizes (on the order of thousands of atoms), circumventing the problems apparent in DFT. This can be achieved by an embedding scheme.

Embedding is a multi-scale framework usually comprising of a synthesis two or more methods of simulation. This generally involves an accurate method (e.g. tight-binding of $\mathcal{O}(N^3)$) governing an inner region of importance—such as a dislocation core—coupled to a more approximate and faster method (e.g. BOP of $\mathcal{O}(N)$) which governs a larger, outer region. This concentrates computational accuracy where its needed, while providing a larger atomic environment such that larger scale effects can be seen.

The gain to be had from this scheme is clear: the parametrisation of tight binding and BOP methods allow flexibility compared to DFT, allowing accurate replication of band structures, electronic bonding interactions and empirically measured quantities, while also obtaining sufficient system sizes such that dislocations and alloyed materials are able to be investigated. This provides a method for: exploration of oxygen and alloying element interaction with dislocation cores; determination of oxide scale growth mechanisms by observation of oxygen/titanium diffusion between the titanium (alloy), oxide layer and a partial pressure oxygen environment and investigation of stress corrosion cracking of the oxide scale by synergetic dislocation and oxygen interactions.

Chapter 2

Dislocations and titanium

2.1 Metallurgy of titanium

Titanium has two main phases: α -phase (Hexagonal close-packed—HCP) and β -phase (Body-Centred Cubic—BCC). Generally, α -phase is seen at lower temperature and β -phase is seen at higher temperatures (above 420K). These two phases can coexist giving excellent microstructural properties when alloyed with α stabilisers (such as oxygen, nitrogen and aluminium) or β stabilisers, like vanadium [?]. This multi-phase coexistence can be seen in Ti-6Al-4V. β -titanium can further transition into ω -phase by absorption of the $q = \frac{d}{3}(1, 1, 1)$ longitudinal phonon [?].

The origin of strength in titanium alloys is the barrier to dislocation motion. Inclusion of oxygen, nitrogen and sulphur increases the strength of the alloy, but this is a double-edged sword, as these inclusion commonly initiate failure due to increases in brittleness [?].

Titanium's incredible corrosion resistance comes from its oxide layer. It has a complex form comprising of rutile TiO_2 and α -case—an oxygen-rich α -phase. As shown by Bagot *et al.*, below the oxide surface of Ti-6Al-4V, the microstructure separates into a nanoscale mixture of the α -Ti and α_2 Ti_3Al phases with the α_2 phase existing as either nanoscale spherical precipitates or as bands [?]. Nitrogen also penetrates the surface, but to a lesser extent, while vanadium separates into α_2 or other phases. Oxygen diffusion is around two orders of magnitude faster in β than in α -phase. This disparity in diffusion was theorised to contribute to the formation of the observed banded microstructures, when α/β interfaces were present. With gradually increasing oxygen content in α -Ti, precipitation of near-spherical α_2 nanoscale precipitates occurred, because of the lack of α/β interfaces. Both of these mechanisms have yet to be seen on an atomistic scale.

Stress corrosion cracking (SCC) is the process whereby a corrosive chemical species and tensile stress act together to cause cracking. Oxygen, hydrogen, chlorides and hydroxides are such substances. Hydrogen SCC has a mechanism whereby hydrogen migrates to the crack tip, embrittling the material, promoting crack growth [?]. Bath *et al.*, found that the concentration of oxygen inside titanium was a large factor in promoting crack growth [?]. In the aforementioned work of Bagot, it was concluded that oxygen stimulates precipitation of α_2 , which explains embrittlement produced in the O enriched layer beneath the oxide [?]. It still remains to be seen how SCC manifests in titanium on an atomistic scale, with stresses induced by dislocations. The diffusion rate of oxygen is much less than that of Hydrogen, so it is not clear if the same mechanism of embrittlement occurs.

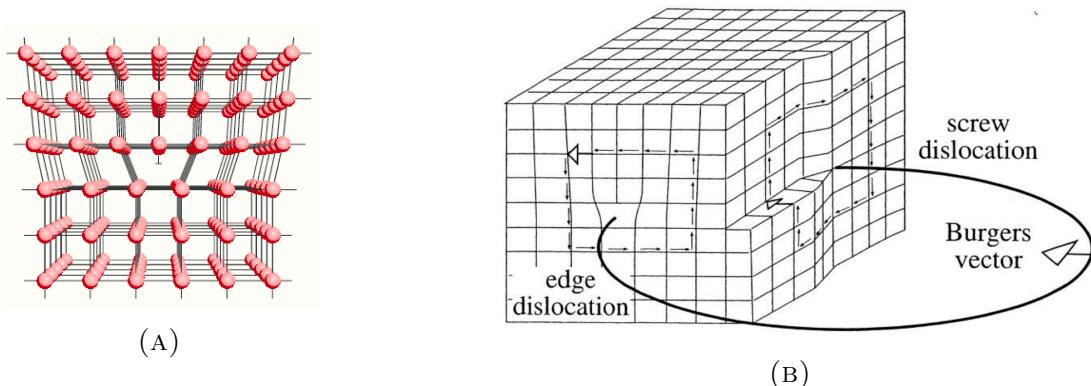


FIGURE 2.1: (A) Diagram of edge dislocation in a simple cubic lattice. The 'T' symbol denotes the dislocation line pointing into the page. The Burgers vector for this dislocation is the magnitude of one lattice spacing and points to the right. (B) Diagram of dislocation with both screw and edge dislocations in a simple cubic lattice, showing the burgers circuit—the small black arrows—and the burgers vector of the system, which is invariant along the dislocation line. The applied shear on the top half is to the left, and is to the right on the bottom half of the block. [? ?].

2.2 Dislocations

Dislocations are extended defects. They can be differentiated into two main types: edge dislocations, where a half plane of atoms is added, or removed, from the lattice, and screw dislocations, whereby a region of the lattice has slipped with regards to the perfect crystal [?]. The magnitude and direction of a given dislocation is characterised by the Burgers Vector, the resulting vector necessary to complete a whole circuit around the dislocation line compared to the corresponding circuit in a perfect crystal. This is an invariant quantity along the dislocation line. See 2.1 for a diagram of both screw and edge dislocations. A dislocation can be a combination of both of the aforementioned types and it is said to have mixed character. Generally, edge dislocations are higher in energy than screw dislocations.

Kinks are discrete steps of the dislocation line of atomic dimension that fully contained in the glide plane of the dislocation. Similarly, a jog is any step of atomic dimension of a dislocation line, but it is not in the glide plane of the dislocation. These are always of edge character as such they can result in point defect absorption and emission. Intersections of edge and screw dislocations will produce one jog per dislocation. Jogs influence the movement of dislocations and provide feedback, giving rise to non-linearity of deformation.

A Peierls potential, of the Peierls-Nabarro Model [? ?], is the periodic potential of the lattice which opposes the shear stress to move a dislocation. Peierls potential maxima and minima are along lines of high symmetry of the lattice. The stress to overcome the Peierls potential is the critical shear stress. Dislocations of mixed character will form kinks instead of their idealised lines to stay in Peierls minima, thus minimising their core energy. Majority edge character dislocations form geometrically necessary kinks of screw character. Kinks in screw dislocations can act as anchor points for movement along new glide planes (cross-slip). See 2.3 for diagram.

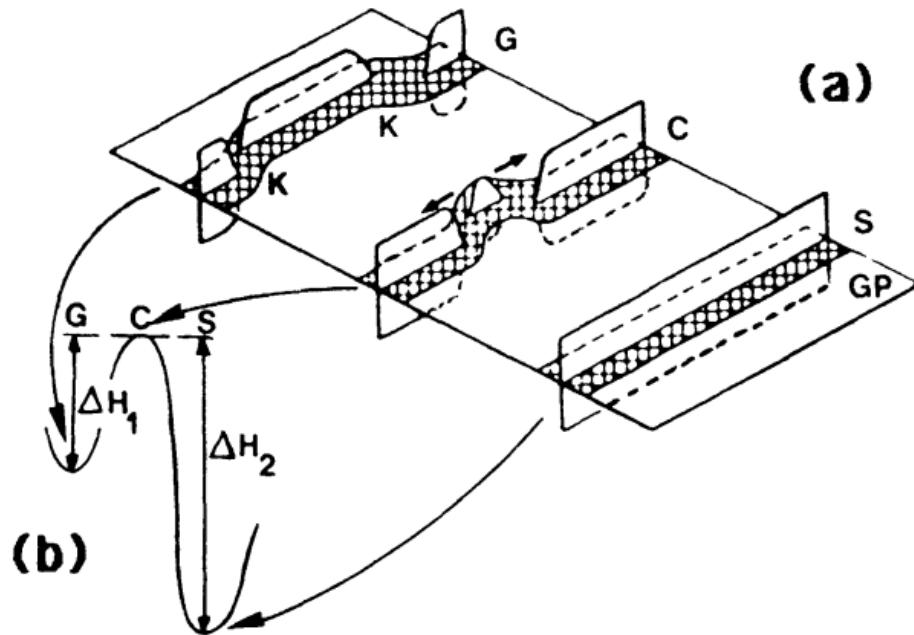


FIGURE 2.2: Mechanism of core spreading by kink migration from sessile to glissile configuration. Dislocation mainly moves in its glide plane and partly normal to it. Thermal barrier ΔH_2 is necessary to overcome before kink pair nucleation at critical position C. The kink pair separates such that the whole dislocation can move one atomic row. Associated energy barrier diagram seen in (b). [?]

Double kinks, a crossing of the dislocation line over to a neighbouring Peierls valley followed by a subsequent kink back to join the original dislocation line, can be nucleated by thermal fluctuation. At low stresses compared to the Peierls stress, these double kinks can separate from each other and traverse the rest of the dislocation line, allowing thermally activated movement of the dislocation and plastic deformation [?]. See 2.2 for diagram.

Screw dislocations are able to cross-slip onto other glide planes, as there are many glide planes parallel to the dislocation line. Stacking fault energies determine the possibility for these dislocations to cross-slip. High stacking fault energies result in a reduced separation between partial dislocations (dislocations of burgers vector not equal to a translational vector of the lattice). This promotes cross-slip, giving rise to wavy slip. Low stacking fault energies give rise to planar slip, from the greater difference in partial dislocations.

Dislocation mobility governs both the levels of stress at which a sample commences plastic flow and the rate of work-hardening [?]. Conservative dislocation motion is where a dislocation moves in its glide plane. Dislocation climb is dislocation movement not in the glide plane assisted by point defects.

2.2.1 Classical models for dislocation simulation

The Peierls-Nabarro Model [? ?], the model of Peierls potentials as described above, was the first model used for simulating dislocations in an elasto-plastic framework, but its

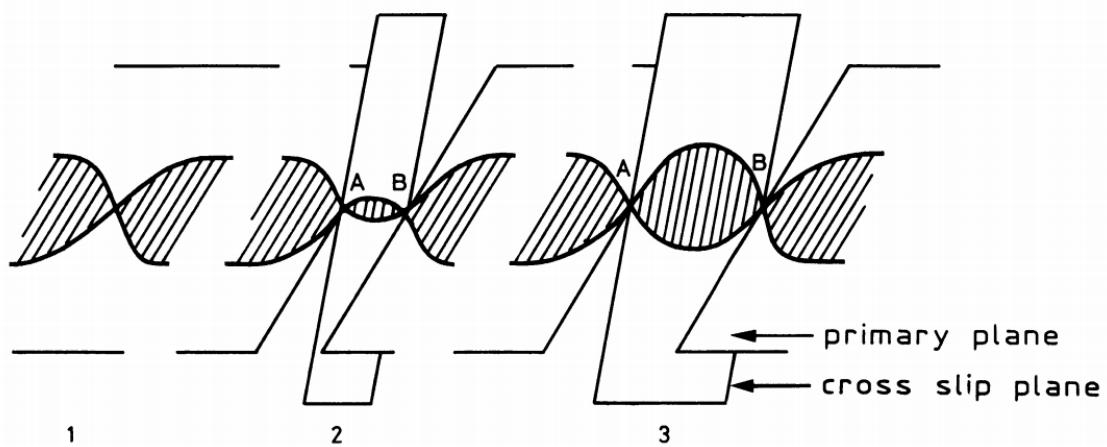


FIGURE 2.3: Figure cross-slip model, devised by Freidel and Escaig [?]. Constriction occurs at 1 and then successive stages of slipping can be seen [?]

description of a dislocation is only maintained in its glide plane [?].

Volterra dislocations, or anisotropic elasticity models, give a well described long range elastic field, calculated from the elastic Green's function [?]. Unfortunately, the core region, which determines most of the dynamics of the system, is not described as it is not in the regime of linear elasticity theory.

Dislocation density calculations and crystal plasticity finite element models do well on the larger scale of grains in the material, but cannot account for solute interaction with dislocation cores.

Interatomic potentials methods have been used to simulate dislocations, but their predictive capability has been limited [?].

2.3 Titanium Slip Systems and Plasticity

HCP titanium has 5 slip systems and 4 twinning modes. Generally, the ease at which a particular slip plane activates can be measured by its critical resolved shear stress (CRSS). A figure can be seen in 2.5. Prismatic $\langle a \rangle$ screw slip $\frac{a}{2}\langle 1\bar{2}10 \rangle\{10\bar{1}0\}$, is the easiest slip mode to activate and thus it accommodates most of the plastic deformation seen in titanium. See 2.4 for figure of slip planes.

Secondary slip systems to the $\langle a \rangle$ screw dislocation are dislocations with Burgers vector $\frac{a}{2}\langle 1\bar{2}10 \rangle$, which can glide in the basal plane, or in the first-order pyramidal plane $\{10\bar{1}\bar{1}\}$. However this basal slip is only seen in high-purity α titanium, implying that the active plane of $\langle a \rangle$ cross-slip is sensitive to oxygen content.

$\langle a \rangle$ dislocations have been shown to cross-slip on the basal plane, and the first-order pyramidal (π_1) planes; these are also secondary slip systems. Basal slip is only seen for large plastic strain. The critically-resolved shear stress of this mode is higher than that of prismatic slip, but this decreases with increasing oxygen content.

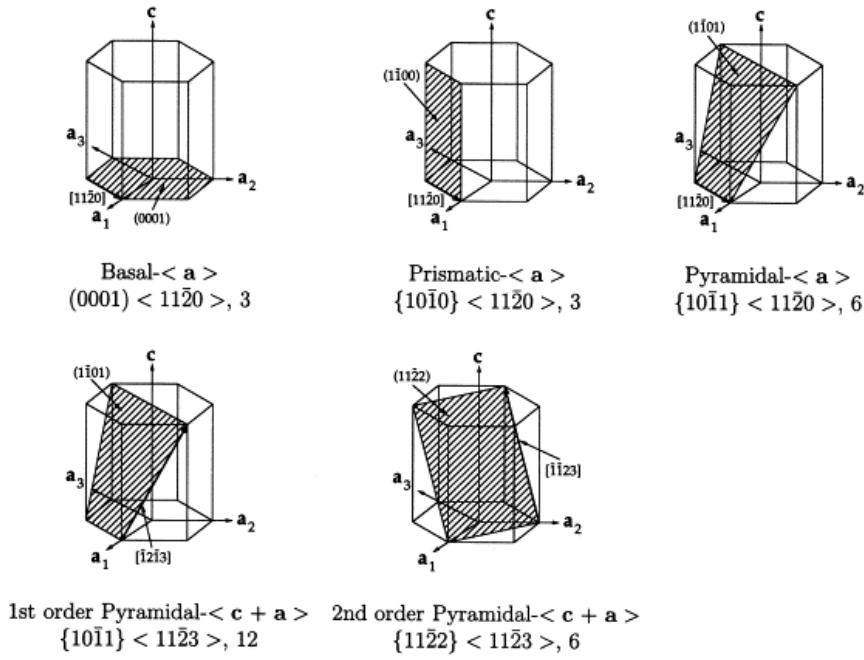


FIGURE 2.4: Figure of slip systems in titanium showing basal, prismatic and pyramidal slip. [?]

Jerky movement of $\langle a \rangle$ screw dislocations is observed at low temperatures, suggesting a 'locking-unlocking' mechanism such as the Freidel-Escaig model of cross slip by constriction [?]. See 2.3 for a diagram of the mechanism.

Edge dislocation mobility is much higher than that of a screw dislocation's at low temperature, so screw dislocations control plastic deformation. Transition of dislocation movement mechanisms, between the Friedel-Escaig model of cross-slip and the Peierls model of double kink migration, has been observed around 300K, from changes in activation volume.

Introduction of aluminium and oxygen into α -Ti promotes slip on the first and second-order pyramidal planes, instead of twinning mode activation, with strain along the $\langle c \rangle$ axis. $\langle c+a \rangle$ slip accommodates most of the strain in along the $\langle c \rangle$ axis. The slip plane that is activated from $\langle c+a \rangle$ depends on the stress direction and also on the character of the dislocation.

2.3.1 Dislocation Cores and Effect of Impurities

The core of a dislocation is controlled by its electronic interactions. These core fields are sensitive to atomic species: configurations with the same valence, for example Zr and Ti, can have the same core configuration, but one may be sessile and another may be glissile [?].

The nature of core spreading of $\langle a \rangle$ screw dislocations remains a contentious issue. There has been consensus of core spreading on the prismatic plane with a possibility of secondary spreading on basal and pyramidal planes. Many atomistic approaches relate a screw dislocation core with basal plane spreading, for high purity titanium [? ?]. Spreading in prismatic

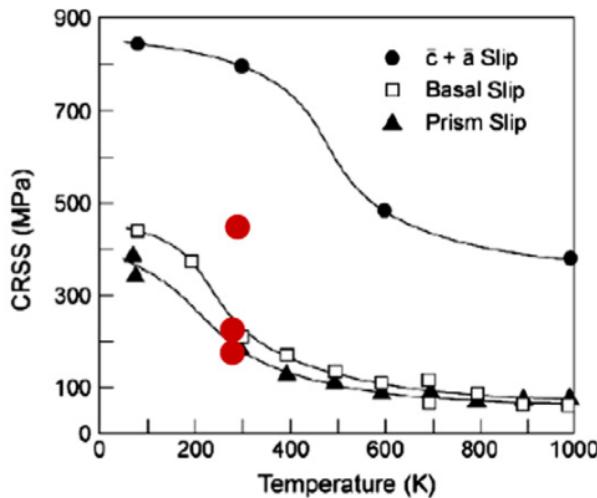


FIGURE 2.5: Diagram showing the variation in CRSS with temperature and the slip systems of HCP titanium. Red dots show data of higher purity titanium Diagram taken from [?].

and pyramidal are stable configurations, but in basal it is not for higher solute concentrations. The spread of dislocations, exhibition of non-planar dissociation, gives evidence to high-lattice friction as suggested by Chaari [?].

Constriction of secondary spread planes of sessile configurations onto glissile configurations in the prismatic plane can be characterized by the recombination energy ΔE of the Freidel-Escaig mechanism 2.3. This recombination energy increases with oxygen content [?].

Churchman back in 1954 suggested that ordering of oxygen in interstitial sites would result in interference of slip on basal and prismatic planes [?]. From empirical evidence and computational simulations, this idea has been confirmed. Atomistic simulations have shown, in high purity titanium, that basal cross slip of $\langle a \rangle$ screw dislocations are observed, even though the first-order pyramidal plane is activated with higher oxygen content [?].

The strain rate sensitivity of stress and the critical resolved shear stress of prismatic slip increases with oxygen content. Suggested hardening processes via oxygen is due to a strong repulsion of oxygen, which modifies the core dislocation structure,

Oxygen content reduces the activation volume of the $\langle c + a \rangle$ slip, as in prismatic slip, suggesting similar mechanisms for these dislocations in terms of mobility to those proposed for the mobility of $\langle a \rangle$ screws.

The nature of $\langle c + a \rangle$ core spreading planes and the actual core structure remains unknown for solute additions.

With oxygen and aluminium alloying additions, planar slip is intensified [?]. It remains to be seen if such results occur with addition of vanadium.

2.4 Computational Simulations of titanium

2.4.1 Dislocations

ab initio calculations to understand the effect of dislocation core spreading have been numerous [? ?]. In recent DFT simulations regarding zirconium by Chaari *et al.* repulsion of oxygen from the dislocation core was seen theorised to result from the destruction of the octahedral interstitial sites by the stacking fault associated with the dislocation dissociation. As a consequence of this repulsive interaction, the screw dislocation cross-slipped to restore the octahedral insertion site. This cross-slip event induced by the oxygen atom, creating two jogs on the screw dislocation. This explains experimental observations of lattice friction acting against screw dislocation glide in zirconium alloys containing oxygen. However, the dynamics of the jogs along the dislocation line and whether they will annihilate or coalesce after unpinning from oxygen atoms, has not been explored. This necessitates three dimensional simulation of dislocations in the presence of oxygen [?].

Recent calculations done by Hao *et al.*, using a plane-wave GGA DFT model, more accurately model the non-linear stresses of the of the dislocation cores while obtaining the correct lattice and elastic constants [?]. However the presence of the vaccuum resulted in numerical instabilities, as well as a discontinuity in electronic density which induced Friedel oscillations [?].

Unfortunately there is a problem with dislocation simulations in DFT. Periodic boundary conditions cause insufficient relaxation of atoms in the dislocation core for small scale simulations of a single dislocation [?]. This results in unresolved glide planes and hence, unphysical dislocation motion. Furthermore, the choice of exchange-correlation function can give vastly different results. For example, the Peierls barrier in Tantalum differs by 20% between the Perdew-Burke-Ernzerhof GGA parameterisation and LDA [?].

There have been attempts to circumvent the poor relaxation of atoms in the dislocation cores in DFT. One method is to adopt flexible boundary conditions. These include a coupling of the system to an empirical potential, or use of a lattice green function region, whereby a harmonic approximation to the displacements induced on each atom by a force acting on the atom from the origin: $u_i(\mathbf{r}) = \sum_j G_{ij}(\mathbf{r})F_j$. This enables iterative *ab initio* relaxation of the core and subsequent relaxation using the empirical potential/lattice green function, with an outer buffer region to prevent external boundary forces. This effectively simulates an infinite lattice and is a method of embedding—more methods will be described in a later section. Other solutions to this problem involve defining a dislocation dipole within a periodic boundary simulation cell such that the stresses from the dislocations cancel to first order, given proper placement of the dislocations. Peach-Kohler forces exist between dislocations, so quadropole configurations have to be adopted for minimisation of image dislocations [?].

Fortunately, DFT can provide insight into possible core structures by use of stacking faults.

2.4.2 Stacking Faults

Many *ab initio* calculations determine the generalised stacking fault energy (γ -surface) associated with the system, formulated by Vitek [?]. These can be calculated by cutting the crystal along a fault plane and the resulting half-crystals are shifted by a fault vector γ belonging to the fault plane. Atoms are then relaxed, usually in the direction perpendicular to the fault plane, to obtain the γ -surface: the 2D fault energy as a function of fault vector. Local minima of the surface suggest stable stacking faults for the decomposition of perfect dislocations into partials.

Because relaxations are allowed only perpendicularly to the fault plane, stable stacking faults may be missed on a γ -surface. An example is the second order $\langle 2112 \rangle$ pyramidal plane in HCP metals [?], where no relevant minimum is found for the splitting of $\langle c+a \rangle$ dislocations when only perpendicular relaxations are allowed. However, a minimum does appear when all atoms, except those just above and below the fault plane, are allowed to fully relax. This can be achieved using a method by Morris [?].

Legrand suggests that the preference for prismatic slip of $\langle a \rangle$ dislocations is related to an increased basal stacking fault energy, seen in transition metals with d fillings between 1.5 and 2.5 electrons. This results in directional d-covalent bonding because of partially filled d bands [?].

Calculations by Yu *et al.* confirmed that $\langle c+a \rangle$ dislocation dissociation onto $\langle a \rangle$ prismatic planes is a means by which α -titanium maintains ductility. Two stable intrinsic (vacancy composed) stacking faults were proposed to account for the lack of stable stacking fault at $\langle c+a \rangle/2$ for the $\langle c+a \rangle$ dislocation. This gave a possible dissociation of $\langle c+a \rangle$ dislocation into three partial dislocations separated by the two intrinsic faults [?].

