Local Atomic Environments and the Reactivity of Clusters: A Theoretical Investigation

Yao Zhao

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Trinity 2024

To my parents.

Acknowledgements

First and foremost, I would like to express my profound gratitude to my supervisor, Professor John E. McGrady, for his invaluable guidance, support, and encouragement throughout the duration of this research. I have gained precious knowledge and skills under his expert mentorship.

I would also like to acknowledge Professor Aleksandar T. Staykov, who first introduced me to the field of theoretical and computational chemistry. Without his influence, I would not have pursued this degree.

My heartfelt appreciation goes to my mother, Lilin, and my stepfather, Tetsuro, for their unwavering love and encouragement. Their constant support has been a source of strength throughout this journey.

I am sincerely grateful to the Aso Group and New College for their generous financial support, which has been crucial in enabling my studies.

Finally, I am deeply thankful for the companionship of my friends, Aidan Manley, for our constant scientific discussions, Sourav Mondal, for our shared conference travels, and all the members of the group for their support. My DPhil journey would not have been possible without them.

Abstract

The reactivity of clusters is known to be strongly influenced by their local atomic environments, and that is the central concept that unites the different chapters in this thesis. I present three distinct theoretical investigations into the electronic structure and reactivity of clusters, covering different aspects of chemical behaviour. The first project concerns the structure of partially-ligated ruthenium carbonyl clusters, and the mechanism of their reactivity in the gaseous phase. The patterns of stable configurations of ligands are identified, and the link between the number and distribution of unsaturated sites on reactivity will be discussed. The second project looks at catalysis from the opposite extreme of unsaturation, specifically for a Ru_3-based 'single-cluster catalyst' and its interaction with a surface support. The electronic structures of several models for this system are examined and the significance of subtle details in the model on the stability will be demonstrated. The final project concerns structural phase transitions and temperature effects on the stability of Zintl clusters in solids. The work emphasises the importance of understanding phonon density of states and the origin of changes of entropy when considering the thermodynamics of phase transitions. While all of the research described in this thesis is motivated by experimental observations, the work itself is done purely with theoretical and computational techniques, the aim being to enhance the information content of the experiments. Such understanding is crucial for formulating universal principles of designing advanced catalysts such as single-cluster catalysts.

The work in Chapter 5 have been published in the following paper:
Thermodynamics of phase transitions in Zintl clusters from density functional theory: making and breaking of bonds in Ba_3Ge_4 [1]
Y. Zhao and J. E. McGrady, *Phys. Chem. Chem. Phys.*, 2024, **26**, 7318
DOI: 10.1039/D3CP05713E