In contrast, analysis of DFT gamma surfaces for stable stacking faults by Yin and Curtin *et al.* suggest that $\langle c+a \rangle$ dissociation into three partials is unlikely [?].

Nudged Elastic Band theory (NEB) [?] can also be used to obtain γ -lines [?]. The sensitivity of the γ -surface on the way atomic positions are relaxed is even more important in more complex crystalline structures with several atoms per primitive unit cell. This is most likely the case with with rutile and anatase TiO_2 .

2.4.3 Embedding Methods

The 'holy grail' of crystal plasticity simulation is to achieve multi-scale modelling. This type of model could naturally bring the effects of atomistic quantum interactions into macroscopic phenomena, such as, in the case of titanium, how oxygen-dislocation interactions increase the hardness of the material and how these dislocations can cause damage to surface oxides of the material.

Embedded atom methods (EAM), second moment approximations to tight binding theory—Finnis-Sinclair-like models—have been used to achieve multiscale simulation techniques. These have used flexible boundary conditions by imposition of a lattice green's function. Embedding models using this scheme have been successfully implemented by people such as Trinkle, via differing formats [?], but these could be made more accurate, in theory, by

using tight binding and larger moment approximations. This can be seen in work by Girshick where bond order potential methods were used and a stable prismatic core was observed, as in DFT [?].

Itakura also imposed a lattice greens function method on a DFT governed dislocation core in BCC Iron, but less than 150 atoms were used [?].

Other multiscale approaches have been achieved by use of molecular dynamics (MD). MD simulations of crack propagation in titanium has been shown by Sheng [?]. Atomistic crack propagation using DFT derived classical potentials in 2-D has been investigated by Kermode [?]. QM/MM (quantum mechanical and molecular dynamics) simulations of crack tips have been investigated in 2-D [?]. But, there has not yet been a multi-scale model of sufficient scale with quantum mechanical methods, that has been able to investigate dislocations and cracking in titanium, or its alloys, in 3-D.

2.5 Numerical Optimisation

The downside to tight binding is finding parameters. This is a lengthy process and requires good physical intuition. Some interesting methods have been developed with regards to obtaining tight binding parameterisations. Cawkwell *et al.* used DFT to fit tight binding models to molecules, using the numerical optimisation method of Simulated Annealing [?]. Another approach is that by Hedge and Bowen: the use of machine learning to obtain the bond integrals and pair potentials, but this was only achieved for trivial cases such as a simple *s*-orbital of copper and *s* and *p* orbitals of diamond [?]. Much success has been found with Schwefel's Genetic Algorithm, employed by Lozovoi and Paxton [?], but it is quite slow for a larger number of species.

The "Curse of Dimensionality" is the pathological affliction of large numerical optimisation problems. This arises in tight-binding as the number of dimensions vastly increases with the number of species, as self-interaction, and interaction with each other species must be defined, in general. So a numerical method that can obtain the optimum set of parameters quickly and efficiently for systems with many species would be of great use.

Evaluation of the objective function, the overall function one would want to minimise for the best model, on a given pass is time consuming. It requires calculation of empirical quantities for all species for a given set of input parameters. In general, many objective function points are necessary for convergence of parameter sets. This could be mitigated by the use of Bayesian optimisation, which 'learns' of the best parts of parameters space to sample using data from previous iterations, to find the minimum of the objective function, and hence the best parameter sets.

Chapter 3

Quantum Mechanics and Numerical Optimisation

3.1 Density Functional Theory

The many-body Schrödinger equation for a system is utterly intractable when trying to describe systems composed of a few electrons, let alone solids. The development of Density Functional Theory (DFT), with ideas developed by Hohenberg, Kohn and Sham, provided a way of finding the ground state energy for a many body system by solving a single particle Schrödinger equation [? ?]. They proposed that an N electron system can be considered to have a density dependent on space $\rho(\mathbf{r})$ which has an associated antisymmetric wavefunction $|\Psi\rangle$. The energy can then be said to be a *functional* of the density, $E[\rho(\mathbf{r})]$, where the function describing the energy depends on the density function fed into it [?]. The central statements of DFT are that all observable quantities of a stationary system can be determined solely from the ground state density and that the ground state density of the interacting system can be calculated from an effective non-interacting system. The ground state energy of the system of interest is calculated with Kohn-Sham orbitals. The resulting Kohn-Sham energies are not physical energies of the system, but if the material in question is weakly correlated, then these energies become a good approximation [?].

3.2 The Hohenberg-Kohn-Sham Functional

The definition of the HKS functional is

$$F^{HKS}[\rho] = T_s[\rho] + E_H[\rho] + E_{xc}[\rho] \quad (3.1)$$

where the use of a reference system of non-interacting electrons is used with density ρ . This defines the kinetic energy of the reference system $T_s[\rho]$. The actual kinetic energy of the functional is $T[\rho] = T_s[\rho] + \Delta T$, where ΔT is included (approximately) in the exchange-correlation energy $E_{xc}[\rho]$. As $E_{xc}[\rho]$ can be approximated rather well by the Local Density Approximation based on the energy of the Jellium model, we can write the real HKS functional

$$E^{\text{HKS}}[\rho] = \underset{\text{K-S KE}}{T_s[\rho]} + \underset{\text{Hartree}}{E_H[\rho]} + \underset{\text{Exchange-correlation}}{E_{xc}[\rho]} + \underset{\text{Electron-ion}}{E_{eZ}[\rho]} + \underset{\text{Ion-ion}}{E_{ZZ}[\rho]} \quad (3.2)$$

3.3 Non-self-consistent Tight Binding

To derive a simple non-self-consistent tight binding model, one can use the first-order Harris-Foulkes functional

$$E^{(1)}[\rho] = \underset{\text{Band Energy}}{\text{Tr} \hat{\rho} \hat{H}^{\text{in}}} - \underset{\text{Hartree}}{E_H^{\text{in}}[\rho]} - \underset{\substack{\text{Removal of } xc \\ \text{from Band Energy}}}{\text{Tr} \hat{\rho}^{\text{in}} V_{xc}} + \underset{\text{Exchange-correlation}}{E_{xc}^{\text{in}}[\rho]} + \underset{\text{Ion-ion}}{E_{ZZ}[\rho]}$$

where ⁽¹⁾ denotes it is first order and ⁱⁿ shows that it is an input charge density. Minimising this with respect to the charge density ρ is equivalent to solving the single particle Schrödinger equation:

$$\hat{H}^{\text{in}} |\psi_n\rangle = \epsilon_n |\psi_n\rangle. \quad (3.3)$$

This is easier to minimise than the HKS functional as this does not have to be solved self-consistently.

Using the basis of atomic orbitals—with the notation $|I\mu\rangle$, where the state describes the μ th orbital of the I th site (atom)—the Schrödinger equation can be rewritten as an eigenvalue problem.

$$\sum_{J\nu} H_{I\mu J\nu}^{\text{in}} C_{J\nu}^n = \epsilon_n \sum_{J\nu} S_{I\mu J\nu}^{\text{in}} C_{J\nu}^n, \quad (3.4)$$

where $S_{I\mu J\nu}^{\text{in}}$ is an element of the overlap matrix between orbitals. Using the definition of the trace we have,

$$E_{\text{band}} = \text{Tr} \hat{\rho} \hat{H}^{\text{in}} = \sum_{J\nu I\mu} \rho^{I\mu J\nu} C_{J\nu}^n H_{I\mu J\nu}^{\text{in}} \quad (3.5)$$

we can define the Bond Energy as the intersite terms,

$$E_{\text{bond}} = \sum_{J\nu I\mu, J \neq I} \rho^{I\mu J\nu} C_{J\nu}^n H_{I\mu J\nu}^{\text{in}} \quad (3.6)$$

where $\rho^{I\mu J\nu}$ is the *Bond Order*.

Tight binding theory rests upon the parametrisation of the Hamiltonian elements $H_{I\mu, J\nu}^{\text{in}}$ (Hopping integrals). These can be empirical, semi-empirical or canonical values. In 1954 Slater and Koster developed general formulae for the matrix elements between orbitals in the LCAO (Linear Combination of Atomic Orbitals) basis. These involve direction cosines l, m and n corresponding to the cosines between the coordinates x_1, x_2 and x_3 respectively, and standard bond integrals between orbitals *i.e ssσ, spσ, ppπ*, for the bonds between atoms' s orbitals, and p orbitals [?].

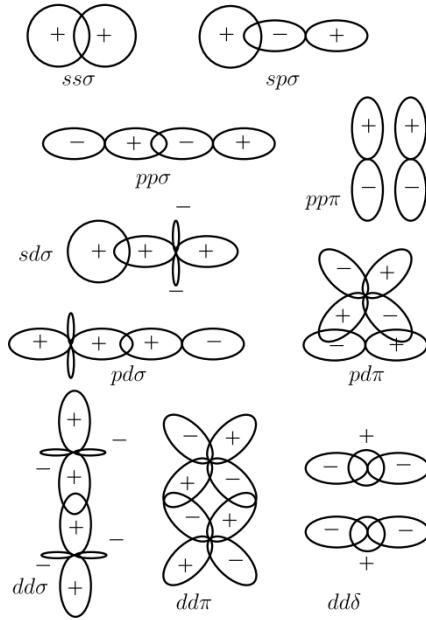


FIGURE 3.1: Figure of overlapping orbitals showing the fundamental bond integrals necessary tight binding. Using these bond integrals and directions cosines one can use the Slater-Koster table to obtain overlap energies of orbitals. [? ?]

To accurately obtain the Hamiltonian for a system we would need an infinite basis set on each atom, but in this LCAO basis, we can get away with what is essentially a minimal set, giving good enough results. A minimal set is necessary for fast computation times.

In a system of n types of bonds between atoms where each atom has m orbitals considered, we have a Hamiltonian matrix that is $(n \times m) \times (n \times m)$. If the system is assumed to be periodic, we use the n bond types apparent in the unit cell of the lattice. Using Bloch's Theorem, phase factors of the form $e^{i\mathbf{k} \cdot \mathbf{d}}$ are introduced, that modulate the bond integral from the Slater-Koster table for each of the bond types in the unit cell, where \mathbf{k} is the momentum and \mathbf{d} is the distance between the centres of the bonded atoms. Most variants of tight binding call for atoms which interact only with their nearest neighbours. The resulting Hamiltonian matrix has many elements of value zero, due to the orbital basis used, decreasing the amount of computational time necessary for diagonalisation to find the energies of each of the orbitals.

3.3.1 Distance Dependence of Bond Integrals

There must be a range cutoff to the bond integrals, to speed up computation time and, in the case of most theories, to keep the interactions between nearest neighbours. Initially decaying exponentials (Born-Mayer) were used but there have been arguments to suggest that the cutoffs can be of the form R^{-n} where n is an integer (usually between 2-5).

3.3.2 Tight Binding Band Model

This model assumes a pair potential for the repulsive terms in

$$E_{\text{band}}^{\text{rep}} = E_{\text{xc}}^{\text{in}} - \rho_{\text{xc}}^{\text{in}} V^{\text{in}} - E_H^{\text{in}} + E_{ZZ}.$$

This gives the band energy to be

$$E_{\text{band}}^{(1)} = \text{Tr} \hat{\rho} \hat{H}^{\text{in}} + \frac{1}{2} \sum_{I \neq J} V_{IJ}. \quad (3.7)$$

The binding energy can then be expressed as

$$E_B = E_{\text{band}} + E^{\text{rep}}, \quad (3.8)$$

where E^{rep} can be approximated by a repulsive pair potential.

There is no strong justification for this assumption. This model has been superseded by the Tight Binding Bond Model (TBBM).

3.3.3 Tight Binding Bond Model

In a different fashion to the Tight Binding Band Model, TBBM assumes a pair potential for the repulsive energies of the functional, but it is derived from the *Cohesive Energy*, not the total energy. This reformulation makes sure that the cohesive energy is invariant to an applied potential.

$$E_{\text{binding}} = -N_a E_{\text{cohesive}} = E^{\text{total}}(\text{Condensed Matter}) - E^{\text{total}}(\text{Free Atoms}) \quad (3.9)$$

The energy of free atoms, neglecting spin polarisation is:

$$E^{\text{Total}}(\text{Free}) = \sum_I [\text{Tr} \rho^I \hat{H}^I - \rho^I V_{\text{eff}}^I + \rho^I V_{\text{ext}}^I + E_{\text{xc}}^I + \frac{1}{2} \rho^I V_H^I] \quad (3.10)$$

where

$$\rho^{\text{in}}(\mathbf{r}) = \sum_I \rho^I(\mathbf{r} - \mathbf{R}_I)$$

and similarly for $V_H^{\text{in}}(\mathbf{r})$ and $V_{\text{ext}}^{\text{in}}(\mathbf{r})$. As the exchange-correlation potential is *not* a linear function of the density, we cannot split $V_{\text{eff}}^{\text{in}}$ into atomic contributions.

The binding energy can then be expressed as

$$\begin{aligned} E_B^{(1)} = & E^{\text{bond}} + \sum_{I\mu\nu} \{ \rho^{I\nu I\mu} H_{I\nu I\mu}^{\text{in}} - \rho^{I\nu I\mu} H_{I\nu I\mu}^I \} - \sum_I \rho^I (V_{\text{eff}}^{\text{in}} - V_{\text{eff}}^I) \\ & E_{\text{xc}}^{\text{in}} - \sum_I E_{\text{xc}}^I + \sum_{I \neq J} \rho^I V_{\text{ext}}^J + \frac{1}{2} \sum_{I \neq J} \rho^I V_H^I + E_{ZZ}, \end{aligned}$$

where we have written the Hartree energy in a pairwise way

$$\begin{aligned} E_H^{\text{in}} &= \frac{1}{2} \rho^{\text{in}} V_H^{\text{in}} \\ &= \frac{1}{2} \sum_I \rho^I V_H^I + \frac{1}{2} \sum_{I \neq J} \rho^I V_H^I \end{aligned}$$

One can then define a promotion energy E^{prom} which corresponds to an energy promotion for electrons going into higher orbitals in becoming part of a solid, compared to that of the free atom. The binding energy can then be expressed as

$$E_B^{(1)} = E^{\text{cov}} + E^{\text{prom}} + E^{\text{rep}}, \quad (3.11)$$

where E^{rep} is generally a repulsive pairwise potential.

3.3.4 Local Charge Neutrality

An elementary form of self-consistency arises in this model due to the constraint that all atoms must be charge neutral. This involves modification to the on-site Hamiltonian terms $H_{I\mu I\mu}^{\text{in}}$ atom by atom such that the sum over expansion coefficients of the density matrix $\sum_\mu \rho^{I\mu I\mu}$ is constant.

3.4 Self-consistent Tight Binding

3.4.1 Self-consistent Charge Transfer Model

The goal of self-consistency is to include the second order term of the HKS functional, namely, in the \mathbf{r} representation,

$$E_2 = \frac{1}{2} \int C_{\text{in}}(\mathbf{r}, \mathbf{r}') \delta\rho(\mathbf{r}) \rho(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (3.12)$$

where the kernel C_{in} is

$$C_{\text{in}} = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{\text{xc}}[\rho]}{\delta \rho \delta \rho'} \Big|_{\rho^{\text{in}}}, \quad (3.13)$$

where $\delta\rho$ is an induced charge.

Assuming that the induced charges are from the difference in Mulliken charges

$$\delta q_I = q_I - q_I^{\text{in}},$$

where

$$q_I = \rho_I + \frac{1}{2} \sum_J \rho_{IJ} \quad (3.14)$$

then

$$E_2 = \frac{1}{2} \sum_I U_I \delta q_I^2 + \frac{1}{2} \sum_{IJ, I \neq J} U_{IJ} \delta q_I \delta q_J, \quad (3.15)$$

where the on-site energy changes are described by the *Hubbard* U : U_I , which is deemed to be constant for a given species. The intersite terms are just Coulomb interactions,

$$U_{IJ} = \frac{1}{R_{IJ}} \quad (3.16)$$

which are generally damped at short range due to overlapping atomic charge distributions. Consistency with the LDA approximation is apparent with just using the first term of 3.13 when $\mathbf{r} \neq \mathbf{r}'$.

This gives the Hamiltonian

$$\begin{aligned} H_{I\mu J\nu} &= H_{I\mu J\nu}^{\text{in}} + (U_I \delta q_I + \sum_{I' \neq I'} U_{II'} \delta q_{I'}) \delta_{IJ} \delta_{\mu\nu} \\ &\quad \frac{1}{2} (U_I \delta q_I + U_J \delta q_J + \sum_{I' \neq I} U_{II'} \delta q_{I'} + \sum_{I' \neq J} U_{JI'} \delta q_{I'}) O_{I\mu J\nu}, \end{aligned}$$

where $O_{I\mu J\nu}$ is an element of the decomposition of the overlap matrix \mathbf{S} : $\mathbf{S} = \mathbb{1} + \mathbf{O}$

3.4.2 Polarisable-ion tight binding

One can extend this model to include electrostatic ion polarisability and crystal field splitting by the formulation of self-consistent polarisable-ion tight binding [? ?]. This extension of tight binding is integral to the development of a model for transition metal oxides, as Mulliken charges always underestimate the valence charges, rendering the previous model inadequate to describe such systems.

This model arises due to an expansion of multipole moment operators

$$\hat{Q}_{\mathbf{R}L} = \hat{r}^l Y_L(\hat{\mathbf{r}}), \quad (3.17)$$

where $Y_L(\hat{\mathbf{r}})$ is a spherical harmonic, l is the angular momentum and L is a composite angular momentum index $L = \{lm\}$. Using the eigenvectors of tight binding, $|\mathbf{R}L\rangle$, the matrix element is

$$\langle \mathbf{R}L' | \hat{Q}_{\mathbf{R}L} | \mathbf{R}L'' \rangle = \Delta_{l'l''l} C_{L'L''L}, \quad (3.18)$$

where $C_{L'L''L}$ are Gaunt coefficients and $\Delta_{l'l''l}$ becomes a new parameter in the model. These are off-diagonal, on-site terms, which are usually zero.

3.5 Bond Order Potentials

The bond order potential method is a faster, but less accurate method, of calculating the energies of atoms. It does this by constructing the density of states from a number of moments, which give information of how energies interact between sites: information of

electronic structure is obtained from p^{th} moment hopping around paths of length p between sites. From this the Green's function is constructed from which energies and forces can be calculated.

3.5.1 Moment Theorem

This theorem shows that the local density of states on a site i and an orbital α , $n_{i\alpha}(E)$, can be constructed from a number of moments $\mu_{i\alpha}^{(p)}$.

$$n_{i\alpha}(E) = \sum_{p=0}^{\infty} \frac{\mu_{i\alpha}^{(p)}}{E^{p+1}} \quad (3.19)$$

where

$$\mu_{i\alpha}^{(p)} = \int E^p n_{i\alpha}(E) dE = \langle i\alpha | \hat{H}^p | i\alpha \rangle. \quad (3.20)$$

3.5.2 Lanczos Algorithm

This algorithm casts the Hamiltonian in terms of a tri-diagonal matrix. The matrix elements are defined as

$$\langle U_m | \hat{H} | U_n \rangle = \begin{cases} a_n & \text{if } m = n \\ b_n & \text{if } m = n - 1 \\ b_{n+1} & \text{if } m = n + 1 \\ 0, & \text{otherwise} \end{cases}$$

where $|U_n\rangle$ are all orthonormal: $\langle U_m | U_n \rangle = \delta_{mn}$.

From this we can get the recursion relation

$$H |U_n\rangle = a_n |U_n\rangle + b_n |U_{n-1}\rangle + b_{n+1} |U_{n+1}\rangle \quad (3.21)$$

We can start from $|U_0\rangle$, for $n = 0$ and

$$a_0 = \langle U_0 | \hat{H} | U_0 \rangle.$$

and consequently obtain the other terms from the recursion relation.

3.5.3 The Green's Function

$$G(z) = \frac{1}{z - H}. \quad (3.22)$$

When inverting $(z - \hat{H})$, which is tridiagonal in the $|U_n\rangle$ basis, by determinant expansion method the matrix element of the Green's function is

$$G_{00}(z) = \{[z - \hat{H}]^{-1}\}_{00} \quad (3.23)$$

$$= \frac{1}{z - a_0 - \frac{b_1^2}{z - a_1 - \frac{b_2^2}{\ddots}}} \quad (3.24)$$

So to obtain a diagonal element of Green's function for an arbitrary state $|\psi\rangle$, is simple. Let $|U_0\rangle = |\psi\rangle$, calculate a_n and b_n and then insert.

From the Green's function, we can directly obtain the bond order and the density matrix as

$$\Theta_{i\alpha,j\beta} = 2\rho_{i\alpha,j\beta} = 2 \sum_{n\text{occ}} \langle i\alpha|n\rangle \langle n|j\beta\rangle \quad (3.25)$$

$$= 2 \int_{-\infty}^{E_F} dE \langle i\alpha| \delta(E - \hat{H}) |j\beta\rangle \quad (3.26)$$

and using the following definition of a delta function

$$\delta(x) = -\frac{1}{\pi} \lim_{\eta \rightarrow 1} \text{Im}\{[x + i\eta]^{-1}\}, \quad (3.27)$$

then

$$\Theta_{i\alpha,j\beta} = 2\rho_{i\alpha,j\beta} \quad (3.28)$$

$$= -\frac{2}{\pi} \lim_{\eta \rightarrow 1} \text{Im}\left\{ \int_{-\infty}^{E_F} dE [G_{++}(E - \hat{H} + i\eta) - G_{--}(E - \hat{H} + i\eta)] \right\}, \quad (3.29)$$

where

$$G_{i\alpha,j\beta}(z) = \frac{1}{2} [G_{++}(z) - G_{--}(z)] \quad (3.30)$$

Using $|i\alpha\rangle$ as the starting state for the Lanczos algorithm, we can get moments in terms of coefficients:

$$\mu^{(0)} = 1 \quad (3.31)$$

$$\mu^{(1)} = a_0 \quad (3.32)$$

$$\mu^{(2)} = a_0^2 + b_0^2 \quad (3.33)$$

$$\mu^{(3)} = a_0^3 + 2a_0b_1^2 + a_1b_1^2 \quad (3.34)$$

$$\vdots \quad (3.35)$$

and so on for higher order moments.

All coefficients are known until N . So must have cut off at coefficients of a_∞ and b_∞ . Can have a terminator $t(z)$,

$$t(z) = \frac{1}{z - a_\infty - \frac{b_\infty^2}{z - a_\infty - \frac{b_\infty^2}{\ddots}}} \quad (3.36)$$

$$= \frac{1}{z - a_\infty - b_\infty^2 t(z)} \quad (3.37)$$

$$\implies t(z) = \frac{1}{b_\infty} \left[\left(\frac{z - a_\infty}{2b_\infty} \right) - i \sqrt{1 - \left(\frac{z - a_\infty}{2b_\infty} \right)^2} \right], \quad (3.38)$$

which is the Turchis square-root terminator.

This is limited to metallic states as electronic density of states can be represented by a single continuous band. With semiconductors with band gap, it will result in poor convergence by oscillatory behaviour of recursion coefficients [?].

Condensing the many body expansion of the bond order to a few lines, we essentially have a bond label matrix where $\Lambda_{j\beta,i\alpha}$ that satisfies $\sum_{i\alpha} \Lambda_{i\alpha,i\alpha} = 1$. From this bond order matrix, we can obtain the full Green's function [?].

Local charge neutrality is necessary as a form of self-consistency within BOP.

3.6 Embedding: tight binding and BOP

The quick computational efficiency of BOP and the more accurate, but more computationally expensive, tight binding, can be used together to efficiently model dislocations in materials, by using embedding [? ?]. This enhances the scaling, allowing for large scale simulations.

To model a dislocation system there are two main methods. Both involve the definition of a dislocation core, the region within which the stress models of dislocations are non-linear. Tight binding is used in this region, as an accurate method is crucial for determination of dislocation slip planes and analysis of solute-core interaction. For the surrounding bulk of the material, BOP is used for efficiency. Inert regions of atoms on the boundaries, as seen in 3.2, enable emulation of an infinite medium, as the Green's function is infinitely ranged, such that bulk properties of the lattice are more accurately replicated. The advantage of this is that the size of the system is large enough such that one is able to forgo periodic boundary conditions, or have to necessitate dislocation dipoles within the simulation cell.

There can be differing formulations of this embedding scheme, as seen in 3.2. Analysis of the effect of oxygen on core structure can be achieved by an embedding scheme of 3.2(a), and the mathematics behind such a method is described below. The method by which one can investigate SCC in titanium (alloys) is given by 3.2(b).

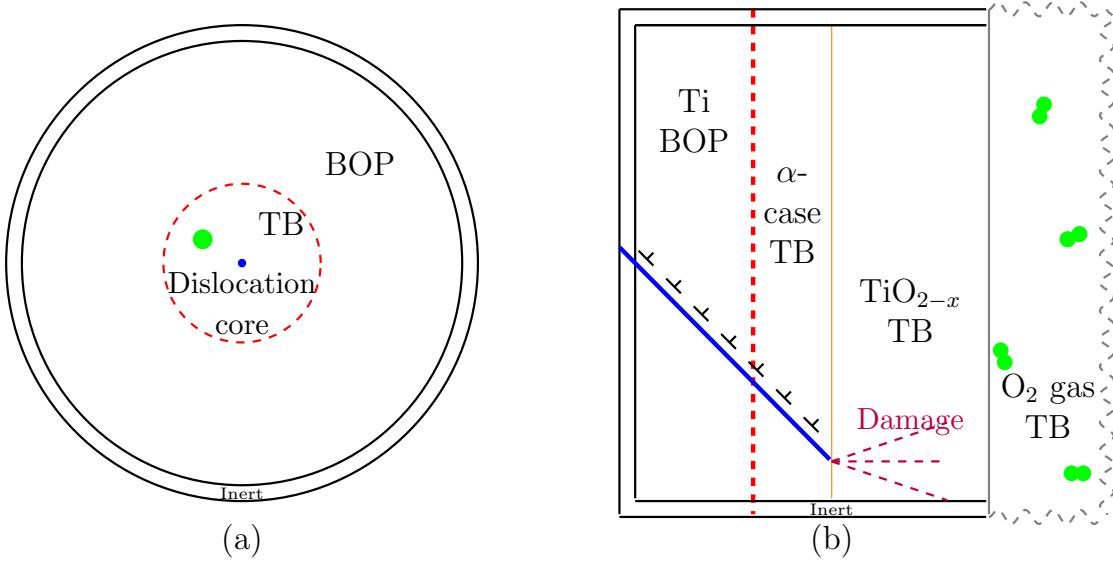


FIGURE 3.2: (a) oxygen interaction with dislocation core. Tight binding (TB) governs non-linearly stressed core while BOP atoms simulate the rest of the bulk material. (b) Dislocation pile-up at $\text{Ti}/\alpha\text{-case}/\text{TiO}_{2-x}$ interface with partial pressure O_2 environment.

3.6.1 Block-matrix Solution

Within embedding, one can see that the Green's matrix is divided up into a larger BOP region, a correlating region between BOP and Tight binding, and a smaller tight binding region. See 3.3. we can formulate solution which is predicated on the fast computation of G_{B0} , the green's matrix of the BOP region alone, due to the inherent computational speed of BOP, and G_{T0} , due to the smaller defined size of the tight binding system.

The diagonalisation of a 2×2 matrix can be written as

$$a^{-1} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^{-1} \quad (3.39)$$

$$= \begin{bmatrix} a_{11}^{-1} + A C A^\dagger & -A C \\ -(A C)^\dagger & C \end{bmatrix}, \quad (3.40)$$

where a_{ij} are matrix subblocks and where A , B and C are:

$$\begin{aligned} A &= a_{11}^{-1} a_{12} \\ C &= (a_{11}^{-1} - a_{12}^\dagger A)^{-1} \\ &\equiv (\mathbb{1} - a_{22}^{-1} a_{12}^\dagger A)^{-1} a_{22}^{-1}. \end{aligned}$$

Changing the inversion to suit the asymmetric matrix $(\mathbb{1} - G_0 \Delta H)$ requires more storage. We can relax the dependence on G_{T0} by allowing Q_{T0} to be used directly instead:

$$G(z) = \begin{bmatrix} Q_{B0}(z) & -H_{BT} \\ -H_{BT}^\dagger & Q_{T0}(z) \end{bmatrix}^{-1}, \quad (3.41)$$

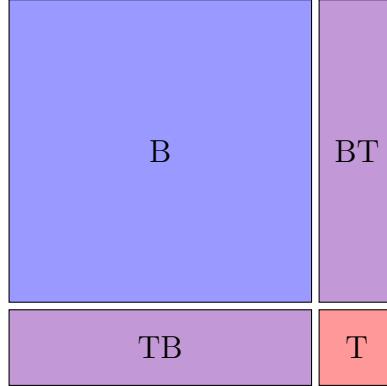


FIGURE 3.3: Diagram depicting the sub-blocks of the matrices used to calculate the Green's matrix, where B and T correspond to BOP and tight binding non-interacting Hamiltonians, and TB/BT correspond to the correlating blocks. T sub-block is small as it only contains the core of the dislocation, whereas B contains the rest of the system.

The expression for C can then be written as,

$$C : \quad G_T = (Q_{T0} - H_{BT}^\dagger G_{B0} H_{BT})^{-1}, \quad (3.42)$$

revealing the energy dependent embedding potential for the cluster region as:

$$\begin{aligned} Q_T &= Q_{T0} - H_{BT}^\dagger G_{B0} H_{BT} \\ H_T &= H_{T0} - H_{BT}^\dagger G_{B0} H_{BT} \\ &= H_{T0} + \Sigma(z). \end{aligned}$$

The matrix products forming A and ACA^\dagger go as $\mathcal{O}(N_B^2)$ as $N_B \gg N_T$. The A product corresponds to $G_B H_{BT}$ and it can be sped up significantly by exploiting the sparsity of H_{BT} . If the system is composed of two semi-infinite slabs, as in and an appropriate permutation of the whole matrix is taken, the non-zero values of H_{BT} would be concentrated in its lower left corner. The product ACA^\dagger , is much harder to optimise.

The energy dependent embedding potential requires full inversion or diagonalisation of the cluster matrix for every evaluation of energy z during the Fermi level search, and at every integration point. This is the cost for having a large system size.

For the interface of titanium (alloy) and titanium dioxide, 3.2(b), the formulation of Green's matrix may have to appear in a different way as seen for a planar defect in [?], but the underlying philosophy to hasten computation is the same.

3.7 Elastic constants

Elastic constants can be found using specific distortion matrices applied to the lattice in question. The energy can then be expanded as a function of volume and the distortion

parameter α ,

$$E(V, \alpha) = E(V_0, \alpha) + V_0 \left(\sum_i \tau_i \alpha_i \xi_i + \frac{1}{2} \sum_{i,j} C_{ij} \alpha_i \xi_i \alpha_j \xi_j \right), \quad (3.43)$$

where i is in Voigt Notation—which replaces xx by 1, yy by 2, zz by 3, xy (and yx) by 6, xz (and zx) by 5, and yz (and zy) by 4. The factor ξ_i takes the value 1 if the Voigt index is 1, 2, or 3 and the value 2 if the Voigt index is 4, 5, or 6. τ_i is an element of the stress tensor, α_i is from the symmetric distortion matrix \mathbf{e} and C_{ij} are the elastic constants [?].

Simultaneous equations are constructed from different strains, for example, say we have the lattice vectors defined as

$$a_1 = \left(\frac{\sqrt{3}}{2}, -\frac{1}{2}, 0 \right) \quad (3.44)$$

$$a_2 = (0, 1, 0) \quad (3.45)$$

$$a_3 = \left(0, 0, \frac{c}{a} \right), \quad (3.46)$$

we will have the matrix \mathbf{R} as

$$\mathbf{R} = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{c}{a} \end{bmatrix} \quad (3.47)$$

so the distortion of the lattice is given by multiplying \mathbf{R} by \mathbf{e} , which is given by

$$\mathbf{e} = \begin{bmatrix} 1 + \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & 1 + \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & 1 + \alpha_{zz} \end{bmatrix}. \quad (3.48)$$

Using the distortion matrix

$$\mathbf{e} = \begin{bmatrix} 1 + \alpha & 0 & 0 \\ 0 & 1 + \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3.49)$$

which changes the size of the basal plane while keeping the z-axis constant, gives

$$E(V, \alpha) = E(V_0, \alpha) + V_0[(\tau_1 + \tau_2)\alpha + (C_{11} + C_{12})\alpha^2]. \quad (3.50)$$

Using the matrix

$$\mathbf{e} = \begin{bmatrix} 1 + \alpha & 0 & 0 \\ 0 & 1 - \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (3.51)$$

which increases the x axis and decreases the y axis, with an equal amount, while keeping the z-axis constant, gives

$$E(V, \alpha) = E(V_0, \alpha) + V_0[(\tau_1 - \tau_2)\alpha + (C_{11} - C_{12})\alpha^2], \quad (3.52)$$

these equations can be solved to find C_{11} and C_{12} . Similarly, using other strain matrices will give combinations of other elastic constants which can subsequently be solved for.

This method is equivalent to expanding the energy per unit volume, W , as a function of α . From this expansion of W , we can relate the elastic constants to the curvature of the W with respect to alpha

$$\frac{1}{2} \sum_{i,j} \xi_i C_{ij} \xi_j = \frac{1}{2} \alpha^2 \frac{\partial^2 W}{\partial \alpha^2}. \quad (3.53)$$

One use a central difference scheme, or fit a polynomial (usually of fifth order), to the $W(\alpha)$ curve, calculate the curvature at $\alpha = 0$, and then proceed as before: use different strains to obtain simultaneous equations for the elastic constants in terms of their curvature and solve the resulting system of linear equations.

3.8 Bayesian Optimisation

Bayesian optimisation is a method for finding optimal parameters for a 'Black-Box' function. Using this method optimal parameters for a tight binding model (i.e. the bond integrals, pair potential etc.) can be found. This would, in theory, decrease the computational time required to fit these parameters compared to other algorithms, as Bayesian techniques do not discard tested data: correlations are made between input and output data to suggest more optimal parameter sets. The basis of this method, as the name suggests, rests on Bayes' Theorem:

$$\frac{P(A|B)}{\text{Posterior}} = \frac{\frac{P(B|A)P(A)}{\text{Likelihood Prior}}}{\frac{P(B)}{\text{Evidence}}}. \quad (3.54)$$

Bayesian optimization algorithms consist of a *prior* model, from which the *posterior* distribution can be obtained (with use of the *likelihood* and *evidence*) to predict the output value of the function at any arbitrary input point, and selection criteria to determine which point to be evaluated next. The use of a prior distribution over the input variables imposes constraints on the input parameters.

3.8.1 Bayesian Linear Regression

The essential objective of regression is to obtain parameters of a function, which can reconstruct, to the best possible degree, a target function from a limited set of data points.

Bayesian linear regression does this by assuming that the function of interest, $y(\mathbf{x})$ can be reconstructed using a linear combination of basis functions ϕ_i and parameters that determine the weight of each of those functions w_i . In general, it is assumed that this function of interest is smeared out by some form of noise such that the actual *target* points that are input into the regression are

$$t(\mathbf{x}) = y(\mathbf{x}) + \epsilon, \quad (3.55)$$

where ϵ is Gaussian noise.

So, given M basis functions,

$$y(\mathbf{x}, \mathbf{w}) = \sum_{j=0}^{M-1} w_j \phi_j(\mathbf{x}) \quad (3.56)$$

$$= \mathbf{w}^T \boldsymbol{\phi}, \quad (3.57)$$

where $\boldsymbol{\phi} = (\phi_1, \phi_2, \dots, \phi_{M-1})$, and $\mathbf{w} = (w_1, w_2, \dots, w_{M-1})$.

Polynomial basis functions have the form

$$\phi_j = x^j \quad (3.58)$$

One limitation of polynomial basis functions is that they are global functions of the input variable, so a change in one region of input space affect all other regions [?].

'Gaussian' basis functions have the form

$$\phi_j = \exp\left\{-\frac{x - \mu_j}{s^2}\right\}, \quad (3.59)$$

where μ_j is the position of the basis function in the input space and s^2 is the width of the basis function. These are separate parameters and can be adjusted for each function.

In general, one can choose almost any function for the basis functions. The linear in this name means that the reconstructed function is a linear combination of the parameters \mathbf{w} .

To fit functions, we can minimise a sum-of-squares error function. It can be proved that this sum-of-squares error function is the maximum-likelihood solution given Gaussian noise.

In the case of this Gaussian conditional distribution, we can write that the expectation of $t|x$

$$\mathbb{E}[t|\mathbf{x}] = \int t p(t|\mathbf{x}) dt = y(\mathbf{x}, \mathbf{w}) \quad (3.60)$$

Assuming that the noise ϵ on the data points $y(\mathbf{x})$ is Gaussian, we can write the predictive probability distribution of obtaining the input points $\mathbf{t} = (t_1, t_2, \dots, t_N)$ given inputs $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$, \mathbf{w} and β , (precision or inverse variance) as

$$p(\mathbf{t}|\mathbf{X}, \mathbf{w}, \beta) = \prod_{n=0}^N \mathcal{N}(t_n | \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_n), \beta^{-1}). \quad (3.61)$$

This is the posterior distribution of obtaining a given output from the previous given inputs. One can maximise this, using Maximum Likelihood (ML) and obtain that the most likely parameters are

$$\mathbf{w}_{\text{ML}} = (\boldsymbol{\Phi}^T \boldsymbol{\Phi})^{-1} \boldsymbol{\Phi}^T \mathbf{t} \quad (3.62)$$

where the Design matrix Φ is defined as

$$\Phi(\mathbf{x}) = \begin{bmatrix} \phi_0(x_1) & \phi_1(x_1) & \cdots & \phi_{M-1}(x_1) \\ \phi_0(x_2) & \phi_1(x_2) & \cdots & \phi_{M-1}(x_2) \\ \vdots & \vdots & \cdots & \vdots \\ \phi_0(x_N) & \phi_1(x_N) & \cdots & \phi_{M-1}(x_N) \end{bmatrix}. \quad (3.63)$$

This was the minimisation of a simple least-squares error function, but in reality this can lead to incredibly large values for parameters if bases are not chosen well enough for the problem, for example if there is a very large degree M for polynomial fitting. To counteract this, one can define a regularised error function

$$E_{\text{reg}} = \sum_{n=0}^N \{t_n - \mathbf{w}^T \phi_n(\mathbf{x})\}^2 - \frac{\lambda}{2} \mathbf{w}^T \mathbf{w}. \quad (3.64)$$

As can be clearly seen, this introduces penalties to the error function if parameter values are too large.

Incorporating this into Bayesian linear regression, we have a closed form solution for the parameters by maximising the likelihood:

$$\mathbf{w}_{\text{ML}} = (\lambda \mathbb{1} - \Phi^T \Phi)^{-1} \Phi^T \mathbf{t}. \quad (3.65)$$

For a fully Bayesian treatment which entertains the idea of constraints on optimisation, we need a prior over parameters which is given by

$$p(\mathbf{w}) = \mathcal{N}(\mathbf{w} | \mathbf{m}_0, \mathbf{S}_0) \quad (3.66)$$

from this, the posterior can be found to be

$$p(\mathbf{w} | \mathbf{t}) = \mathcal{N}(\mathbf{w} | \mathbf{m}_N, \mathbf{S}_N), \quad (3.67)$$

where

$$\mathbf{m}_N = \mathbf{S}_N (\mathbf{S}_0^{-1} \mathbf{m}_0 + \beta \Phi^T \mathbf{t}) \quad (3.68)$$

$$\mathbf{S}_N^{-1} = \mathbf{S}_0^{-1} + \beta \Phi^T \Phi, \quad (3.69)$$

where \mathbf{m}_0 , \mathbf{S}_0 and \mathbf{m}_N , \mathbf{S}_N are the mean vectors and covariance matrices for prior and posterior distributions respectively. In a sequential approach, the posterior mean and covariance matrix become the prior of the next iteration. For each input vector used, the resulting prior distribution over parameters for the next iteration becomes tighter; so for a large number of input vectors, there will be convergence to one point in parameter space.

3.8.2 Gaussian Processes for Regression

The probability distribution of a function $y(\mathbf{x})$ is a Gaussian process if for any finite selection of points $\mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \dots, \mathbf{x}^{(N)}$, the density $P(y(\mathbf{x}^{(1)}), y(\mathbf{x}^{(2)}), \dots, y(\mathbf{x}^{(N)}))$ is a Gaussian.

Taking into account the noise on the target variables

$$t_n = y_n + \epsilon_n,$$

where the noise variable is chosen independently for each observation n . We will consider noise distributions which have a Gaussian distribution such that

$$p(t_n|y_n) = \mathcal{N}(t_n|y_n, \beta^{-1}) \quad (3.70)$$

The beta precision parameter is independent for each data point so we have a joint distribution of the target values

$$\mathbf{t} = (t_1, \dots, t_N)^T$$

which are conditioned on the values of,

$$\mathbf{y} = (y_1, \dots, y_N)^T$$

given by an isotropic Gaussian of the form

$$p(\mathbf{t}|\mathbf{y}) = \mathcal{N}(\mathbf{t}|\mathbf{y}, \beta^{-1}\mathbf{I}_N), \quad (3.71)$$

where \mathbf{I}_N is an $N \times N$ unit matrix.

Due to the definition of a Gaussian process we have the marginal distribution $p(\mathbf{y})$ which is given by a Gaussian with a mean of zero and a covariance defined by the Gram matrix \mathbf{K} , such that

$$p(\mathbf{y}) = \mathcal{N}(\mathbf{y}|\mathbf{0}, \mathbf{K}). \quad (3.72)$$

This means that the kernel function that determines \mathbf{K} is chosen to express the property that for points \mathbf{x}_n and \mathbf{x}_m that are similar, the corresponding values of $y(\mathbf{x}_n)$ and $y(\mathbf{x}_m)$ will be more strongly correlated than for more dissimilar points.

A parametric kernel has the form

$$k(\mathbf{x}_n, \mathbf{x}_m) = \theta_0 \exp\left\{-\frac{\theta_1}{2} \|\mathbf{x}_n - \mathbf{x}_m\|\right\} + \theta_2 + \theta_3 \mathbf{x}_n^T \mathbf{x}_m, \quad (3.73)$$

where θ_i are hyperparameters.

To find the marginal distribution $p(\mathbf{t})$, conditioned on the input vectors $\mathbf{x}_1 \dots \mathbf{x}_n$, then we need to integrate over \mathbf{y} . We find that the the marginal distribution of \mathbf{t} is given by,

$$p(\mathbf{t}) = \int p(\mathbf{t}|\mathbf{y})p(\mathbf{y})d\mathbf{y} = \mathcal{N}(\mathbf{t}|\mathbf{0}, \mathbf{C}), \quad (3.74)$$

where the covariance matrix \mathbf{C} has elements

$$C(\mathbf{x}_n, \mathbf{x}_m) = k(\mathbf{x}_n, \mathbf{x}_m) + \beta^{-1} \delta_{nm}. \quad (3.75)$$

To get the predictive distribution $p(t_{N+1}|\mathbf{t}_N)$. We need to find the conditional distribution first $p(t_{N+1}|\mathbf{t})$, so we write down the joint distribution $p(\mathbf{t}_{N+1})$ where $\mathbf{t}_{N+1} = (t_1, \dots, t_N, t_{N+1})^T$. So the joint distribution is

$$p(\mathbf{t}_{N+1}) = \mathcal{N}(\mathbf{t}_{N+1} | \mathbf{0}, \mathbf{C}_{N+1}), \quad (3.76)$$

where \mathbf{C}_{N+1} is an $N+1 \times N+1$ covariance matrix. We partition the \mathbf{C} matrix as

$$\mathbf{C}_{N+1} = \begin{bmatrix} \mathbf{C}_N & \mathbf{k} \\ \mathbf{k}^T & c \end{bmatrix}, \quad (3.77)$$

where \mathbf{C}_N is a $N \times N$ covariance matrix for $n, m = 1, \dots, N$, the vector \mathbf{k} has elements $k(\mathbf{x}_n, \mathbf{x}_{N+1})$ for $n = 1, \dots, N$ and the scalar c is given by $c = k(\mathbf{x}_{N+1}, \mathbf{x}_{N+1}) + \beta^{-1}$. Therefore the conditional distribution $p(t_{N+1}|\mathbf{t})$ is given by a Gaussian distribution with a mean and covariance given by

$$\begin{aligned} m(\mathbf{x}_{N+1}) &= \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{t} \\ \sigma^2(\mathbf{x}_{N+1}) &= c - \mathbf{k}^T \mathbf{C}_N^{-1} \mathbf{k} \end{aligned}$$

The only restrictions that the kernel functions have are that they need to be positive semi-definite (all eigenvalues are greater than zero).

The mean can also be written in the form of

$$m(\mathbf{x}_{N+1}) = \sum_{n=1}^N a_n k(\mathbf{x}_n, \mathbf{x}_{N+1}), \quad (3.78)$$

where a_n is the n^{th} component of $\mathbf{C}_N^{-1} \mathbf{t}$. So if the kernel function $k(\mathbf{x}_n, \mathbf{x}_m)$ only depends on the distance $\|\mathbf{x}_n - \mathbf{x}_m\|$ then we obtain an expression in terms of radial basis functions.

The central computational operation in using Gaussian processes will involve the inversion of a matrix of size $N \times N$, for which standard methods require $\mathcal{O}(N^3)$ computations. By contrast, in the basis function model we have to invert a matrix \mathbf{S}_N of size $M \times M$, which has $\mathcal{O}(M^3)$ computational complexity. Note that for both viewpoints, the matrix inversion must be performed once for the given training set. For each new test point, both methods require a vector-matrix multiplication, which has cost $\mathcal{O}(N^2)$ in the Gaussian process case and $\mathcal{O}(M^2)$ for the linear basis function model. If the number M of basis functions is smaller than the number N of data points, it will be computationally more efficient to work in the basis function framework. However, an advantage of a Gaussian processes viewpoint is that one can consider covariance functions that can only be expressed in terms on an infinite number of basis functions.

3.8.3 Expected Improvement

A key requirement in Bayesian Optimisation is being able to intelligently find the next point to sample, such that we are more likely to find a global minimum in parameter space. This can be done by a variety of methods, the most simple being the probability of improvement, which is just the maximum likelihood maximisation [?]. However, one can improve on the probability of improvement by Expected Improvement given by

$$\text{EI}(\mathbf{x}; x_n, y_n, \theta) = \sigma(\mathbf{x}; x_n, y_n, \theta)(\gamma(\mathbf{x})\Phi(\gamma(\mathbf{x})) + \mathcal{N}(\gamma(\mathbf{x}); 0, 1)), \quad (3.79)$$

where

$$\gamma(\mathbf{x}) = \frac{f(\mathbf{x}_{\text{best}}) - \mu(\mathbf{x}; x_n, y_n, \theta)}{\sigma(\mathbf{x}; x_n, y_n, \theta)}$$

where Φ is the cumulative distribution and $f(\mathbf{x}_{\text{best}})$ is the best (minimum) objective function value [?]. This provides a measure for how likely a point is to find a better minimum by using the estimated standard deviation and mean at that point. Some interesting research has been done by Durant *et al.* where they have analysed other methods for achieving Bayesian Optimisation using methods other than Expected Improvement, such as lower confidence bounds, which have been known to have similar performance [?].