Contents

List of Figures

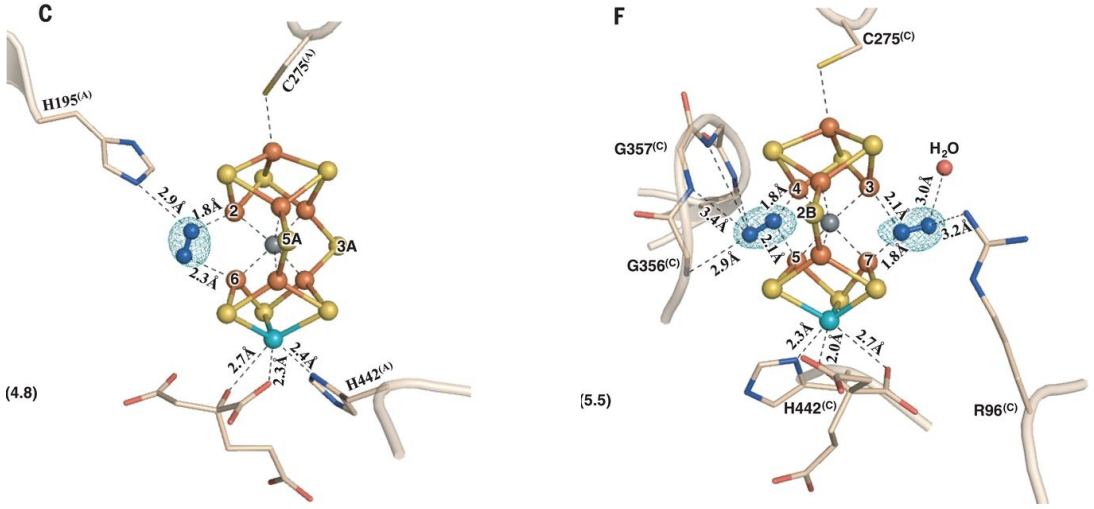
Chapter 1

Introduction

1.1 Cluster chemistry

Clusters are a diverse class of chemical compounds that are typically composed of multi-atom aggregates of a size intermediate between molecules and nanoparticles. Examples of clusters come from many different regions of the periodic table, with a wealth of different types of chemical bonds. The International Union of Pure and Applied Chemistry (IUPAC) defines a cluster as 'a number of metal centres grouped close together which can have direct metal bonding interactions or interactions through a bridging ligand, but are not necessarily held together by these interactions.' [3] Examples that meet this definition can be found in transition metal carbonyl clusters, transition metal halide clusters and as well as bio-inorganic cofactors in nitrogenase. The term can also extend to groups of main-group elements like fullerenes, boranes and Zintl clusters. In terms of properties, we can usefully differentiate clusters from nanoparticles in so much as the latter retain some physical and chemical characteristics of their bulk counterparts, while clusters show unique properties. In the cluster size regime, even minor changes in the number and composition of atoms can significantly impact a wide range of characteristics including reactivity, magnetism, photoluminescence, and properties.

Perhaps the most widely studied application of cluster chemistry is in catalysis, where the presence of multiple metal centres can open up new pathways that are not available to their single-site counterparts. A high profile example from nature is the Fe_7MoS_9C cluster found in the enzyme nitrogenase, which catalyses the reduction of nitrogen (N_2) to ammonia (NH_3) [4], a crucial process for life's essential building blocks. While the mechanistic details of the catalytic reaction by nitrogenase are still not clear, a wide range of reaction pathways have been proposed in the literature. The first type of reaction pathways involves intermediates with terminally-bonded,



(a) Singly substituted cluster (b) Doubly substituted cluster

Figure 1.1: Experimentally reported structures of FeMo cofactor in nitrogenase during the nitrogen fixation process. Here, sulfide ligands from the parental cluster are (a) singly or (b) doubly substituted with nitrogen molecules. The presence of multiple metal atoms gives rises to various complex bonding modes of nitrogen, which are not typically found in the mononuclear transition metal complexes. Due to its complexity in structures, the mechanisms of the nitrogen reduction by nitrogenase are still a central topic of debates in bioinorganic chemistry. From Reference [7]. Reprinted with permission from AAAS.

or 'end-on', nitrogen molecules on an active centre, a structure commonly found in transition metal dinitrogen complexes. The nitrogen molecule then may be catalysed to be hydrogenated in either the distal pathway, in which the terminal nitrogen atom is reduced completely to ammonia first, or the alternative pathway, in which both nitrogen atoms are hydrogenated alternatively. There also has been a proposal of a new pathway, called 'enzymatic', that involves intermediates with η^2-bonded, or 'end-on', nitrogen molecules, which are unusual in mononuclear complex species but can be found in bimetallic complexes or metal surfaces. Availability of multiple-metal centres adds another complexity of possible bonding modes of nitrogen molecules and provides new reaction pathways. This 'enzymatic' pathway of nitrogen fixation has inspired several studies exploring the potential of artificial clusters as catalysts in ammonia production [5, 6].

In the laboratory, synthesis of clusters and investigations of their reactivity began with small-sized clusters in the gaseous phase, using experimental techniques such as mass spectrometry, laser ablation, and sputtering. These fundamental studies have enhanced our understanding of the link between structure and reactivity, for exam-