Chapter 4

Results and Discussion

4.1 Tight Binding Band Model: oxygen.

To model a partial pressure environment, one had to obtain a model for oxygen. A simple tight binding model of diatomic oxygen (O_2) was successfully improved upon using a 'semi-analytic' pair potential. The Hamiltonian was constructed in the Linear Combination of Atomic Orbitals basis (LCAO), with basis orbitals of s , and p_x , p_y and p_z orbitals for each atom, using the $2s$ and $2p$ orbitals of the valence shell. This produced an 8×8 Hamiltonian. See 4.1 for its construction, as from Harrison [?]. The on-site energies of the orbitals run on the diagonal of the matrix, given by ϵ_s and ϵ_p for s and p orbitals respectively. The off diagonal terms represent the coupling between orbitals centred on different atoms (the two center integrals). The zeros are from couplings of an atomic orbital to a different orbital to itself, which are of course zero.

The energies of the off-diagonal overlaps can be calculated using the the Slater-Koster formulae table with the appropriate bond integrals, $V_{ab\gamma}(r)$, and direction cosines, l . m and n , which correspond to resolutions of a unit distance between the atoms on the x , y and z coordinates. As the atoms are equivalent, the energies of the corresponding overlaps of an orbital of atom a with atom b and vice versa are the same. With two different species this would not be the case.

Simplifying the Hamiltonian as much as possible, to obtain reasonably simplified eigenvalues, the p_x bonds of the atoms were simulated to be collinear along the x -axis, corresponding to direction cosines of $l = 1$ and $m = n = 0$.

$$H(r) = \begin{bmatrix} \epsilon_s & E_{ss} & 0 & 0 & 0 & -E_{sx} & -E_{sy} & -E_{sz} \\ E_{ss} & \epsilon_s & E_{sx} & E_{sy} & E_{sz} & 0 & 0 & 0 \\ 0 & E_{sx} & \epsilon_p & 0 & 0 & E_{xx} & E_{xy} & E_{xz} \\ 0 & E_{sy} & 0 & \epsilon_p & 0 & E_{xy} & E_{yy} & E_{yz} \\ 0 & E_{sz} & 0 & 0 & \epsilon_p & E_{zx} & E_{yz} & E_{zz} \\ -E_{sx} & 0 & E_{xx} & E_{xy} & E_{xz} & \epsilon_p & 0 & 0 \\ -E_{sy} & 0 & E_{xy} & E_{yy} & E_{yz} & 0 & \epsilon_p & 0 \\ -E_{sz} & 0 & E_{zx} & E_{yz} & E_{zz} & 0 & 0 & \epsilon_p \end{bmatrix}, \quad (4.1)$$

where the r dependence of the energies have been dropped.

To find the available energy states, the Hamiltonian was diagonalised find the eigenvalues E_n , where the subscript n is to denote the n^{th} eigenvalue.

$$E_n(r) = \begin{cases} \epsilon_p - V_{pp\pi}, \\ \epsilon_p - V_{pp\pi}, \\ \epsilon_p + V_{pp\pi}, \\ \epsilon_p + V_{pp\pi}, \\ 1/2(\epsilon_p - V_{pp\sigma} + \epsilon_s + V_{ss\sigma} - \sqrt{(\epsilon_p - V_{pp\sigma} + \epsilon_s + V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s + V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 - \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}), \\ 1/2(\epsilon_p - V_{pp\sigma} + \epsilon_s + V_{ss\sigma} + \sqrt{(\epsilon_p - V_{pp\sigma} + \epsilon_s + V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s + V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 - \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}), \\ 1/2(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma} - \sqrt{(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s - V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}), \\ 1/2(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma} + \sqrt{(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s - V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}) \end{cases} \quad (4.2)$$

For clarity, the explicit r dependence of the bond integrals has been dropped, and will be assumed as to be r -dependent from here.

Oxygen has 6 electrons in its valence shell: 2 electrons in the $2s$ orbital and 4 in its $2p$ orbital. This gives 12 electrons in total for the diatomic system. To find the E_{band} we sum over the occupied electronic states. Each orbital has a maximum occupancy of 2 electrons, thus to find the ground state we fill the lowest 6 energy states with two electrons each. The lowest eigenvalues are given to be $E_n, n \in \{2, 3, 4, 5, 6, 7\}$, which gives the band energy

$$E_{\text{band}} = 9\epsilon_p + 3\epsilon_s + V_{ss\sigma} + 2V_{pp\pi} - V_{pp\sigma} - \sqrt{(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s - V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}.$$

The bond integrals were of Goodwin-Skinner-Pettifor form, from the paper of Lozovoi and Paxton *et al.* [?]. These are defined as, and have derivatives of

$$V_{ab\gamma}(r) = V_{ab\gamma}^0 \left(\frac{r_0}{r} \right)^n \exp \left\{ n \left(- \left(\frac{r}{r_c} \right)^{n_c} + \left(\frac{r_0}{r_c} \right)^{n_c} \right) \right\}, \quad (4.3)$$

$$\frac{dV_{ab\gamma}}{dr} = -\frac{n}{r} \left(1 + n_c \left(\frac{r}{r_c} \right)^{n_c} \right) V_{ab\gamma}(r), \quad (4.4)$$

$$\frac{d^2V_{ab\gamma}}{dr^2} = -\frac{n}{r^2} V_{ab\gamma}(r) \left(n_c(n_c - 1) \left(\frac{r}{r_c} \right)^{n_c} - 1 \right) \quad (4.5)$$

$$- n \frac{dV_{ab\gamma}}{dr} \frac{1}{r} \left(1 + n_c \left(\frac{r}{r_c} \right)^{n_c} \right) \quad (4.6)$$

where a and b are orbitals, and γ denotes the molecular bonding orbital between them (σ or π in this simple case).

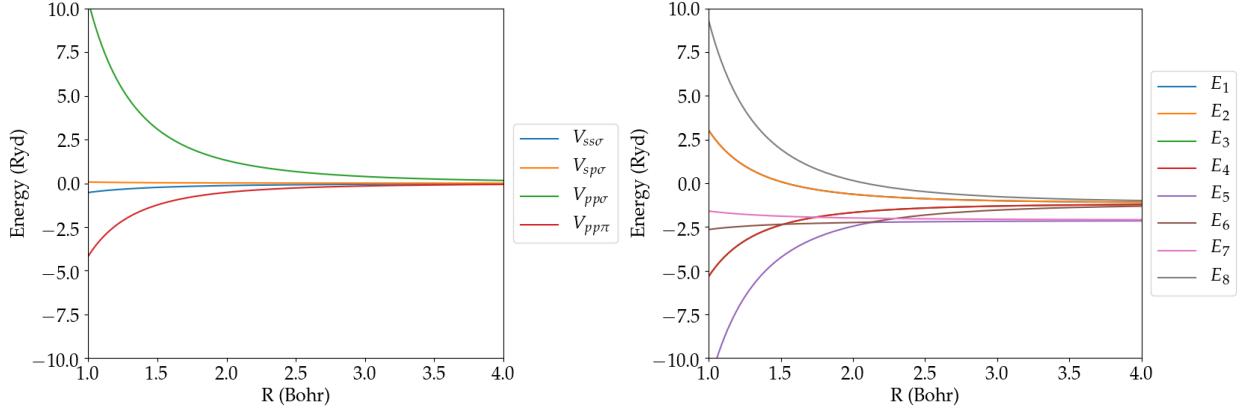


FIGURE 4.1: Left: Plot Energy versus separation of atoms with curves corresponding to the bond integrals. Right: The analytic solution to the eigenvalues of the above Hamiltonian using equation (4.2). From the equations it can be seen that the first and second eigenvalues are degenerate, along with the third and fourth.

To find the binding energy, one needs to include a pair potential that encapsulates the repulsive energy between the atoms. To replicate what is seen empirically, we fit the position of the minimum of the binding energy to the average bond length of an oxygen atom, 2.282 Bohr. In addition, to have the correct dynamics we fit the curvature of this minimum to correspond to the natural vibrational frequency of the oxygen atom: 1552.5cm^{-1} . See 4.2.

The derivatives of the band energy can be shown to be:

$$\frac{dE_{\text{band}}(r)}{dr} = V'_{ss\sigma} + 2V'_{pp\pi} - V'_{pp\sigma} \\ - \frac{((V'_{pp\sigma} - V'_{ss\sigma})(\epsilon_p + \epsilon_s + V_{pp\sigma} - V_{ss\sigma}) + 2(-\epsilon_s V'_{pp\sigma} + \epsilon_p V'_{ss\sigma} + 2V'_{sp\sigma} V_{sp\sigma} + V'_{pp\sigma} V_{ss\sigma} + V_{pp\sigma} V'_{ss\sigma}))}{\left((\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p \epsilon_s - V_{pp\sigma} \epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma} V_{ss\sigma})\right)^{\frac{1}{2}}}$$

$$\frac{d^2E_{\text{band}}(r)}{dr^2} = V''_{ss\sigma} + 2V''_{pp\pi} - V''_{pp\sigma} \\ + \frac{\left(((V'_{pp\sigma} - V'_{ss\sigma})(\epsilon_p + \epsilon_s + V_{pp\sigma} - V_{ss\sigma}) + 2(-\epsilon_s V'_{pp\sigma} + \epsilon_p V'_{ss\sigma} + 2V'_{sp\sigma} V_{sp\sigma} + V'_{pp\sigma} V_{ss\sigma} + V_{pp\sigma} V'_{ss\sigma}))\right)^2}{2\left((\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p \epsilon_s - V_{pp\sigma} \epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma} V_{ss\sigma})\right)^{\frac{3}{2}}} \\ - \frac{((V''_{pp\sigma} - V''_{ss\sigma})(\epsilon_p + \epsilon_s + V_{pp\sigma} - V_{ss\sigma}) + (V'_{pp\sigma} - V'_{ss\sigma})^2 + 2(-\epsilon_s V''_{pp\sigma} + \epsilon_p V''_{ss\sigma} + 2V''_{sp\sigma} V_{sp\sigma} + 2V'^2_{sp\sigma} + V''_{pp\sigma} V_{ss\sigma} + 2V'_{pp\sigma} V'_{ss\sigma} + V_{pp\sigma} V''_{ss\sigma}))}{2\left((\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p \epsilon_s - V_{pp\sigma} \epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma} V_{ss\sigma})\right)^{\frac{1}{2}}},$$

where primed quantities are derivatives with respect to r .

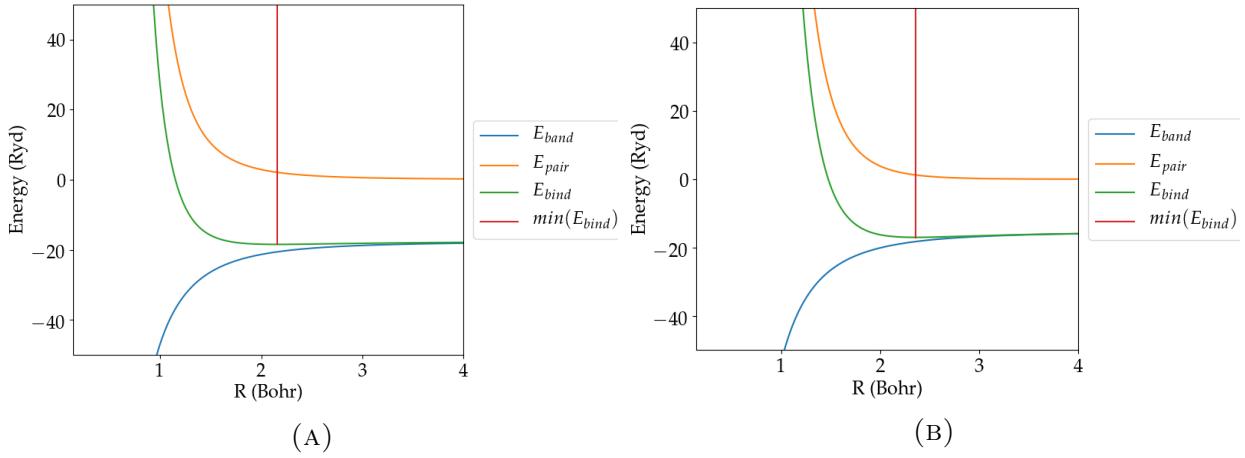


FIGURE 4.2: The band, pair potential and binding energies of diatomic oxygen (a) and nitrogen (b). The minimum of the binding energy is shown by the vertical red line. The pair potential, E_{pair} , was of the form $E_{\text{pair}} = \frac{A}{r^n} e^{-br}$, with parameters as shown in 4.3. The binding energy, $E_{\text{bind}} = E_{\text{band}} + E_{\text{pair}}$ is the sum of the pair potential energy and the band energy. The minimum separation of the binding energy curves are at the average bond lengths of oxygen and nitrogen, 2.282 Bohr and 2.382 Bohr respectively. The curvature was adjusted to agree with the primary vibrational modes of both molecules: 1552.5cm^{-1} and 2358cm^{-1} for oxygen and nitrogen respectively.

The pair potential used was of exponential power law form:

$$E_{\text{pair}}(r) = \frac{A}{r^n} e^{-br}. \quad (4.7)$$

This gave appropriate decay and curvature characteristics compared to simple sum of exponentials and power laws, while having a reduced number of parameters to fit. Differentiating the above equation twice and equating to the gradient and curvature of the binding energy at the equilibrium bond length, $r_0 = 2.82$ Bohr, respectively, one finds solutions for the b and A coefficients to be

$$b = \frac{\kappa}{2} - \frac{n}{r_0} + \sqrt{\kappa^2 - 4\frac{n}{r_0^2}} \quad (4.8)$$

$$A = \frac{r_0^{n+2} e^{br_0} \kappa \frac{dE_{\text{band}}}{dr}|_{r=r_0}}{n + (br_0 + n)^2}, \quad (4.9)$$

where

$$\kappa = \frac{(\gamma_{\text{O}_2} - \frac{d^2 E_{\text{band}}}{dr^2}|_{r=r_0})}{\frac{dE_{\text{band}}}{dr}|_{r=r_0}}, \quad (4.10)$$

and γ_{O_2} is the natural frequency of O_2 in units of Ryd/Bohr². These equations gave the correct r_0 for a given n , but n had to be adjusted until the correct curvature was observed. However, it was seen that n would have to be large, $n \geq 5.75$, for the curvature to match. This was above the expected pure power law dependence argued in literature to be $2 < n < 5$ in [?], and quite large compared to pair potentials in the original Paxton paper (which

Molecule	A	b	n	Method	r_0	r_0 Exp.	Wav ⁿ	Wav ⁿ Exp.
O_2	160	0.685	3.8552	tbe	2.231	2.282	1553.9	
				Python	2.2304		1577.9	1552.5
N_2	710	0.96586	4.75955	tbe	2.357	2.362	2351.1	
				Python	2.3606		2351.3	2358.0

FIGURE 4.3: Table for results of pair potential derived from analytic results for O_2 and N_2 . A , b and r_0 are in units of Bohr. Wavⁿ is the wavenumber derived from the curvature of the binding energy minima and is in units cm^{-1} . Exp. are the experimental quantities for both the bond length and the wavenumber. Analytic results were used in self-consistent tight binding tbe code, and gave good agreement, even though the oxygen wavenumber is a couple of percent off in tbe.

actually did not give the correct curvature or r_0). As such, adjustment to the parameters was done by hand, such that the curvature matched more closely. A table of the results can be seen in 4.3.

4.2 Tight Binding Band Model: nitrogen

For nitrogen, one can obtain similar results. The band energy is given by the sum over the lowest 5 eigenvalues of 4.1. Using the same bond integrals as for oxygen, this gives

$$E_{\text{band}} = 7\epsilon_p + 3\epsilon_s + V_{ss\sigma} + 4V_{pp\pi} - V_{pp\sigma} - \sqrt{(\epsilon_p + V_{pp\sigma} + \epsilon_s - V_{ss\sigma})^2 + 4(-\epsilon_p\epsilon_s - V_{pp\sigma}\epsilon_s + V_{sp\sigma}^2 + \epsilon_p V_{ss\sigma} + V_{pp\sigma}V_{ss\sigma})}.$$

Using the same procedure as for O_2 : evaluating E_{band} derivatives, using equations (4.8), (4.9) and (4.10) (substituting γ_{O_2} for γ_{N_2}) for estimation and then adjusting parameters by hand, one finds parameters for the pair potential be $A = 710$, $b = 0.96586$ and $n = 4.75955$. A table of the fitting and comparison with tbe can be seen in 4.3.

The pair potential derived from simple calculations in python give almost exactly the same results within the self-consistent polarisable-ion tight binding framework of tbe, as expected.

4.3 Tight Binding Band Model: titanium

Like for oxygen and nitrogen, a Hamiltonian for HCP titanium was constructed using the LCAO basis. Diagonalisation of the s , p and d orbital Hamiltonians was achieved and gave band energies. As an infinite lattice was being simulated, invocation of Bloch's theorem was necessary: due to translational invariance along principal lattice vectors, the problem could be reduced to a unit cell with Bloch phase factors in momentum space. HCP structures have a two-atom unit cell, with the first atom being at the origin, and the second atom at the position $\mathbf{r} = (\frac{1}{2}a, \frac{1}{2\sqrt{3}}a, \frac{1}{2}c)$. As before, we have one s -orbital, three p orbitals (x , y and

z) and five d -orbitals (xy , yz , zx , $x^2 - y^2$ and z^2), so the s , p and d matrices in isolation are:

$$H_s(\mathbf{k}) = \begin{bmatrix} \epsilon_s & g_0 V_{ss\sigma} \\ g_0^* V_{ss\sigma} & \epsilon_s \end{bmatrix} \quad (4.11)$$

$$H_p(\mathbf{k}) = \begin{bmatrix} \epsilon_p & 0 & 0 & g_0 E_{xx} & g_0 E_{xy} & g_0 E_{xz} \\ 0 & \epsilon_p & 0 & g_0 E_{xy} & g_0 E_{yy} & g_0 E_{yz} \\ 0 & 0 & \epsilon_p & g_0 E_{zx} & g_0 E_{yz} & g_0 E_{zz} \\ g_0^* E_{xx} & g_0^* E_{xy} & g_0^* E_{xz} & \epsilon_p & 0 & 0 \\ g_0^* E_{xy} & g_0^* E_{yy} & g_0^* E_{yz} & 0 & \epsilon_p & 0 \\ g_0^* E_{zx} & g_0^* E_{yz} & g_0^* E_{zz} & 0 & 0 & \epsilon_p \end{bmatrix} \quad (4.12)$$

$$A = \begin{bmatrix} \epsilon_d & 0 & 0 & 0 & 0 \\ 0 & \epsilon_d & 0 & 0 & 0 \\ 0 & 0 & \epsilon_d & 0 & 0 \\ 0 & 0 & 0 & \epsilon_d & 0 \\ 0 & 0 & 0 & 0 & \epsilon_d \end{bmatrix}$$

$$B = \begin{bmatrix} g_0 E_{xy,xy} & g_0 E_{xy,yz} & g_0 E_{xy,zx} & g_0 E_{xy,x^2-y^2} & g_0 E_{xy,z^2} \\ g_0 E_{xy,yz} & g_0 E_{yz,yz} & g_0 E_{yz,zx} & g_0 E_{yz,x^2-y^2} & g_0 E_{yz,z^2} \\ g_0 E_{xy,zx} & g_0 E_{yz,zx} & g_0 E_{zx,zx} & g_0 E_{zx,x^2-y^2} & g_0 E_{zx,z^2} \\ g_0 E_{xy,x^2-y^2} & g_0 E_{yz,x^2-y^2} & g_0 E_{zx,x^2-y^2} & g_0 E_{x^2-y^2,x^2-y^2} & g_0 E_{x^2-y^2,z^2} \\ g_0 E_{xy,z^2} & g_0 E_{yz,z^2} & g_0 E_{zx,z^2} & g_0 E_{x^2-y^2,z^2} & g_0 E_{z^2,z^2} \end{bmatrix}$$

$$H_d(\mathbf{k}) = \begin{bmatrix} A & B \\ B^* & A \end{bmatrix},$$

where

$$g_0 = \sum_{\mathbf{d}} e^{i\mathbf{k}\cdot\mathbf{d}},$$

where \mathbf{d} is the distance to a given atom and \mathbf{k} is a wavevector.

Upon diagonalisation of these matrices one can get the energies of the s , p and d bands. Similar work was done for an FCC structure, but for brevity, and the relative simplicity of FCC—it only has one atom per primitive unit cell and hence half the bands—the results have been omitted.

4.4 Fitting of Parameters for Pure titanium

The first quantity to fit was the band structure. The bond integrals and their decay must give the correct energies for the electrons. Using a Spanjaard-Desjonquères exponential decay model for the bond integrals we can get the right bond integrals.

A general scheme was used to fit the tight binding bands to DFT. The Empirical Tight Binding and LMTO packages within the Questaal suite were used for this [?]. DFT calculations were done to see how much s , p and d character each of the bands had. From this character data, one could sort the bands according to character at any k point, say the

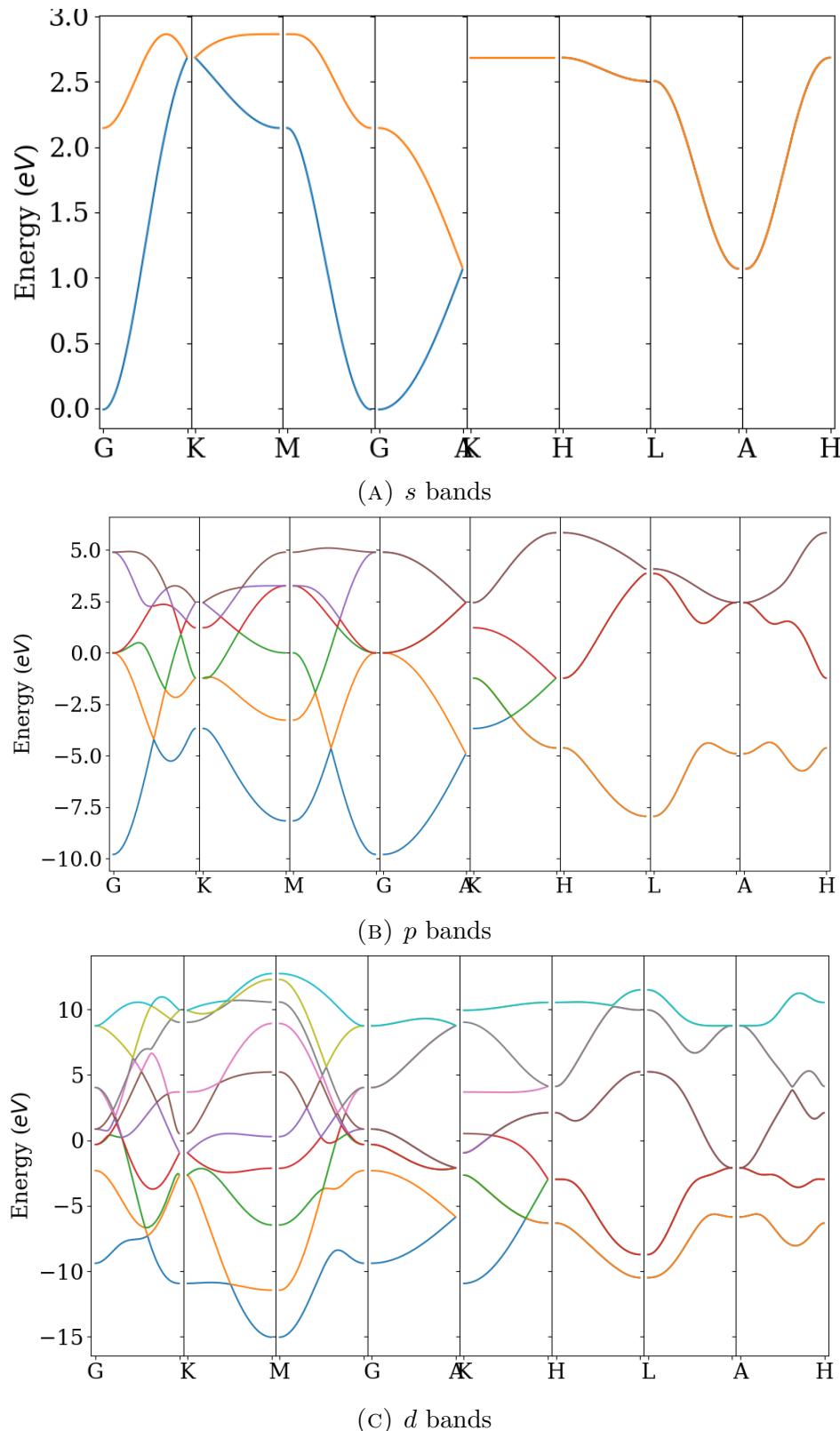


FIGURE 4.4: *s*, *p* and *d* bands for HCP titanium scaled to HCP bands shown in Andersen and Jepsen [?]. Different colours representing different eigenvalues of the hamiltonian. The letters correspond to high symmetry points in the walk along k -space.

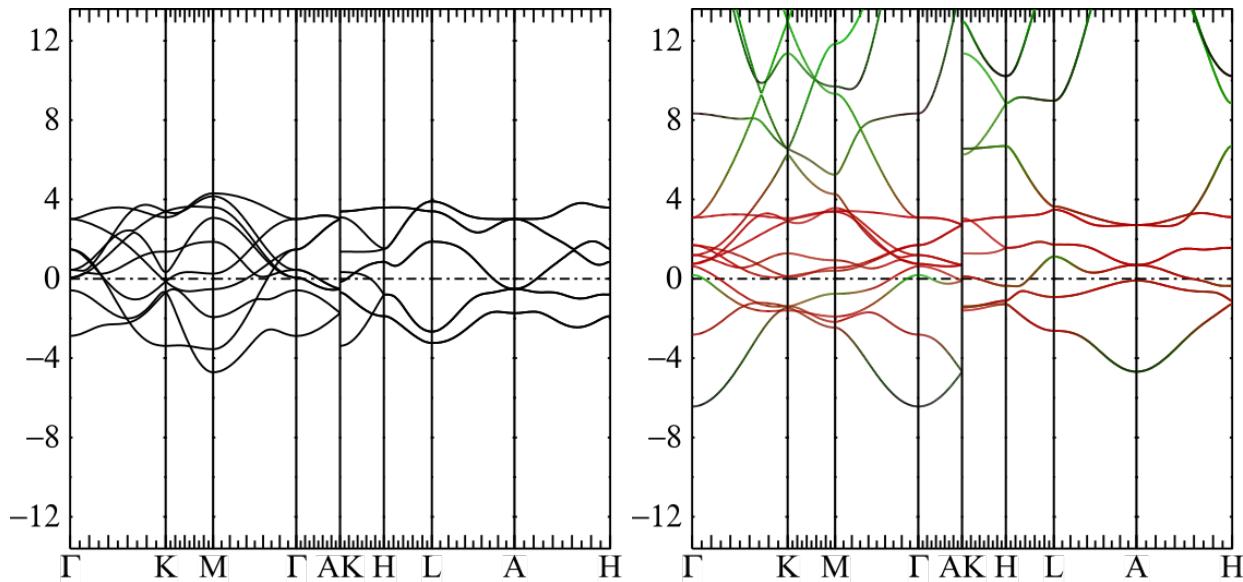


FIGURE 4.5: Band structures for orthogonal d band model in Tight Binding (left) and for DFT (right) for HCP titanium. The energy (y) axis is in eV and the x-axis is the walk along k -space with letters denoting traversal of high symmetry points on the Brillouin zone irreducible wedge. The DFT bands are coloured according to how much character the bands had. Red denotes d -orbital character, minority d/p -orbital character is in green and black is majority s -orbital character. The TBE bands were fitted to the Γ point of the bands. The DFT results were validated by reference to Jafari 2012 [?]

Gamma point, and then calculate the associated bandwidth and then fit the corresponding tight binding bands by varying a global scaling coefficient using a binary search scheme.

The these bands make sense for titanium. The s and p orbitals would have a large overlap as they are extended in space, hence the large band width. As the d -orbitals of the metal are compact, they would have little overlap, hence there is a small bandwidth associated with their structure. As this is a metal one expects no band gap as is shown here (the Fermi level is at 0 in the band diagrams).

To get a good model for titanium, we need to fit to empirical quantities of the metal. The bond integrals must give the correct band structure for titanium. With the addition of the pair potential there must be correct a and c/a parameters. Crucially, for dislocation dynamics we need the right bulk modulus and elastic constants.

To obtain elastic constants to fit to, it was necessary to fit to data of the elastic constants' temperature dependence and then extrapolate to zero temperature. Points from below the Debye temperature of 420K were used to avoid the transition from phonons to independent lattice vibrations [?]. Using Fisher and Renken's data the curves were extrapolated to 0K [?]. Data from the more recent paper of Ogi and Kai was obtained, but there were not enough data points for a reasonable regression line [?]. See 4.6 for figure.

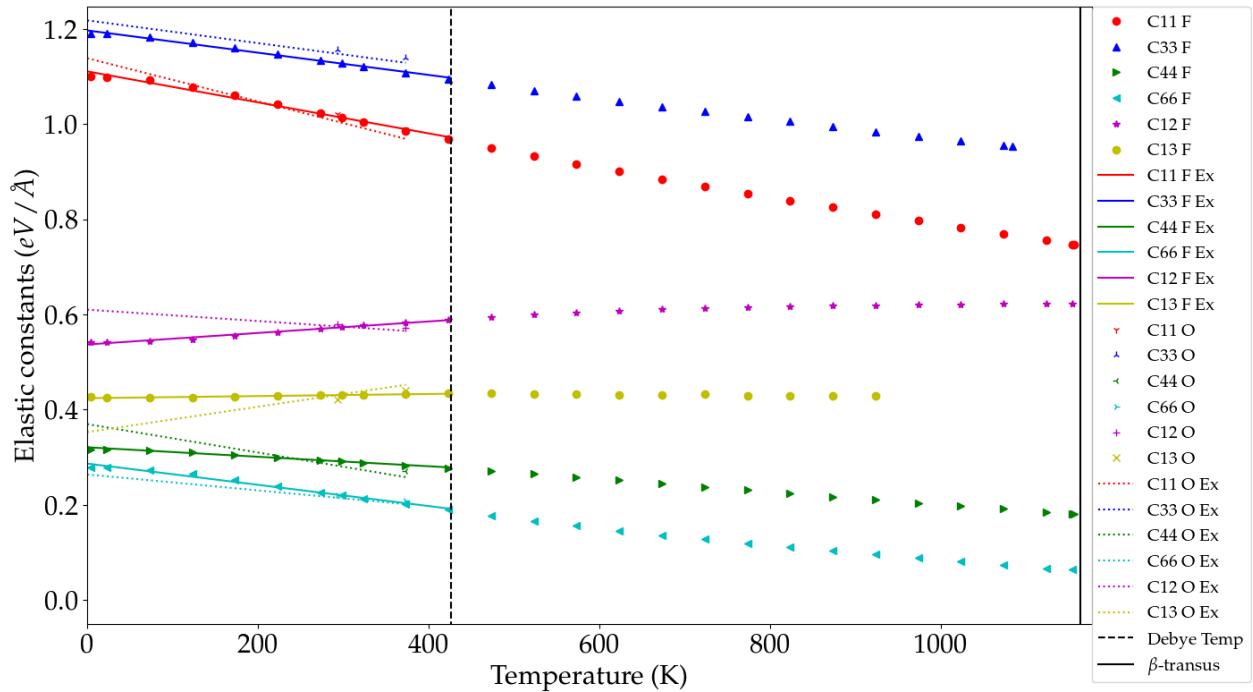


FIGURE 4.6: Plot of Elastic constants using the data of Fisher, Renken *et al.*, and Ogi and Kai [? ?]. The curves were extrapolated from data below the Debye Temperature. The β transus temperature is marked to make sure points are not taken from a different polymorph of the material.

4.5 Numerical Optimisation: Tests

To optimise the model parameters, Bayesian Linear Regression and Gaussian process regression were used. An advantage of these methods are that the standard quadratic regulariser for regression arises naturally in Gaussian Process Regression and has closed form solution in Bayesian Linear Regression [? ?].

Trialling simple forms of Bayesian optimisation initially, one used a Bayesian Linear regression method. The speed of the regression fit is quick, as the matrix one has to invert is only $N \times N$, where N is the number of basis functions however, there is a caveat for a good fit: the underlying form of the function of which we are trying to fit to must be able to be well represented by the basis functions. 'Gaussian' and polynomial bases were trialled. To test the results of the algorithm, replication of an objective function was achieved by an analytic function in 1-D. These were chosen to be: a polynomial quartic and a sum of sine curves of differing frequencies and magnitudes, with additional Gaussian noise.

For polynomial regression, within a polynomial basis, the Bayesian Linear Regression converged quickly, as long as the number of bases (the degree of the polynomial the Bayesian method is trying to fit) were increased to more than 15. However, fitting with Gaussian basis functions (each with a specific width and mean), was seen to be unreliable in giving a faithful representation of underlying functions, given that the number of input basis functions is fixed as each point is added. Furthermore, if optimization of this method were to be achieved, the optimal spacing and width of each of the Gaussian basis functions would have to be estimated, increasing the number of parameters rapidly if a large number of Gaussian

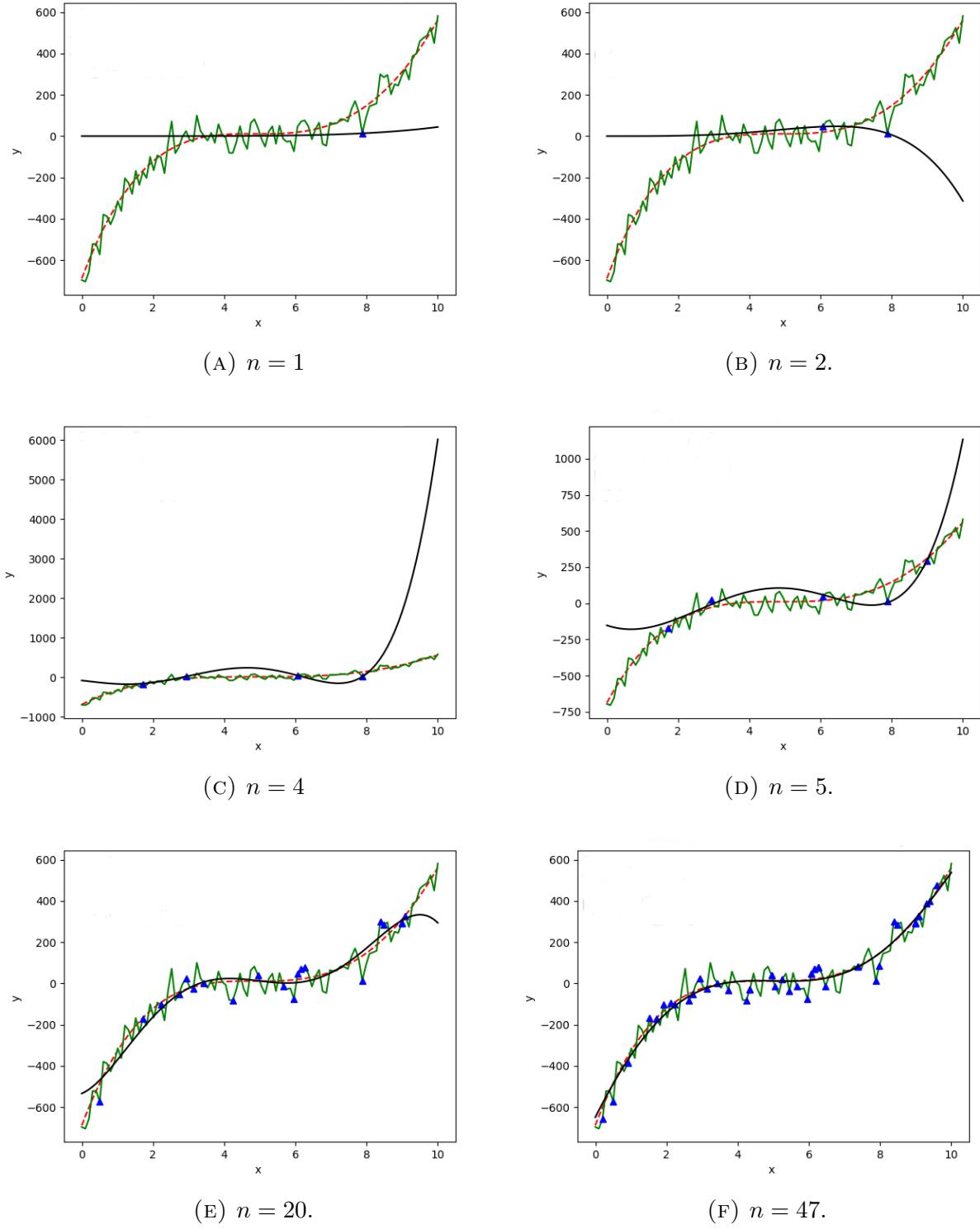


FIGURE 4.7: Bayesian Linear Regression test for a quartic polynomial of $t = -0.5x^4 + 5x^3 + 2x^2 + 11 + \epsilon$, where ϵ Gaussian noise with a variance of 50, using a 5 polynomial basis functions. The number of sampled points is given by n . The mean of the distribution is given in black. The red dashed line is the underlying function of which noise is added to obtain the green curve.

Blue points correspond to random points sampled from the green curve.

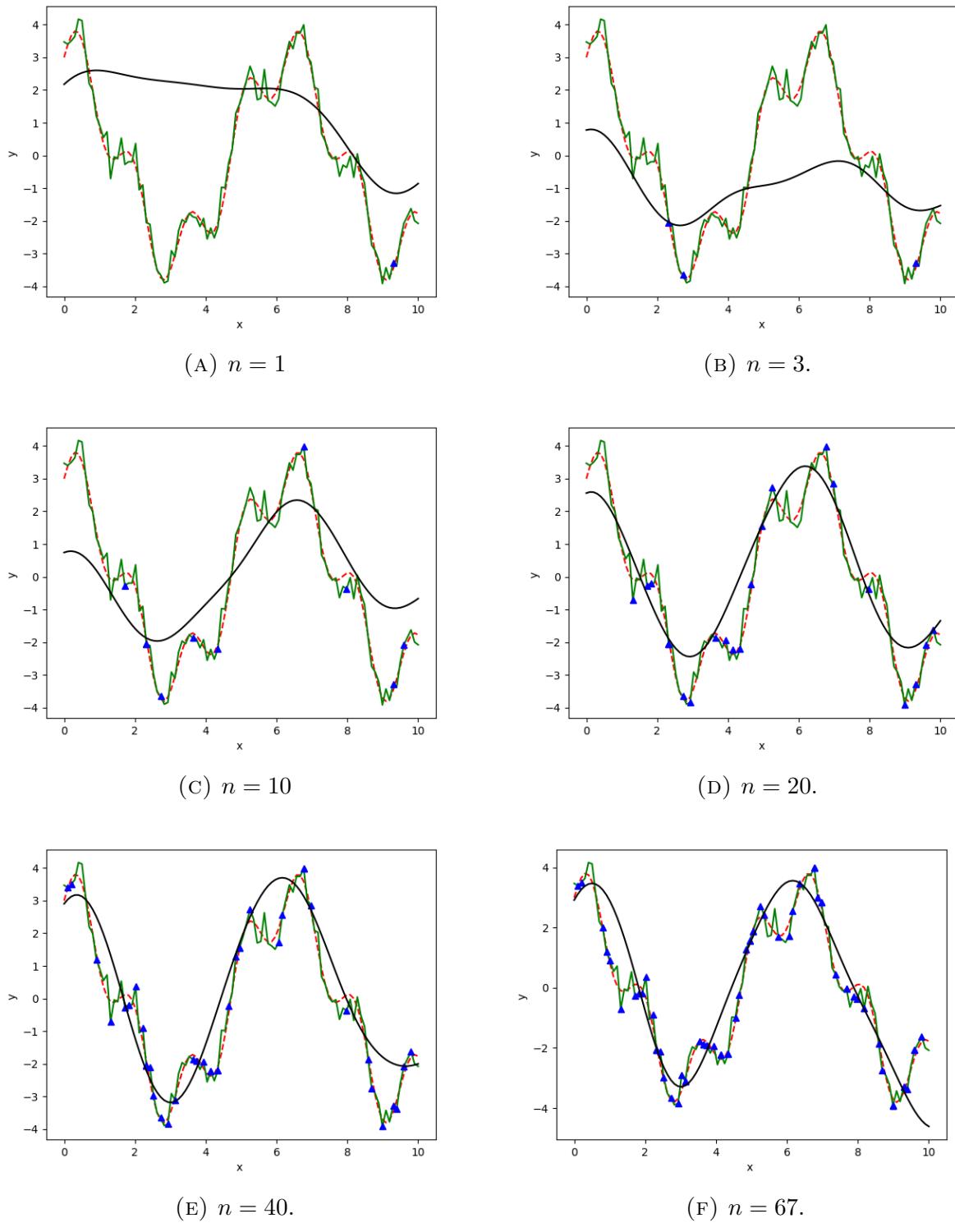


FIGURE 4.8: Bayesian Linear Regression test to fit $y = \sin(4x) + 3\cos(x) + \epsilon$ where ϵ is Gaussian noise of standard deviation 0.3 using a basis of 40 equally spaced Gaussians and $\lambda = 0.1$. The number of sampled points is given by n . The mean of the distribution is given in black. The red dashed line is the underlying function of which noise is added to obtain the green curve. Blue points correspond to random points sampled from the green curve.

basis functions were necessary for good representation. This implies poor construction of the objective function hypersurface in a higher-dimensional parameter space, if each of the parameters in each of the basis functions are not optimised, making minimisation to a real global minima doubtful.

Gaussian process regression reduced the time taken for the reconstruction of the analytic function. With number of points $n > 10$, the unoptimised parametric kernel function, represented the sum of sines better than that of Bayesian linear regression. This implies better performance in objective function minimisation. As can be seen in Cyan, the expected improvement of finding a minimum of the function at a particular point was shown. This found the global minimum.

4.6 Numerical Optimisation: titanium

The algorithm to find the minima of the objective function has two modes: exploration, where (random) sampling of the parameter space is done to ascertain a rough objective function landscape, and exploitation, taking advantage of expected improvement from the samples to find the global minimum. There has to be the right balance between exploration of parameter space and minimisation using expected improvement to find the global minimum.

The algorithm draws correlations between each of the input parameters and the objective function. The expected improvement for each input is computed for each of the allowed configurations of parameter space. The sum of the expected improvements is then maximised with regards to the possible parameters find the next best point to sample to find the minimum of the objective function.

The Bayesian optimisation algorithm was used to fit a sum of exponentials pair potential ($Ae^{-br} + Ce^{-dr}$) and bond integrals of exponential (Spanjaard Dejonquères) form. These are the input parameters from which one hopes to reproduce experimental elastic constants, bulk modulus, a and c/a for HCP titanium. The choice of whether to sample input space at random, or to use Expected improvement was simply reduced to sampling randomly twice, and then performing Expected Improvement. To test the optimisation algorithm, the `optimize` method from the python library `scipy` was used with a Polak-Ribière non-linear conjugate gradient method. Two objective functions were used. The first was a general error function of

$$f_{\text{obj}}^{(1)}(q(\mathbf{x})) = \sum_{n=0}^N \{t_n - q_n(\mathbf{x})\}^2, \quad (4.13)$$

where $q_n(\mathbf{x})$ is a quantity we want to fit—which is a function of the input parameters \mathbf{x} —to the target values t_n .

This had unphysical values for quantities, and elastic constants became negative with this approach, so a regularised error was put in, of the form

$$f_{\text{obj}}^{(2)}(q(\mathbf{x})) = \sum_{n=0}^N \{t_n - q_n(\mathbf{x})\}^2 - \frac{1}{2} (\mathbf{q}(\mathbf{x}))^2, \quad (4.14)$$

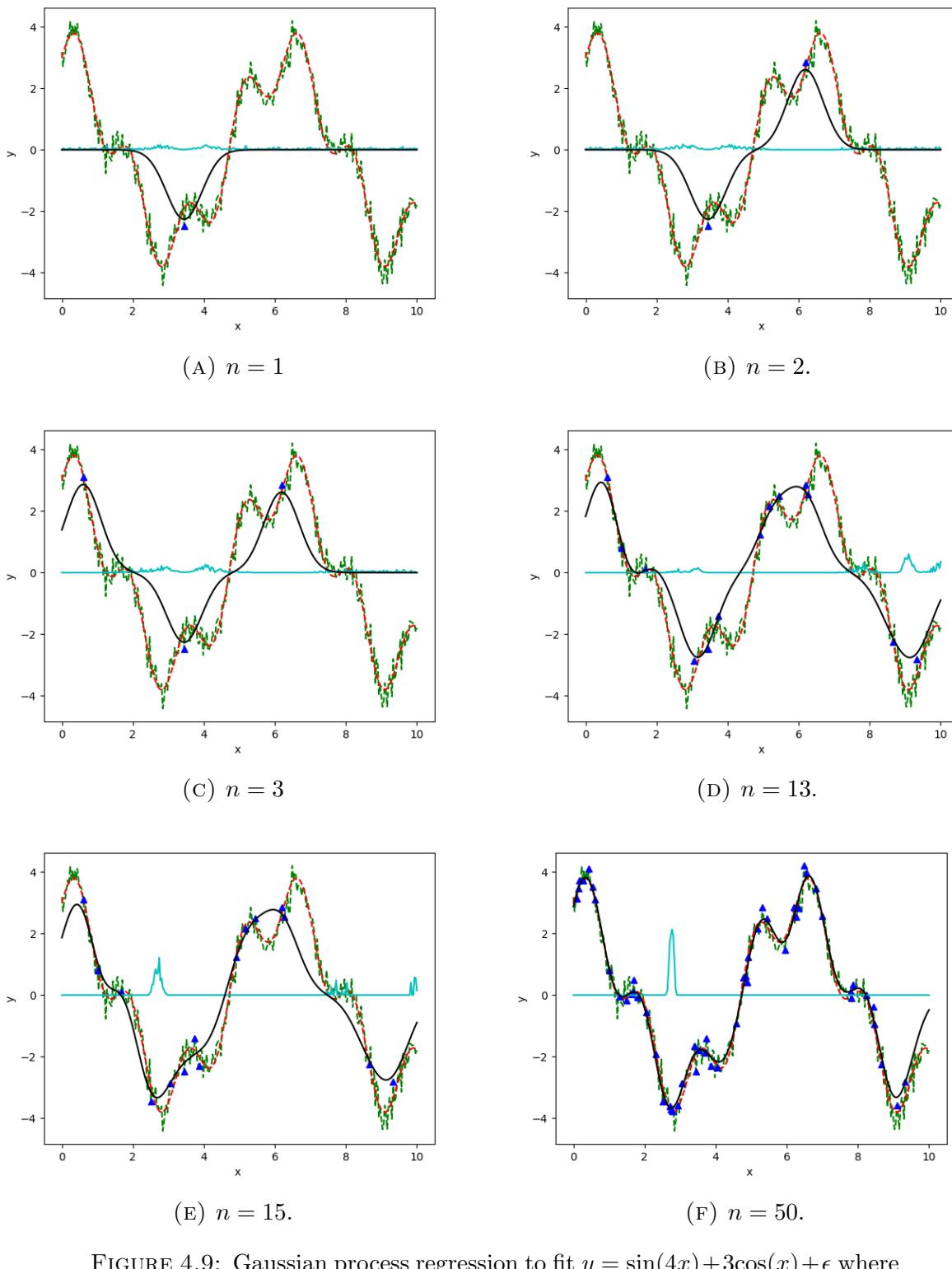


FIGURE 4.9: Gaussian process regression to fit $y = \sin(4x) + 3\cos(x) + \epsilon$ where ϵ is Gaussian noise of standard deviation 0.3. The hyperparameters were unoptimised and of values $\theta_1 = 1, \theta_2 = 4$ and $\theta_3 = \theta_4 = 0$ for the parametric kernel. The number of sampled points is given by n . The mean of the distribution is given in black. The red dashed line is the underlying function of which noise is added to obtain the green curve. Blue points correspond to random points sampled from the green curve of which the fit is constructed from. The cyan line is the Expected Improvement in finding a global minima of the function being fitted, from the Gaussian process regression. it can be seen that it is able to distinguish between local and global minima and it finds the true minima of the function as can be seen in (f)

Fit quantities	PR (1)	PR (2)	Bayesian (1)	Bayesian (2)	t_n Experiment
a	4.887	4.904	6.149	6.462	5.585
c/a	1.533	1.527	1.709	1.630	1.587
C_{11}	2.763	3.873	0.904	4.573	1.110
C_{33}	6.481	7.757	1.926	4.910	1.196
C_{44}	0.970	2.208	0.053	0.937	0.321
C_{66}	-0.029	1.420	-0.535	1.458	0.287
C_{12}	2.017	1.294	2.343	1.641	0.537
C_{13}	2.533	1.810	0.737	1.351	0.424
K	670519.2	3469910.6	111.1	302.7	110.0

FIGURE 4.10: Table of results using both Polak-Ribière (PR) method from `scipy.optimize` and the Bayesian optimisation for (1) unregularised and (2) regularised objective functions. a is in units of Bohr. K is the bulk modulus and is in units of GPa. The elastic constants C_{ij} are in units of eV/Å

which penalised large values of parameters. In addition, to make sure that the elastic constants produced were positive, for m negative elastic constants produced, the total elastic constant error became

$$(1 + m) \sum_{n_C=0}^{N_C} \{t_{n_C} - q_{n_C}(\mathbf{x})\}^2,$$

for N_C elastic constants. Each optimisation was run for three days.