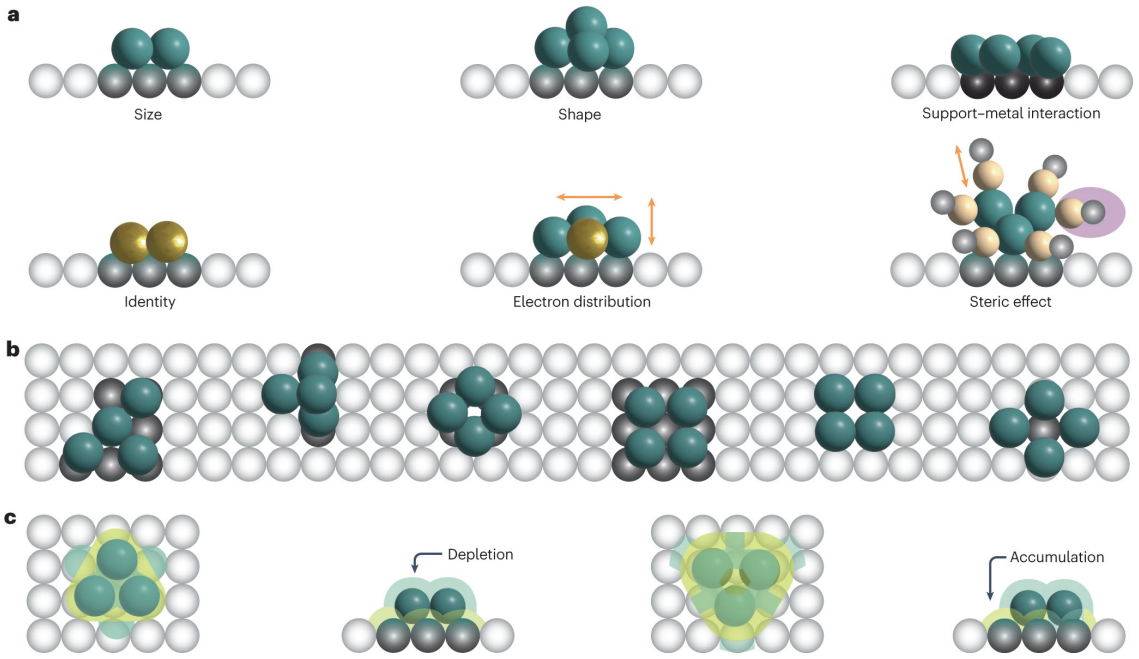


Figure 1.2: Structures and properties of supported clusters on substrates, summarised by Li *et al.* [9]. (a) A list of various factors that changes the properties of supported clusters on substrates, including size, shape, support-metal interaction, identity of the atom, and presence of ligands. (b) Illustrations of different possible configurations of tetramer cluster on substrate. (c) Charge transfer effects between clusters and substrates, as a result of interaction. Adapted with permission from Reference [9], Springer Nature.

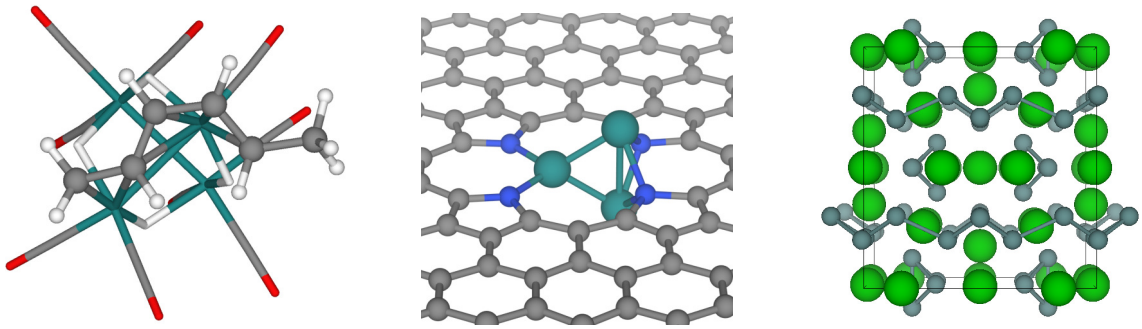
ple in the study of the reactivity of gold clusters with pre-adsorbed hydrogen towards molecular oxygen bond scission, despite gold's chemical inertness in bulk form [8]. The study found that the highly exposed structure of gold clusters and under-coordination of atoms allowed the steady adsorption of both hydrogen and oxygen, and the supply of charges from the cluster to antibonding orbitals on oxygen molecules facilitates the breakage of O-O bonds. However, 'naked' clusters (*i.e.* those without stabilising ligands) generated in this manner also often exhibit high reactivity and instability, prone to aggregation into larger, catalytically inactive species, limiting their practical industrial applications. The concept of 'supported' clusters on substrates has emerged as a means to stabilise the clusters while still retaining a sufficient degree of unsaturation to enable them to serve as active sites for catalysis. The combination of clusters and substrates has led to the development of a very wide range of new heterogeneous catalysts, many of which have gone on to find applications in the chemical industry.