For a given pass, the Bayesian optimisation approach, when calculating the Expected Improvement, was slower, as integration at every point in the target parameter space has to be achieved for the calculation of Expected Improvement. But, evaluation of the Hessian of Polak-Ribière took multiple evaluations of the objective function, so in reality, both methods took a similar time to take steps in parameter space.

As can be seen, neither optimisation method fared well. The Polak-Ribière method from `scipy.optimize` gave frankly farcical bulk moduli, even with regularisation. This was likely a problem in finding the optimal structure resulting from bad input parameters, as a and c/a are at the edges of the constraints set. The Bayesian approach worked far better and gave a lower overall error for the objective function. The best agreement with experiment is that of the unregularised objective function Bayesian. This gave a bulk modulus very close to that of experiment, even with the elastic constants being off. The extra penalties to negative elastic constants can be deemed to be too harsh, as the resulting elastic constants for both minimisation methods are far higher than that of the original results with the first objective function.

Due to the Polak-Ribière method halting after three days, it is possible that the optimisation routine for the Bayesian routine was not run for a sufficient time to converge on the best result possible: it was stopped prematurely such that real-time comparison to the `scipy` method could be made. The actual objective function might be of error due as can be seen from the non-physical results obtained from both methods. Perhaps there is an error with the underlying methods of obtaining the target quantities themselves as neither optimisation method converged to the desired result.

Improvements can be made to this Bayesian routine. The kernel function hyperparameters were not optimised, as can be achieved by maximisation of the posterior for the \mathbf{C} matrix, given imposition of a prior over hyperparameters. This could have resulted in poor reconstruction of the objective function surface, so slower convergence to an optimal set would have been likely.

As for the fitting of models during this PhD, if the Bayesian optimisation approach proves unreliable, or too slow for evaluation of total Expected Improvement over a very high dimensional parameter space, then Schwefel's Genetic algorithm may be used, but hastened by parallelization. Currently, the method within the Questaal code is sequential, even though genetic algorithms are inherently parallelizable.

Chapter 5

Future Work

5.1 Outline of Strategy

The general strategy of this PhD will be to first, obtain good models for titanium and titanium dioxide from numerical optimisation. A large amount of effort will be to obtain parameters for Ti-O interactions that are consistent with both pure titanium and oxygen interactions—for accurate modelling of oxygen-core interaction—and first neighbour interactions of TiO_2 . Special attention will be paid to the Ti-Ti parameters, as it must describe both bulk titanium and the second-neighbour interactions of TiO_2 , which is necessary for a good representation of TiO_2 crystal structure.

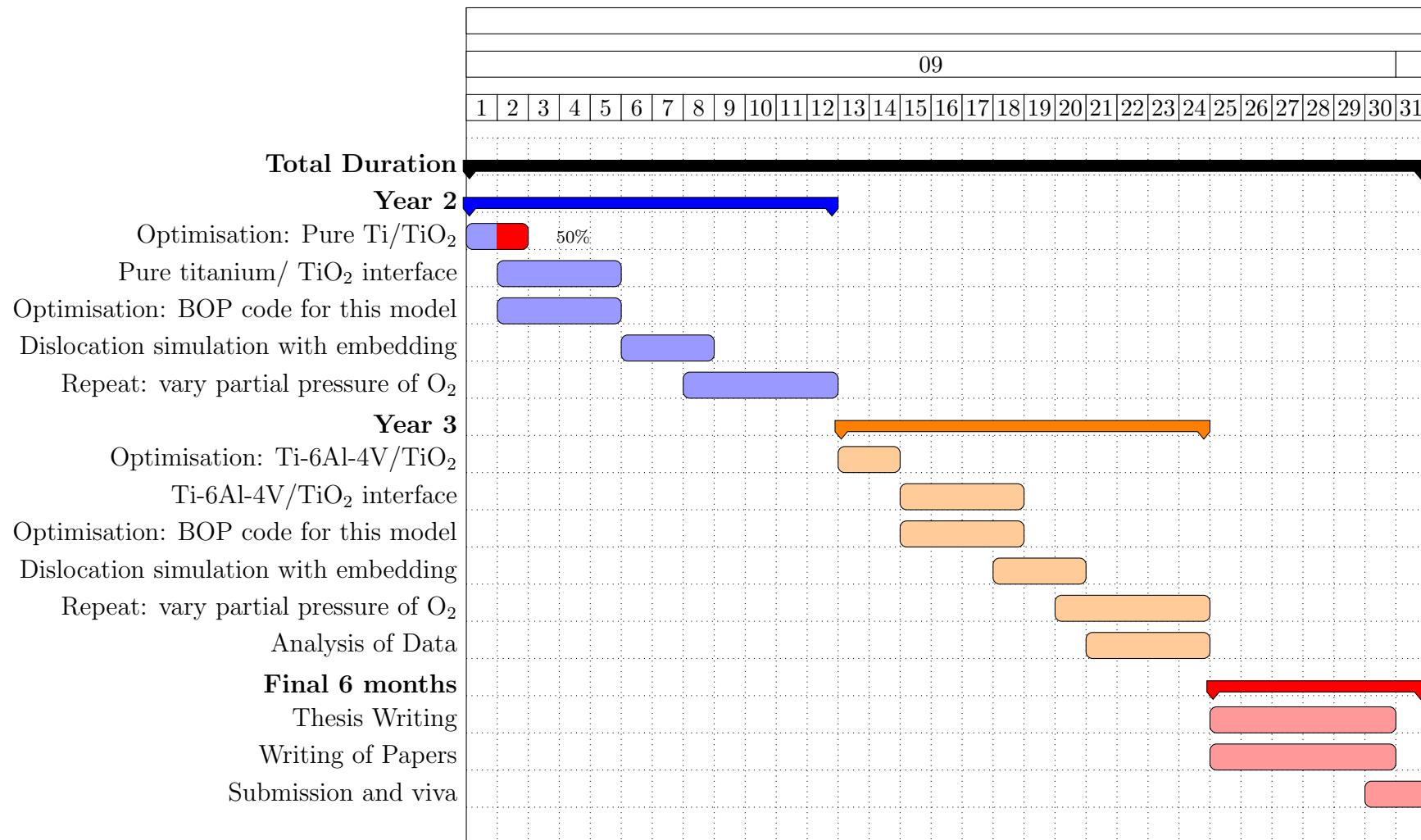
After this initial stage, simulations can begin. First, from the advice of Dave Rugg, pure titanium will be interfaced with titanium dioxide, in an environment of differing partial pressures of oxygen (and possibly nitrogen). From this, one can see how oxygen diffuses through the oxide layer, or if titanium actually migrates up into the oxide layer, causing growth. It would be interesting to see how the partial pressure affects vacancy migration of oxygen and how these results compare with the classical description of oxygen diffusion through alumina grain boundaries, as investigated by Tautschnig and Finnis [?]. The structure of the surface (e.g. if it were to be (110) rutile or (101)/(001) anatase with some number of surface and subsurface oxygen vacancies) will have important effects on the migration of oxygen, and dissociation of oxygen molecules [? ?].

One can then investigate how oxygen interacts with dislocations in isolation. Initially in screw dislocations in bulk titanium ($\langle a \rangle$ and $\langle c + a \rangle$). The effect of oxygen on double kink migration will be studied. Subsequent investigation of dislocation interactions at the interface of bulk titanium and titanium dioxide can then commence, to see how SCC manifests. The effect of dislocation pile-up on the extent of damage at the Ti/TiO_2 interface will be ascertained.

Once these systems have been studied in detail, the analysis of alloyed systems can then begin. Tight binding and BOP models for Ti-6Al-4V will be made by keeping the parameter set as before, for pure titanium and titanium dioxide, while adding more interactions with Ti-Al and Ti-V. Again, the effect of these solutes on the core structure of dislocation will be investigated. Dislocation interaction at the interface between the titanium alloy and the surface oxide will be simulated, and differences between apparent mechanisms of stress corrosion cracking seen in the pure titanium model to that of the alloyed model, will be

made clear. Modifications to the oxide layer may be made, such that one can see how α -case and α_2 $TiAl_3$ phases change oxygen diffusion, solute-core interactions and surface oxide damage.

5.2 Gantt Chart



Chapter 6

Conclusion

In this report, a general outline of what is expected during this PhD has been shown. Motivation for the PhD arises from the unknown mechanisms of interaction between oxygen and dislocation cores, which in turn affects the plasticity of titanium. To accurately ascertain the effects of oxygen on plasticity and corrosion in a titanium/titanium dioxide system, one proposed the use of a multi-scale atomistic approach of "embedding": a technique whereby a self-consistent polarisable-ion tight binding model governs the energies of atoms in the non-linear stressed cores of dislocations, while a bond order potential model—which provides a faster ($\mathcal{O}(N)$ compared to $\mathcal{O}(N^3)$) but less accurate, approach to modelling atomic energies—governs the regions of atoms outside of the core. This provides incredibly good scaling properties, in contrast to DFT, such that both the long-ranged stress fields of the dislocation—which are usually described classically—and the non-linear stressed core region—which originates from quantum mechanics—are accurately described by a single green's matrix formalism. The effects of interstitial oxygen on the dynamics of dislocations, and interactions between dislocation systems of various types, in both titanium (alloys) and titanium dioxide, will be studied. Enhanced stress fields, generated from dislocation pile up at the interface between titanium dioxide and titanium, and its effect on Stress Corrosion Cracking, will be described. Parameter fitting, necessary for the implementation of tight binding and bond order potential methods, will be based on efficient Bayesian Optimisation techniques. This will result in an acceleration of the fitting procedure, allowing for expansion of the simulations from a pure titanium medium, to an alloyed medium. Simulations of the alloyed system in a partial pressure O_2 environment will provide insight into how oxygen diffuses into the the material. The difference of dislocation propagation, interaction and pile up at the alloy/oxide interface will be accomplished. Comparison of crack morphologies, crack formation and susceptibility between alloyed and pure titanium mediums, will be drawn.

Chapter 7

Notes

7.1 Notes on BZ integration

<https://wiki.bnl.gov/CFN-Computation/images/3/34/K-point.pdf>

Integrals over the BZ are necessary to obtain the density of states, the charge density, the matrix elements. This needs a weighted sum over k-points.

In metals a dense grid is necessary as the occupancy is discontinuous at E_F . This implies a large number of fourier components are necessary such that it can be reconstructed well.

This leads to the idea of Fermi-Dirac smearing at E_F .

This has the added effect that now the Energy is not variational with respect to partial occupancies. So we need a free energy.

The free energy is

$$F = E - T \sum_n \sigma S(f(n)),$$

where σ is a smearing parameter which is generally equal to $k_B T$. The entropy is defined to be

$$S(f) = -[f \log(f) + (1 - f) \log(1 - f)]$$

σ can be interpreted as finite temperature. As has been done by Mermin,

Can have gaussian smearing but variational functional $F(\sigma)$ differs from $E(0)$. forces are calculated as derivatives of the variational quantity ($F(\sigma)$). not necessarily equal to forces at $E(0)$.

Method of Methfessel and Paxton (1989) Idea: expansion of stepfunction in a complete set of orthogonal functions term of order 0 = integral over Gaussians generalization of Gaussian broadening with functions of higher order.

MP of order N leads to a negligible error, if $\chi(\epsilon)$ is representable as a polynomial of degree $2N$ around ϵF . linewidth σ can be increased for higher order to obtain the same accuracy entropy term ($S = \sigma \sum_n S_N(f_n)$) describes deviation of $F()$ from $E()$. if S_i few meV then $\hat{E}(\sigma) \approx F(\sigma) \approx E(\sigma) \approx E(0)$. forces correct within that limit. in practice: smearings of order $N=1$ or 2 are sufficient

Tetrahedron integration tetrahedra can break the symmetry of the Bravaislattice at least 4 k-points are necessary must be included linear interpolation under- or overestimates the real curve

Corrections by Blochl (1993) Idea: linear interpolation under- or overestimates the real curve for full-bands or insulators these errors cancel for metals: correction of quadratic errors is possible:

7.2 Why are electron states in the band gap

<https://www.quora.com/What-does-an-electronic-state-in-the-band-gap-mean-How-can-electron-states-exist-in-a-forbidden-energy-region>

Forget the $u(r)u(r)$ part, think of them as plane waves, and we can talk about what's happening to the electrons.

Let's take a simple example. Say all your carriers are conduction-band electrons. Put them all in states close to a band energy minimum near $|k| = 0$. Make the problem isotropic. Then you can pretend you're dealing with electrons in free space. (I.e., energy $E(k)$ is proportional to $|k|^2$.) This is the free electron model.

As you've pointed out there are no localized states. A free electron can't just sit localized in free space. It would disperse! And we know the minimum-energy state in free space, it's a plane wave that has zero $|k|$.

Keep pretending we're in free space. But now introduce a defect. Specifically, put a positive charge $|e|$ at $x, y, z = 0, 0, 0$. Now I still see plane waves as electron states*, but I also see bound states. Specifically I see the bound states of the hydrogen atom, spread out below the minimum energy for free electrons.

7.3 What is a polaron

A polaron is another quasiparticle which describes the distortion of the ionic lattice due to a negative electron moving the lattice of ions about. The whole system of distortion and electron can be thought of one particle moving about the lattice with a large effective mass.

7.4 Electron-hole Interaction

Can think of an electron that has just jumped out into the conduction band of metal as having a state that is bound to the hole state it just created. This is an exciton which is a bound state and thus it has decreased energy compared to how it would otherwise be thought of. This is why it is thought that fitting to the fundamental band gap from DFT is better than that of the optical band gap from experiment as the latter includes an electron-hole interaction.

7.5 Stoner Magnetism

For stoner magnetism in metals it requires a high density of states at the fermi level, a small atomic distance so the electron can hop and change spin such that there is spin polarisation and that the d band is partially filled.

The reason why the electron stays in is that there is essentially more energy that can be gained from hopping. This effectively creates bonding and antibonding for the \uparrow and \downarrow spin states. Where the \uparrow states are lower in energy. This means that the density of states is higher in the \uparrow than in the \downarrow . The Fermi level equilibrates too.

The stoner criterion:

$$I \times \text{DoS} > 1$$

then spontaneous polarisation of unpolarised state $P = 0$ becomes non-zero where

$$P = (N_{\uparrow} - N_{\downarrow})/N$$

An electron can flip spin. Then the exchange force acts like a dam and allows for imbalance.

7.6 Friedel Oscillations

This is the quantum mechanical analog to the screening of a charge. Think of a charge and then loads of opposite point charges coming around to screen it. High density of opposite charges at surface and then further away the density decreases due to screening.

Quantum mechanically, this is what happens when some defect is in the material. This is due to electrons trying to screen the defect, but because electrons are extended in space, this means that there are *oscillations*, such that there is oscillations of positive and negative charge that surround the defect as ripples.

7.7 Surface States

These states decay away into the vaccuum (they have an imaginary value of k). This means that they live in a gap of energies. Think of a gap of k^2 against energy and think of a gap between $k < 0$ and $k > 0$. This is where the surface states live.

7.8 Mott -Hubbard Insulator

Think of coulomb force as U , and the bandwidth as W .

$U = 0$, Expect in a metal that the d band is localised around the fermi level. This is the density of states.

Increasing $U/W = 0.5$, start to have splitting where the d band spreads out.

Increasing more, have two smaller peaks on either side of large peak of DOS at the fermi level. $U/W = 1.2$

Increasing to $U/W = 2$ then the two peaks have separated and there is now an insulating state.

As the value of the coulomb interaction is greater than the bandwidth we expect both Vanadium and Titanium dioxide to have mott insulating behaviour. This means that it is a spin fluctuation rather then charge fluctuation dominated system.

To have reliable O₂ molecule maybe we need spin polarisation?

7.9 Notes on reversible integrators

[?]

Liouville operators are essentially a neat way of doing molecular dynamics. The dynamics are both reversible and preserves the volume in phase space (which is Liouville's Theorem)

Nosé-Hoover Thermostat is a canonical ensemble such that molecular dynamics can be achieved at a finite temperature. This acts to keep finite temperature by a velocity dependent force which acts like a drag to keep temperature constant.

There is not a hamiltonian associated with the system but there is a conserved quantity.

BOP can do molecular dynamics but it probably isn't as stable as in Tight binding.

7.10 Polarisable-ion Tight Binding

Monopole moments usually describe Mulliken charge, but there are extra contributions at play.

These are distortions of the spherical symmetry of the orbitals from electrostatic contributions. This actually adds charge by the expansion of the multipole moments which can contribute to the charge.

This theory couples orbitals of different angular momenta on the same site. This can break the degeneracy of the s, p and d orbital energies, inducing effects such as crystal-field splitting, lattice instability and ferroelectricity.

What couples these orbitals is the higher angular momentum components of the electrostatic potential.

This means that polarisability of the ionic valence electrons can occur.

The $l = 4$ components in the expansion of the electrostatic potential lead to the splitting of d levels in cubic, octahedral or tetrahedral crystal fields in transition metal oxides.

7.11 Universal Tight Binding III

Titanium in this model only has 3d states in this model, as in TiO_2 , 4s states just into the conduction band by overlap with the oxygen 2s orbital.

Octahedral coordination makes crystal field splitting of the d band possible into its e_g (z^2 and $x^2 - y^2$) and t_{2g} (xy , yz and xz). e_g is higher in energy than the other states. Can think of the reason for this originating from the actual coordination of the atom, Think of the orbitals surrounded by the ligands and then think about the $x^2 - y^2$ orbitals being on the axis and being head on with coordination in comparison to xy orbitals which are the other way. This means that the ligands can get closer to the t_{2g} orbitals thus lowering their energy.

Due to the d orbitals, within tight binding, due to the multipole expansion of the operators in a self-consistent polarisable-ion tight binding theory, we can class it as quadrupole, Oxygen is a dipole (think of the p orbital)

(Oxygen 2s states hybridise with O 2p states in water). These 2p states are defined by the TiO_2 valence band.

Due to the ordering of the energy surfaces

$$E_{\text{surf}} : (110) < (100) < (001) \quad \text{for rutile}$$

$$E_{\text{surf}} : (101) < (100) < (001) \quad \text{for anatase}$$

it was said that the rehybridisation of the surface bonds had been replicated successfully.

When oxygen is removed there are released electrons that occupy central states at nearby Ti.

These are in the band gap \Rightarrow localised

DFT underestimates atoms and too close to the conduction band which increases the delocalisation of the states.

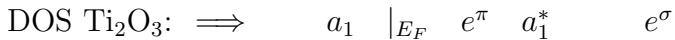
((Tamm state is tight binding state of surface. Potential barrier due to work function of the metal.. Crystal electron motion is restricted if in the forbidden zone)))

7.12 Electronic structure of reduced titanium dioxide

TiO_2 and Ti_2O_3 DFT calculations [?].

Distortion of Ti_2O_3 octahedra further split the t_{2g} states into bonding and antibonding subbands

Two spin degenerate bands in the a_{1g} manifold (containing one electron per Ti atom below the Fermi level) and the four spin-degenerate bands in the e_g^π manifold above the Fermi level.



Removal of a bridged oxygen means that the charge of the oxygen is redistributed.

Mott insulating behaviour.

In non-stoichiometric TiO_2 have conduction band mainly composed of 3d Ti with some 2p O.

Ti_2O_3 has although formally one electron per Ti atom is donated into the conduction band, these are delocalized and there is no spin moment generated. Ti_2O_3 is probably, however, a Mott-Hubbard insulator because of the narrowness of the a_{1g} band split off by the crystal field due to the distorted octahedra in the corundum structure. In the case of the reduced surface, the 3d manifold is rather broader and correlation effects are probably negligible.

This means that oxides of Ti are not dominated by electron correlation.

7.13 Calculation of TiO_2 Surface and Subsurface Oxygen Vacancy by the Screened.

Li, Huo and Robertson [?]

Vacancies by removal of bridging oxygen seem to be stable on the surface of rutile.

GGA \implies

- decreased d electron localisation
- decreased lattice distortion at vacancies
- decreased depth of defect states in band gap

We would want defect states that are in the band gap as they are localised near the conduction band edge of the original insulator.

Screened hybridised functional is able to get the right band gap and the right lattice constant and the deep vacancy state in the band gap.

O vacancies cause polaronic effects for small supercells. So large supercells are necessary to reduce the defect-image interaction.

$\text{V}_\text{O}''+1$ state distorts the lattice which leaves one electron localised on one of the three adjacent sites. Assuming here to be titanium.

In Ti rich environment, the stoichiometry is not $\text{Ti}:\text{TiO}_2$, but it is actually $\text{Ti}_2\text{O}_3:\text{TiO}_2$.

Rutile surface had stable vacancies from the removal of bridging oxygen.

Dangling bonds on surface create gap states in the band gap of bulk TiO_2 .

Removal of oxygen (or some oxygen vacancy I guess) creates 2 unpaired electrons (for a neutral defect) and dangling Ti bonds.

The oxygen vacancy creates localised on near Ti mainly and NN Ti, unlike in TiO_2 bulk.

7.14 Surface and subsurface oxygen vacancies in anatase TiO_2 and differences with rutile.

Cheng Selloni [?]

O vacancy diffusion pathways in Anatase show low energies for diffusion between surface and subsurface, which is why generally vacancies are subsurface in Anatase. This means a high equilibration rate at typical annealing temperatures, which is why vacancies aren't seen on surface.

7.15 A model for time-dependent grain boundary diffusion of ions and electrons through a film or scale, with an application to alumina

Tauschig and Finnis [?]

Electrostatic potential method with reaction equations and concentrations of vacancies. Derivation of power laws associated with concentrations of vacancies. Derived from nominal charges on the ions, not ones from DFT (this was a classical system). Schottkey equilibration between electrons and holes at surface of the interface and also equilibrium constant,. Found that the permeation rate of oxygen was proportional to the difference in vacancy concentration due to the pressure difference between both sides of the film Surface charge dominates the reaction at large times and large pressure

7.16 Bayesian Optimisation Spiel

In Bayesian optimisation, the method is to use the data from all previous inputs to find the most likely point that will minimise the objective function. It does this by use of Bayes' Theorem, where one can think of these t's as target quantities such as your elastic constants, bulk modulus and so forth.

Using this optimisation within a gaussian process regression framework we have the freedom to recreate objective function surfaces that can be composed from an infinite number of basis functions.

So, using bayes theorem we can estimate the probability of obtaining a particular target quantity given the data obtained from the previous target quantities. This data is encapsulated by the use of a kernel function, which measures correlations between sets of data which makes this C matrix.

We can see that this probability of obtaining the next target quantity is predictive, so it depends on the value of the next input vector that you feed it \mathbf{x}_{N+1} . These input vectors are the actual parameters that you give to the model, so it would be the bond integrals and the pair potentials for the tight binding model.

So for each of the target quantities we can calculate some expected value given a particular point in parameter space. From this we can calculate the expected improvement, if we used this point for the next sample from the parameter space.

So, doing this for all input quantities, we can calculate the total expected improvement for all quantities. The maximum of this is the next best point to sample.

7.17 Additional BOP Notes

Chapter 8

TBE to BOP Ti

8.1 Ti Dimer

To check that the functions for the BOP are working as expected I investigated the case of a Ti dimer.

There is good agreement between the pair potential and the binding energies. Differences in E_{band} between tbe and BOP can be attributed to the differences in definition between the two models. The binding energy can be defined as

$$\begin{aligned} E_{\text{bind}} &= E_{\text{band}} - E_{\text{atom}} + E_{\text{pair}} \\ &= E_{\text{bond}} + E_{\text{prom}} + E_{\text{pair}}. \end{aligned}$$

But one can have different definitions for the promotion energy:

$$\begin{aligned} E_{\text{prom}}^{\text{BOP}} &= \sum_{RL} q_{RL} H_{RLRL} - N_{RL} H_{RLRL}^0 \\ E_{\text{prom}}^{\text{tbe}} &= \sum_{RL} (q_{RL} - N_{RL}) H_{RLRL}^0, \end{aligned}$$

So the band energy definition has to be changed to accommodate the differences.

Now using Tony's parameters for Titanium in his Spanjaard-Desjonquieres format in the BOP.

In figure 8.1 there are small discrepancies between the pair potential energy, which also manifest themselves into differences between the binding energy. This is solely down to unit conversion.

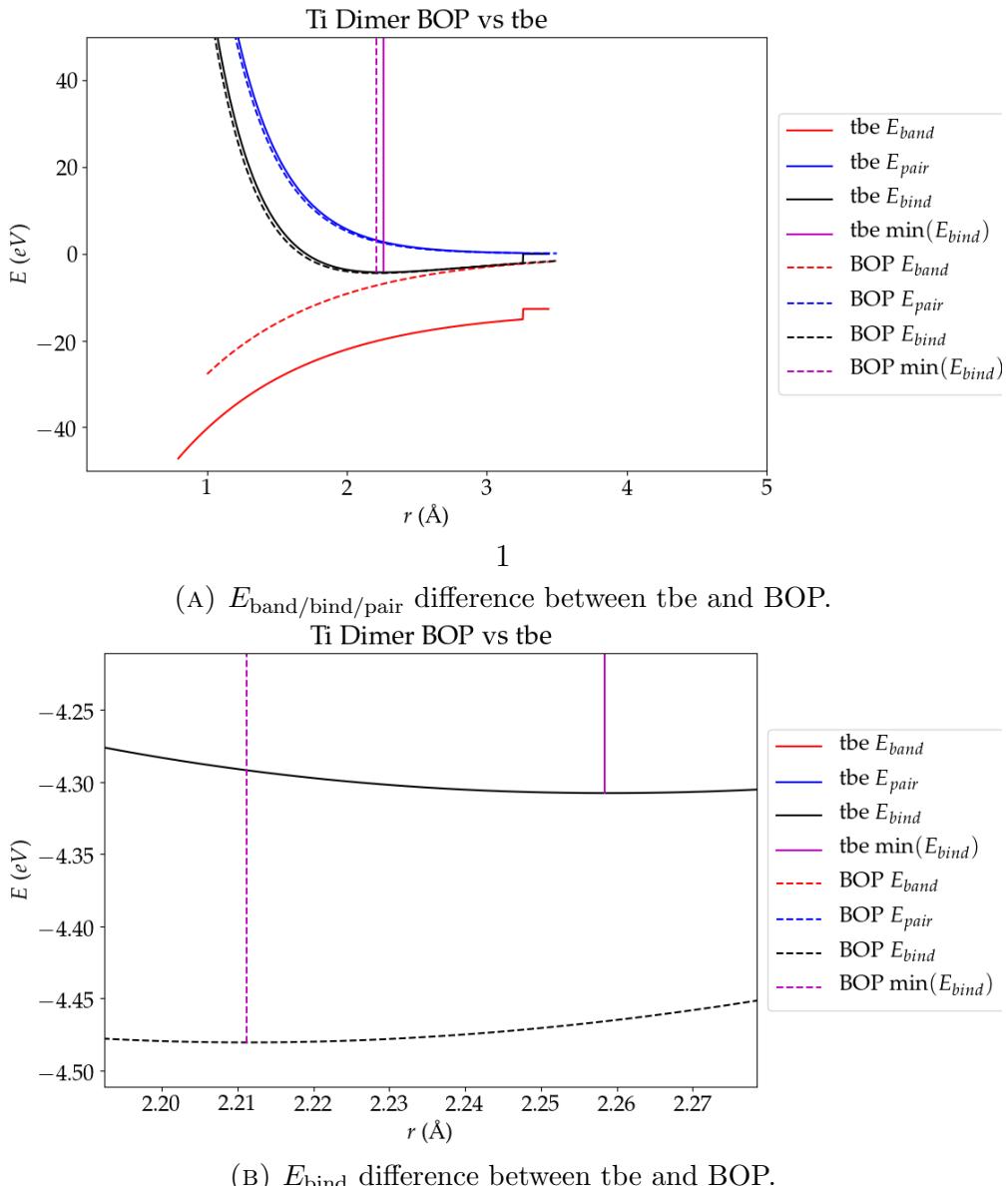


FIGURE 8.1: Comparison between BOP and tbe with the number of recursion levels set to 5.

Parameter	tbe	BOP
r	2.2584 \AA	2.211 \AA

FIGURE 8.2: Values of r in \AA at the minimum of the binding energies obtained from tbe, using Tony's 5 parameter tbe model, and BOP, with $n_{\text{rec}} = 5$.

8.2 Elastic and lattice constants in BOP

To find the elastic constants and lattice constants, we revert back to the bulk case.

There are 15 applied strains from which we can obtain the elastic constants and the bulk modulus. The code used was developed by Alexey Girshick, and it assumes that the minimum of the binding energy curve against applied strain is at zero deformation (which is correct for a structure that is right), however, as can be seen in the figures below, this was not observed. If the structure is not at a minimum, pressure terms are added into the effective elastic constants of the lattice, as shown by Wallace.

8.2.1 Four recursion levels

For $n_{\text{rec}} = 4$ it was seen that some of the elastic constants turned out to be negative even with minimised lattice parameters.

Lattice parameters were found to be:

Parameter	Value at minimum of E_{bind}
a	2.857142
c	4.4
c/a	1.54

FIGURE 8.3: Table of minimum lattice parameters from BOP using Tony's 5 parameter tbe model, with $n_{\text{rec}} = 4$.

This c value is at the lower bound of the grid search used to get the lattice parameters, from this it is reasonable to assume that the structure wants to collapse inwards due to insufficient repulsion from the pair potential.

C_{ij}	EC in 10^{11}Pa . $n_{\text{rec}} = 4$		
	tbe min lattice par	BOP min lattice par	exp.
C_{11}	1.967	2.423	1.761
C_{12}	0.493	0.856	0.868
C_{13}	0.588	0.698	0.682
C_{33}	2.060	3.072	1.905
C_{44}	-0.088	0.272	0.508
C_{66}	-0.668	-4.738	0.450
K	1.037	1.377	1.101
R	0.705	1.103	0.618
H	0.729	0.859	0.489

FIGURE 8.4: Table of elastic constants obtained from BOP using Tony's 5 parameter tbe model at the minimum lattice parameters derived from the BOP code and from tbe, with the number of recursion levels set to 4.

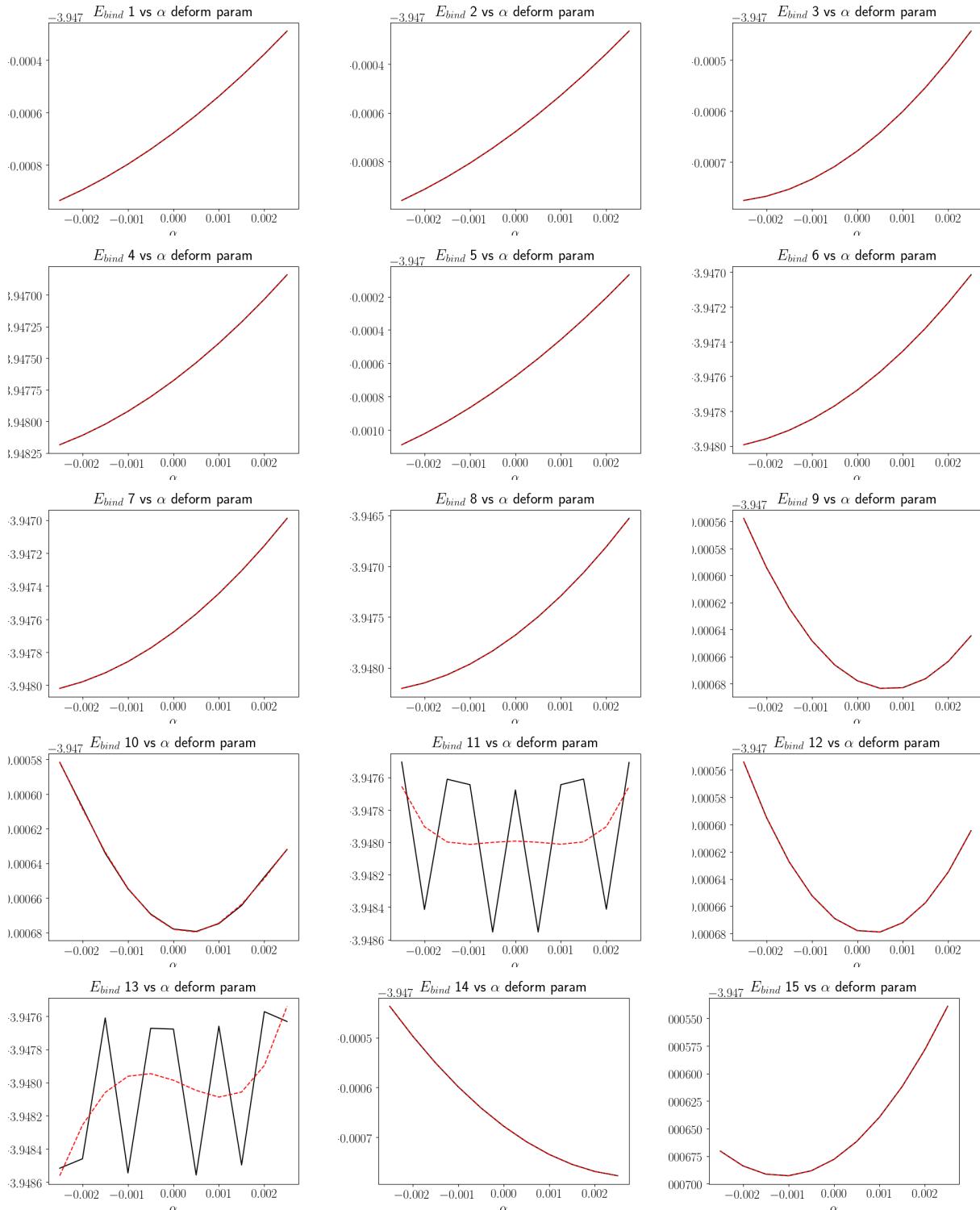


FIGURE 8.5: Binding energy curves from BOP using Tony's titanium parameters with $n_{\text{rec}} = 4$ using the minimum parameters obtained from minimisation in tbe. Energy is in eV

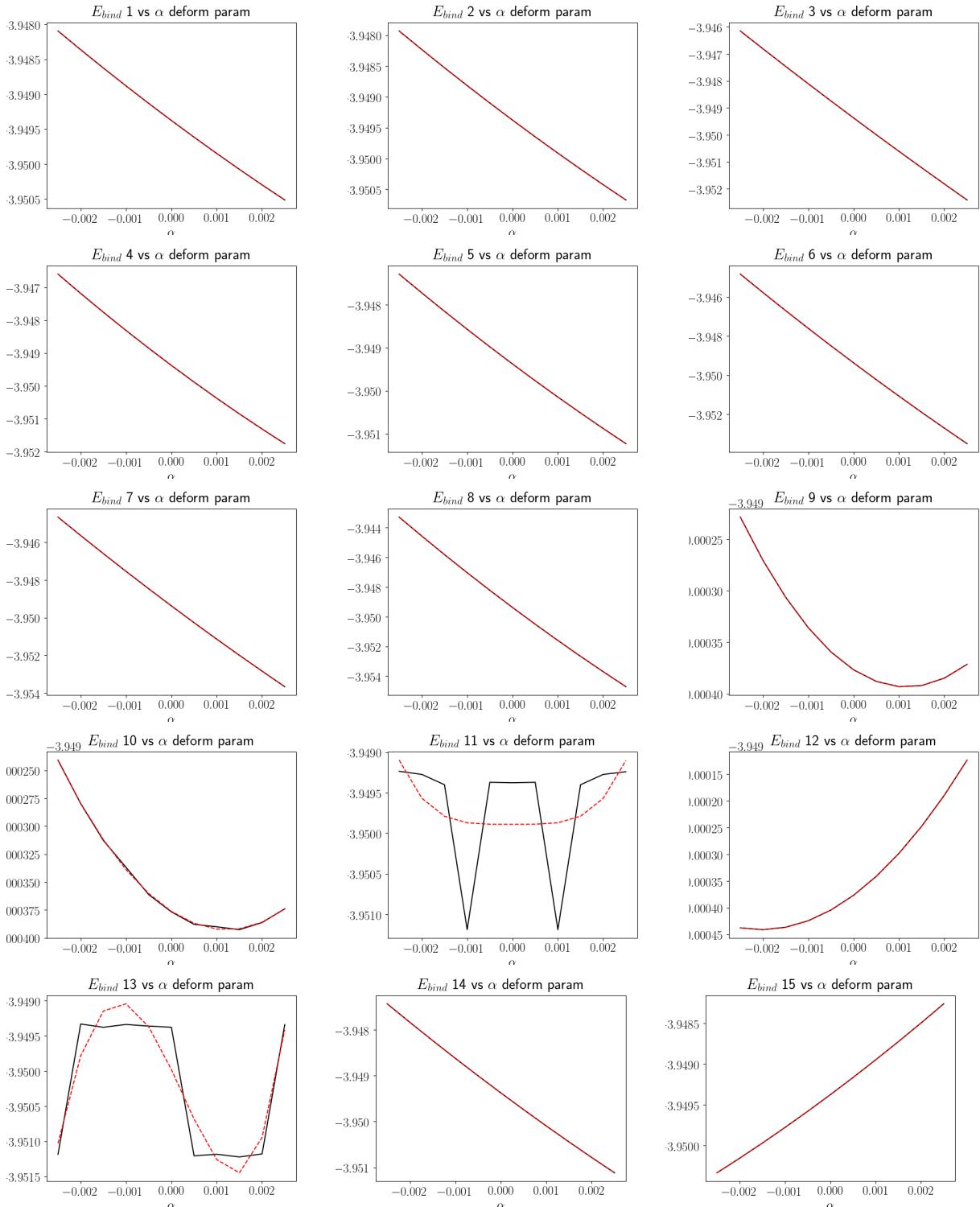


FIGURE 8.6: Binding energy curves from BOP using Tony's titanium parameters with $n_{rec} = 4$ using the minimum parameters obtained from minimisation in BOP code. Energy is in eV

8.3 Five recursion levels

Parameter	Value at minimum of E_{bind}
a	2.8684210
c	4.52631578
c/a	1.57798165

FIGURE 8.7: Table of minimum lattice parameters from BOP using Tony’s 5 parameter tbe model, with $n_{\text{rec}} = 5$.

C_{ij}	EC in 10^{11} Pa. $n_{\text{rec}} = 5$		
	tbe min lattice par	BOP min lattice par	exp.
C_{11}	2.163	2.636	1.761
C_{12}	0.390	0.570	0.868
C_{13}	0.527	0.669	0.682
C_{33}	2.120	3.045	1.905
C_{44}	1.028	-5.456	0.508
C_{66}	2.239	-1.799	0.450
K	1.038	1.349	1.101
R	0.782	1.104	0.618
H	0.876	1.079	0.489

FIGURE 8.8: Table of elastic constants obtained from BOP using Tony’s 5 parameter tbe model at the minimum lattice parameters derived from the BOP code and from tbe, with the number of recursion levels set to 5.

For $n_{\text{rec}} = 5$ one obtains.

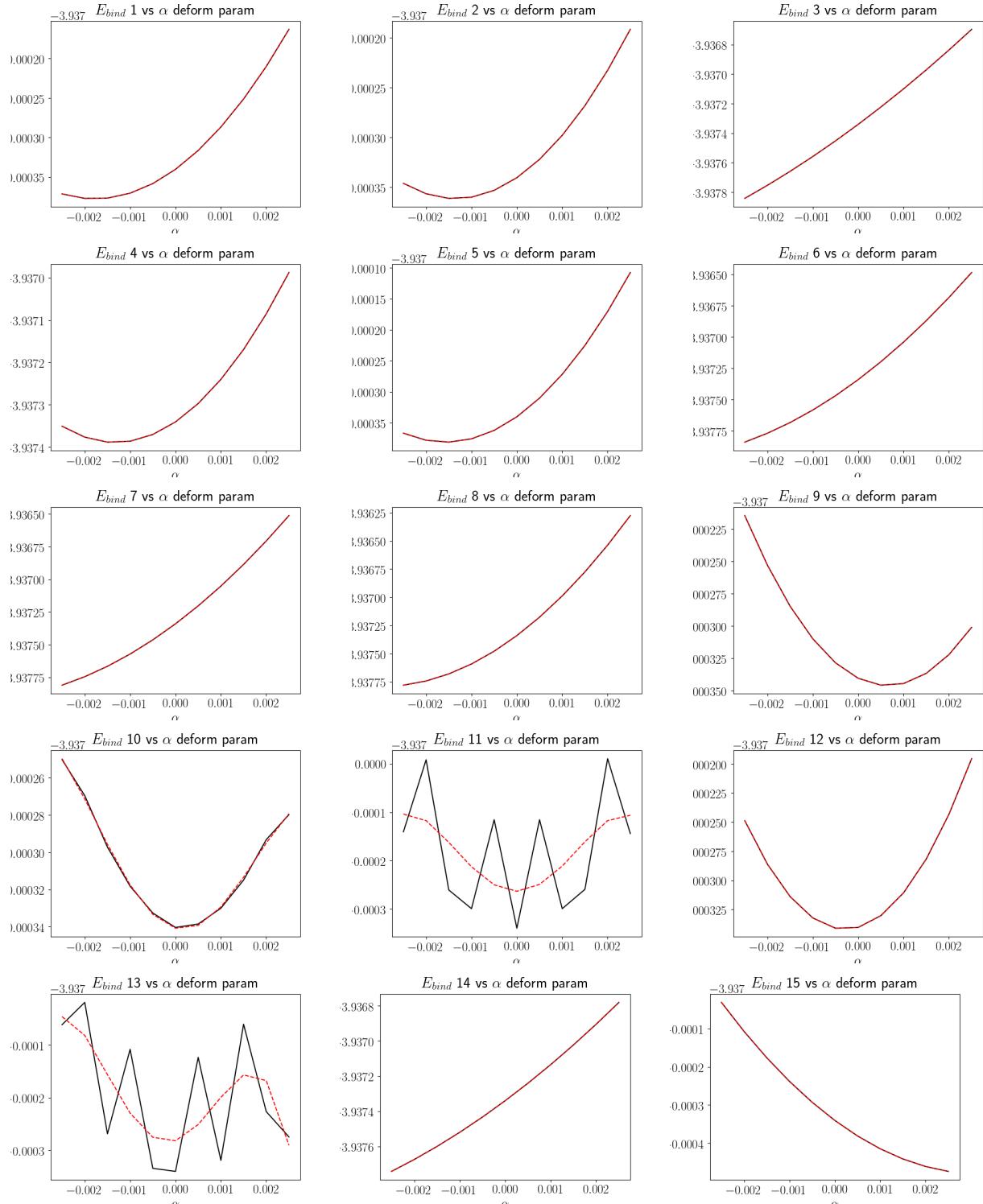


FIGURE 8.9: Binding energy curves from BOP using Tony's titanium parameters with $n_{rec} = 5$ with minimised the lattice parameters. Energy is in eV

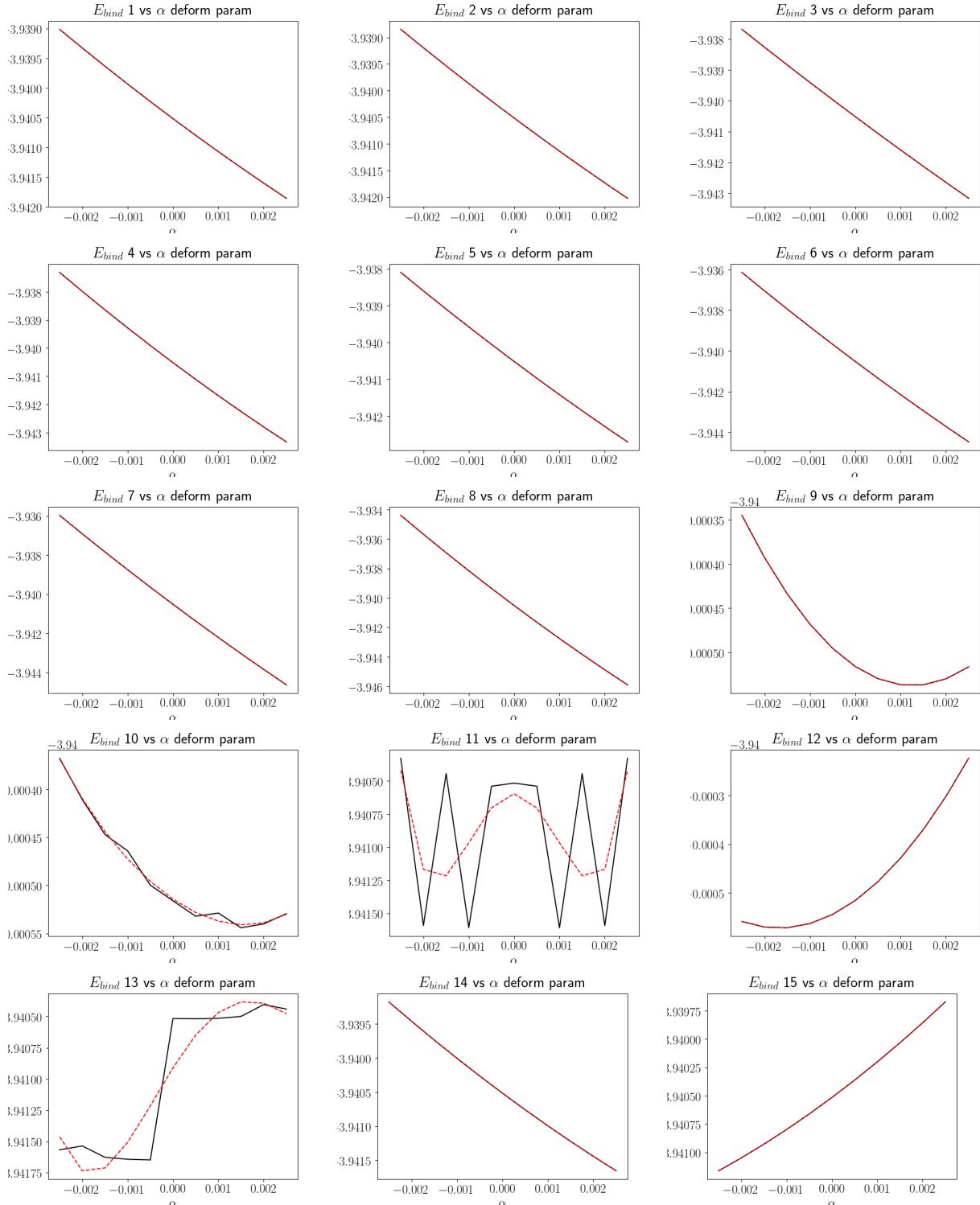


FIGURE 8.10: Binding energy curves from BOP using Tony's titanium parameters with $n_{\text{rec}} = 5$ with minimised bop lattice parameters. Energy is in eV

8.4 Six Recursion Levels

For $n_{\text{rec}} = 6$:

Parameter	Value at minimum of E_{bind}
a	2.868421
c	4.526315
c/a	1.577981

FIGURE 8.11: Table of minimum lattice parameters from BOP using Tony's 5 parameter tbe model, with $n_{\text{rec}} = 6$.

C_{ij}	EC in 10^{11}Pa . $n_{\text{rec}} = 6$		
	tbe min lattice par	BOP min lattice par	exp.
C_{11}	*	62.565	1.761
C_{12}	15.413	*	0.868
C_{13}	-5.989	10.377	0.682
C_{33}	*	*	1.905
C_{44}	-6.114	22.305	0.508
C_{66}	-2.245	10.772	0.450
K	-1.863	4.175	1.101
R	4.578	11.678	0.618
H	*	16.024	0.489

FIGURE 8.12: Table of elastic constants obtained from BOP using Tony's 5 parameter tbe model at the minimum lattice parameters derived from the BOP code and from tbe, with the number of recursion levels set to 6.

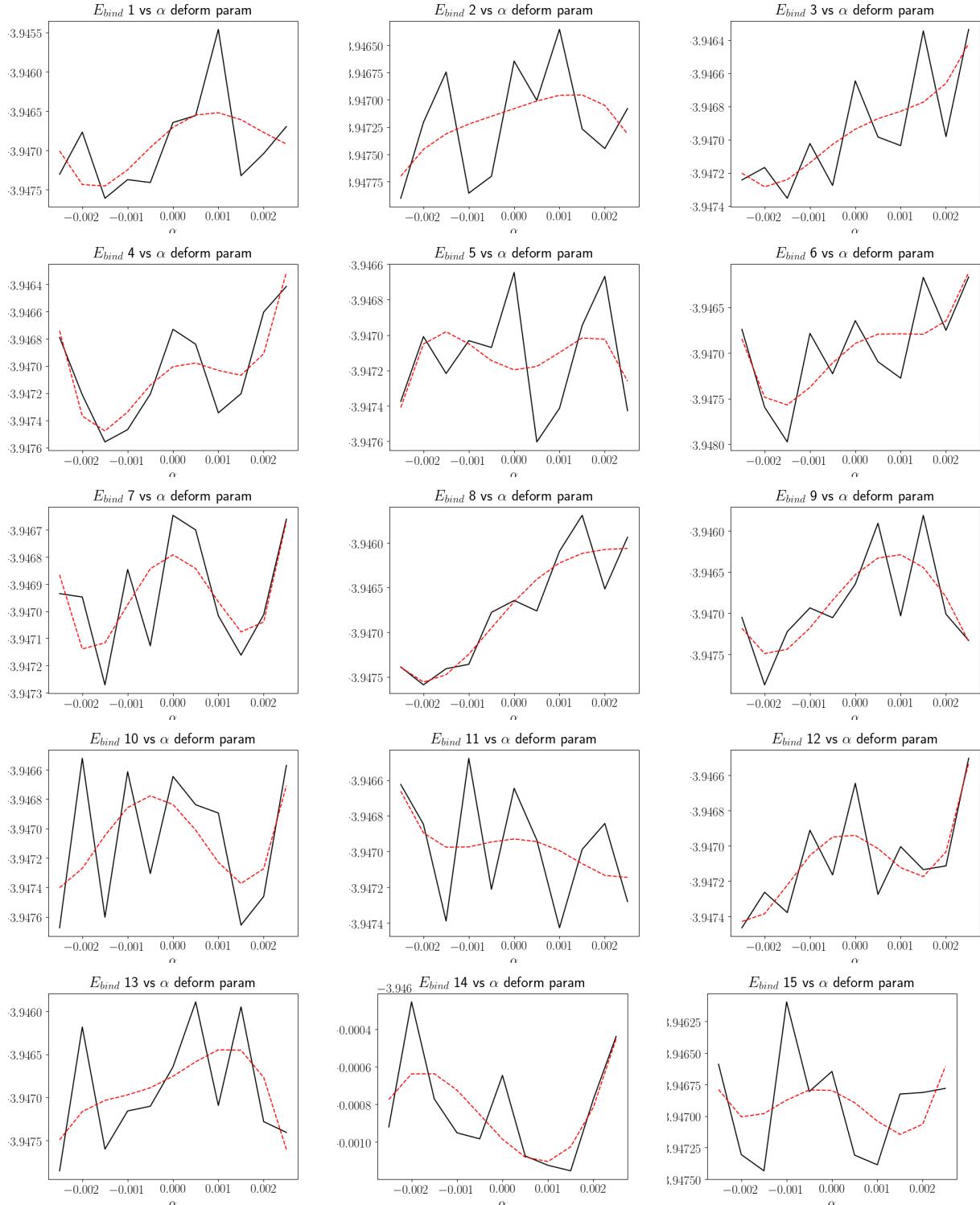


FIGURE 8.13: Binding energy curves from BOP using Tony's titanium parameters with $n_{\text{rec}} = 6$ with minimised tbe lattice parameters. Energy is in eV

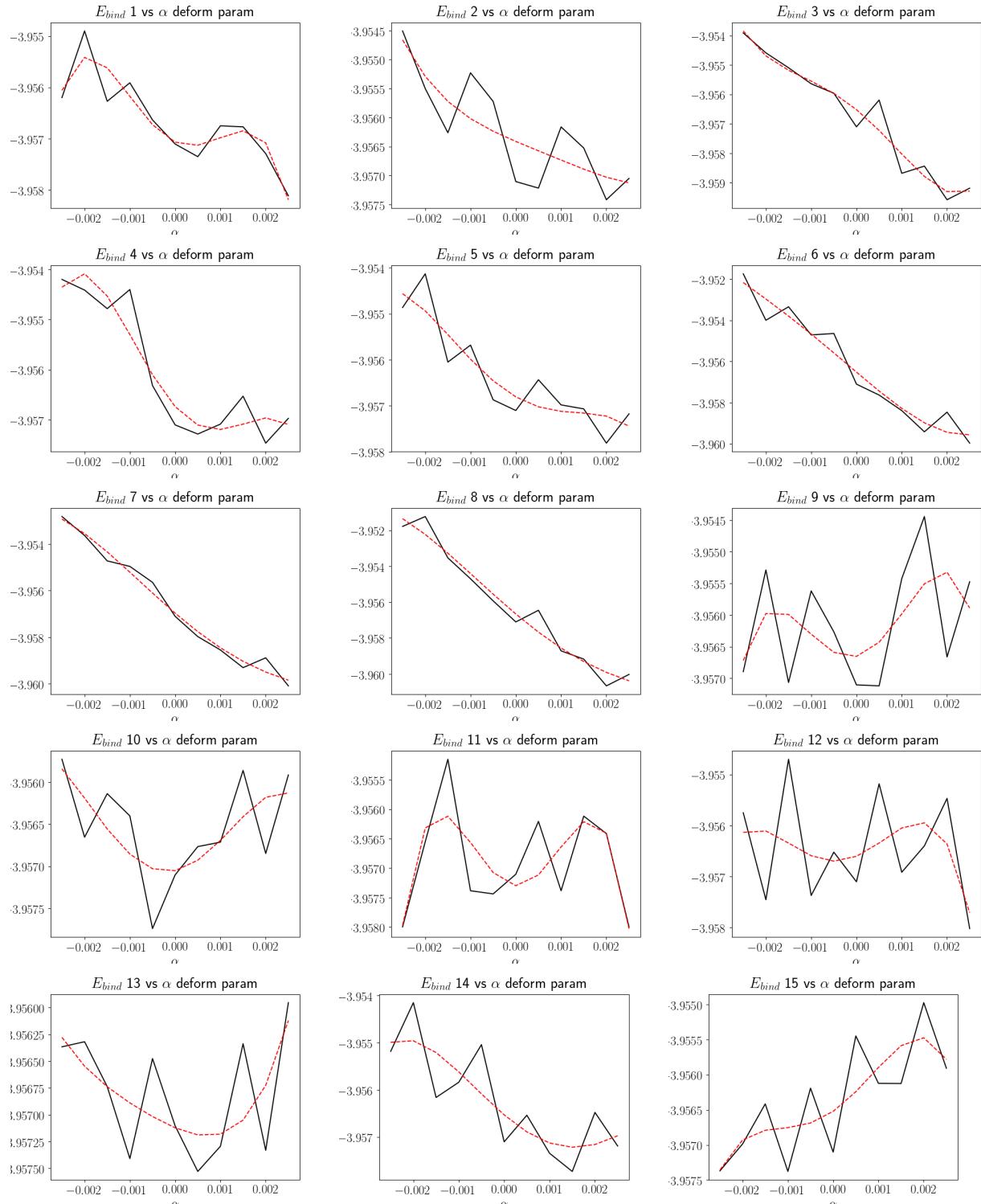


FIGURE 8.14: Binding energy curves from BOP using Tony's titanium parameters with $n_{\text{rec}} = 6$ with minimised bop lattice parameters. Energy is in eV

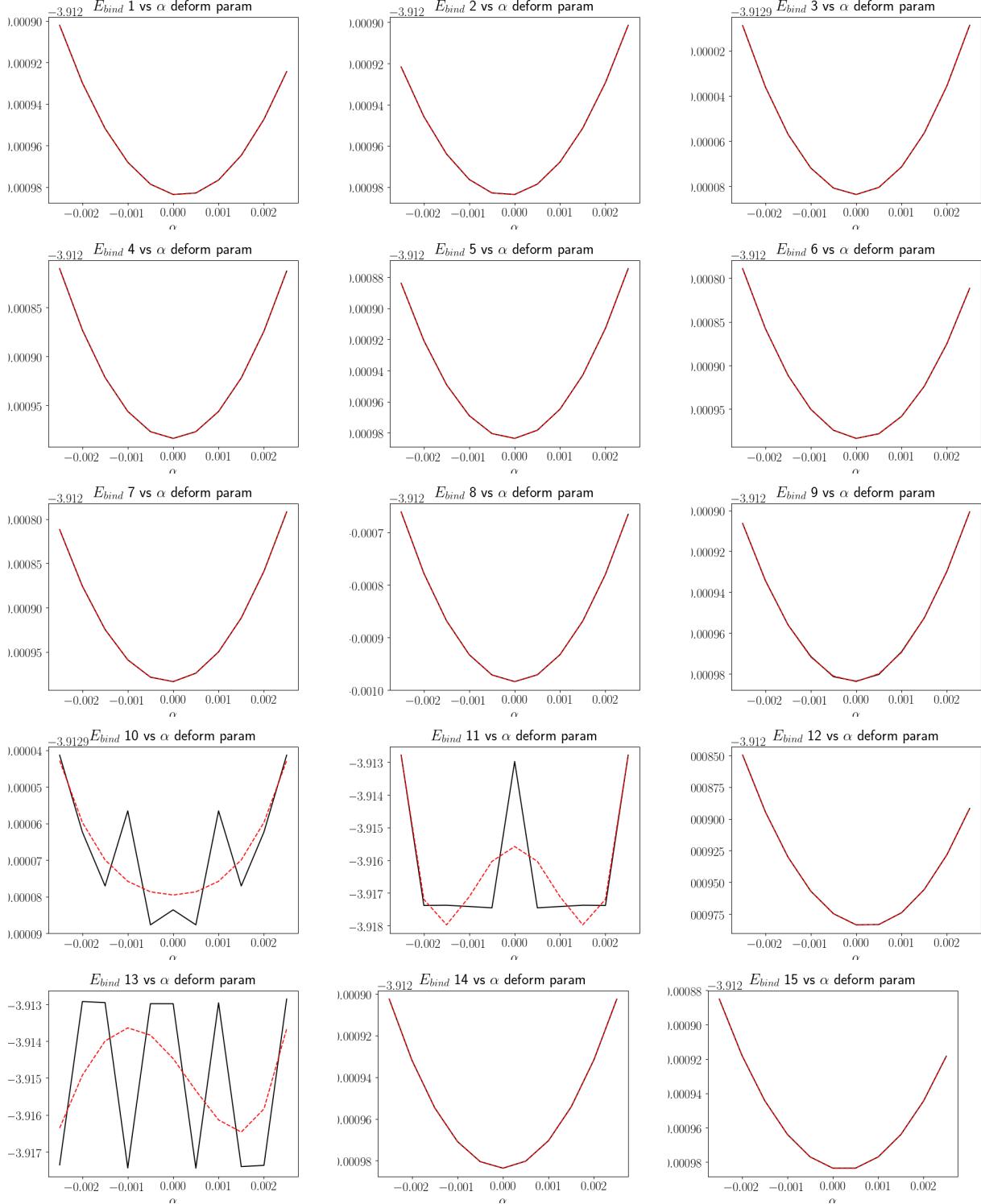


FIGURE 8.15: Binding energy curves from BOP using Tony's titanium parameters with $n_{\text{rec}} = 6$ with minimised bop lattice parameters. Energy is in $eV\text{\AA}^{-3}$

8.5 Anisotropic Elasticity Solutions for Hexagonal material

For fun, I decided to code up the Stroh Formalism solution for anisotropic elasticity based on the work by Bartek. However I made it more general, such that one can input any elastic constant matrix and obtain results, not just for displacements, but also for strain and stress.

Weirdness at c a vs ebind when looking at nrec = 4 between

Return: $a_l \dots a_u = 2.906449052859575 \dots 2.9064502735627005$ $c_l \dots c_u = 4.652707926432292 \dots 4.652720133463543$

Reason was found to be that the r_{prune} was too small. As such this effectively truncated the number of moments necessary

8.6 Checks for other codes

$$P_5(x) = P_2(x)(x - r_2)^3$$

$$P_2(x) = a(x - r_1)^2 + b(x - r_1) + c$$

$$a = \frac{1}{(r_1 - r_2)^5} \left\{ \frac{1}{2} (r_1 - r_2)^2 f''(r_1) - 3(r_1 - r_2) f'(r_1) + 6f(r_1) \right\}$$

$$b = \frac{1}{(r_1 - r_2)^4} \left\{ f'(r_1) * (r_1 - r_2) - 3 * f(r_1) \right\}$$

$$\frac{1}{(r_1 - r_2)^5} x$$

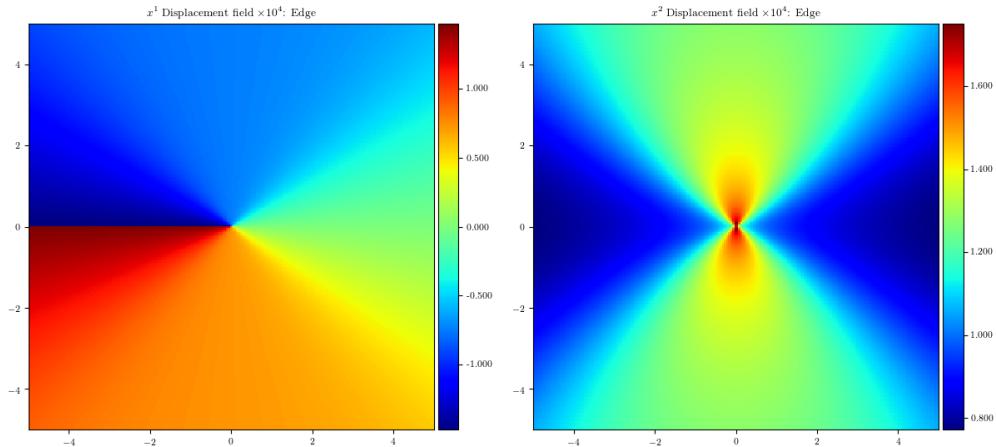
$$c = \frac{f(r_1)}{(r_1 - r_2)^3}$$

$$\mathbf{R}_I(t + \delta t) = 2\mathbf{R}_I - \mathbf{R}_I(t - \delta t) + \frac{f_I(t)}{M_I} \delta t^3 \mathcal{O}(\delta t^4)$$

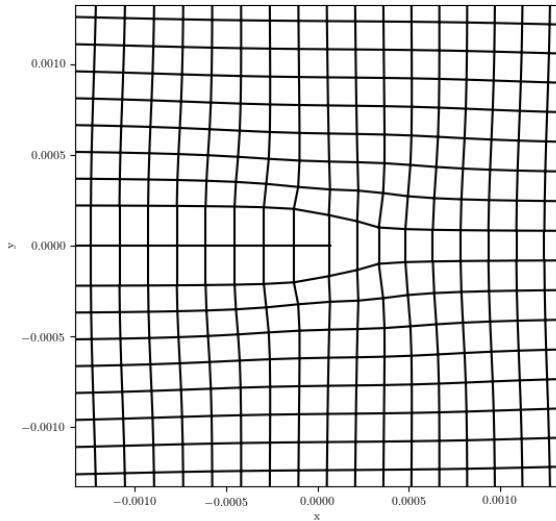
$$\dot{\mathbf{R}}_I(t) = \frac{1}{2\delta t} [\mathbf{R}_I(t + \delta t) - \mathbf{R}_I(t - \delta t)] + \mathcal{O}(\delta t^3)$$

$$\mathbf{R}_I(t + \delta t) = \mathbf{R}_I(t) + \dot{\mathbf{R}}_I \delta t + \frac{f_I(t)}{M_I} \delta t^3 \mathcal{O}(\delta t^3)$$

$$\dot{\mathbf{R}}_I(t + \delta t) = \dot{\mathbf{R}}_I(t) + \frac{1}{2M_I} [\mathbf{f}_I(t + \delta t) + \mathbf{R}_I(t)] + \mathcal{O}(\delta t^3)$$



(A) Displacement heatmap for x (left) and y (right) components of a pure edge dislocation calculated by solutions to the sextic equation by the Stroh Formalism using Ti elastic constants.



(B) Edge dislocation using displacements from above used to distort cubic lattice (to mimic the classic edge dislocation diagram).

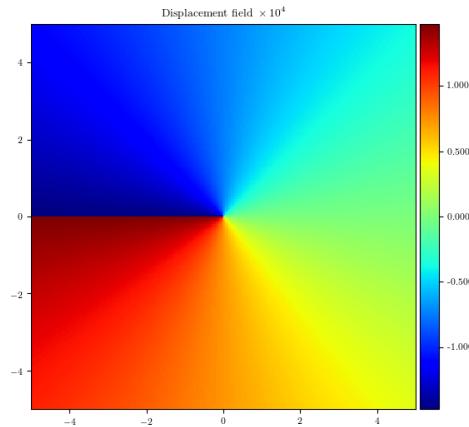
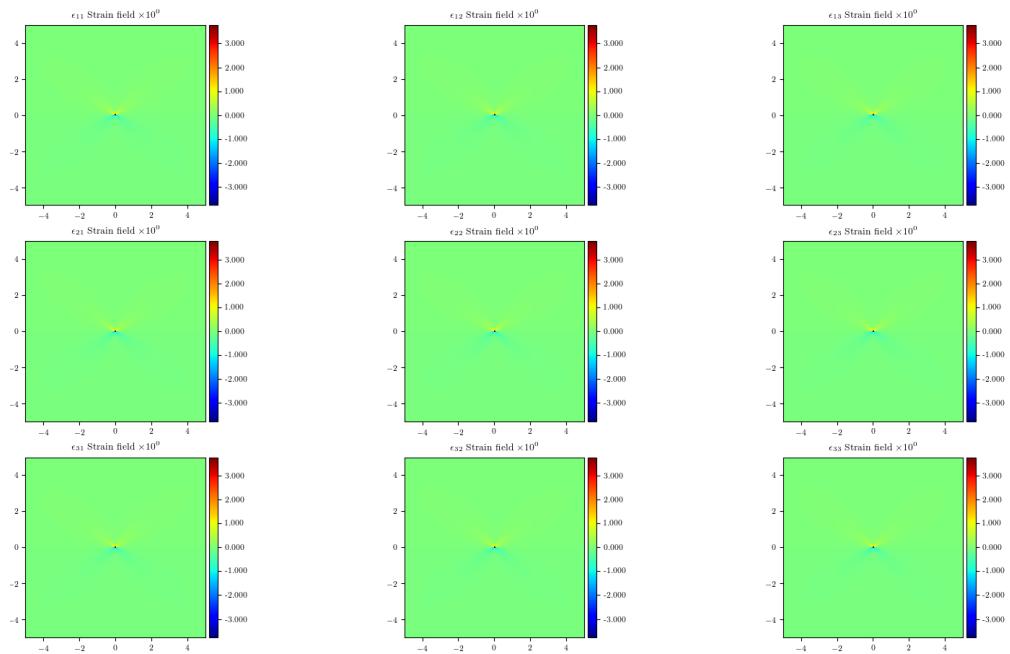
FIGURE 8.17: z displacement for a pure screw dislocation.

FIGURE 8.18: Strain field for a pure edge dislocation.

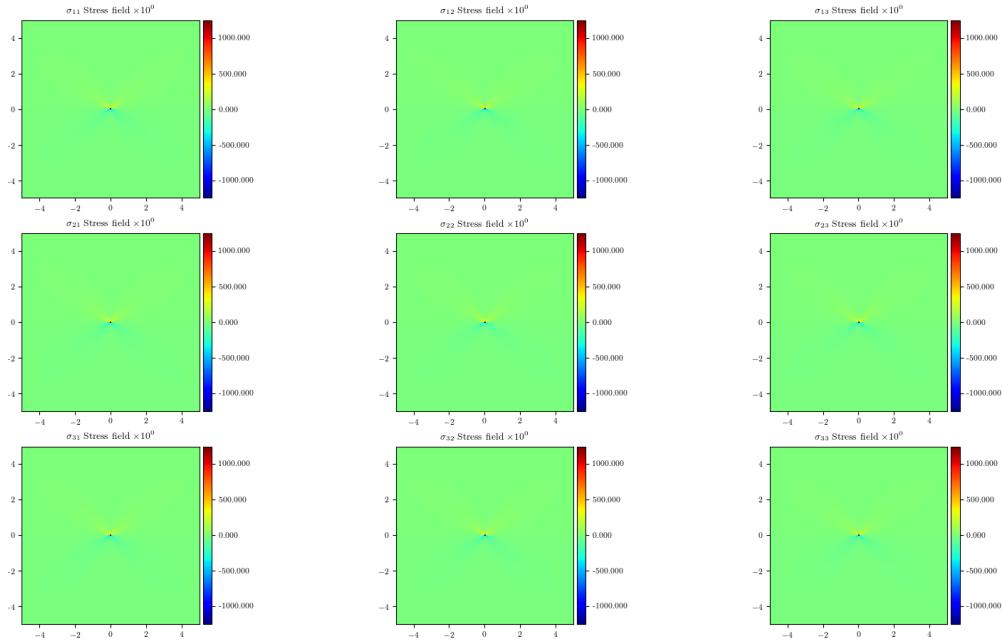


FIGURE 8.19: Stress field for a pure edge dislocation.

**** How to calculate the forces - Use the Hellmann-Feynman theorem. -

$$\mathbf{f}_I = \langle \psi_{\mathbf{R}}^0 | \frac{\partial}{\partial \mathbf{R}_I} \hat{H} | \psi_{\mathbf{R}}^0 \rangle$$