Perhaps unsurprisingly, the interactions between clusters and substrates can significantly influence their reactivity and catalytic activities. The support can be con-

sidered in some sense as a multidentate ligand, saturating some of the vacant coordination sites of the cluster and reducing its activity. These cluster-support interactions become increasingly pronounced as the cluster size decreases, and covalent bonding and/or electrostatic interactions can lead to significant structural reconstructions at the interface [10]. For instance, Pt_4 clusters have been shown to retain a tetrahedral structure on a MgO (100) surface but adapt a square planar configuration on TiO_2 (110) [11], and this structural change has a profound impact on the availability of bonding sites for adsorbates. These geometric effects are driven by electronic factors including charge transfer, electron confinement, and orbital rehybridisation that can make the properties of the supported cluster very different from its gas-phase parent. The above comments highlight the importance of identifying accurately the local atomic environments of clusters and the structure-property relationships that underpin catalytic properties. Key features of the local atomic environment of supported clusters include the coordination number, the identity of neighbouring atoms, the long-range order of substrates, the presence of other ligands, and the nature of the reactive media [9, 12]. Experimental techniques such as scanning transmission electron microscopy (STEM) and X-ray absorption near edge structure (XANES) have provided vital insight into substrate-supported catalysts, but precise measurements are often limited by complications such as tip convolution effects and signal averaging. The identification of active species therefore remains a considerable challenge to experiment. In such circumstances, computational modelling and simulation has a vital role to play in aiding the interpretation of experiment as well as identifying key features of electronic structure and catalytic behaviour of the active sites.

1.2 Thesis outline

The aim of increasing our understanding of the electronic structure and properties of clusters in varying states of unsaturation is the underlying theme that links the chapters of this DPhil thesis. In Chapters 3-5, I will present three distinct computational studies which focus on the stability and reactivity of both transition-metal and main-group clusters in different local atomic environments. Each project is built upon experimental observations of the compounds that examine aspects of their structural or catalytic behaviour. Throughout the majority of the thesis, DFT serves as the computational method of choice, but higher levels of theory are also utilised when DFT proves insufficient for accurate energy calculations.



(a) Chapter 3 (b) Chapter 4 (c) Chapter 5

Figure 1.3: An overview of a list of studies presented in this thesis. (a) Chapter 3 studies the reactivity of partially ligated tetra-ruthenium hydrido carbonyl cluster anions in the gaseous phase. (b) Chapter 4 studies the electronic structures of supported Ru_3 on a carbon nitride matrix. (c) Chapter 5 studies the thermal effects on the solid-state reactions of clusters in the Zintl phase Ba_3Ge_4. The local atomic environments being investigated in each chapter are (a) coordination number and geometry, (b) species of supports, and (c) packing structures of ions in crystal, respectively.

The first project concerns the reactivity of partially ligated tetra-ruthenium hydrido carbonyl cluster anions in the gaseous phase. The experiments, conducted by Professor Scot McIndoe and his team at the University of Victoria, Canada, have shown that the activation of the anionic cluster $[H_3Ru_4(CO)_{12}]^-$ *via* collision-induced dissociation (CID) leads to reactions with a variety of molecules in the gas phase including alkenes, arenes, alkanes, silanes, and oxygen. In particular, a significant increase in the reactivity toward 1-hexene was observed when precisely three or more carbonyl ligands were removed from the cluster. As is the case with all mass spectrometry, the experiments give information on its mass (and therefore composition) but not on structure, so a key task is to establish geometries of clusters that correspond in composition to the peaks in the mass spectrometry. This provides important insight into the ways that clusters respond to increasing degrees of electronic and coordinative unsaturation. We also propose the reaction mechanism of 1-hexene dehydrogenation with these unsaturated species, rationalising the distinctive sensitivity of reactivity to the number of carbonyl ligands. The potential role of metal-metal bonds in catalysis is also discussed. While all these reactions take place in the gaseous phase, and the active species are close to the saturated limit, this study highlights the importance of coordination number and geometry, all of which are crucial to the understanding of the reactivity of supported clusters.

The second project is a closely-related one, focusing on the electronic structures of supported Ru_3 on a carbon nitride matrix. This project is based on an experimental report by Ji *et al.* [13] and co-workers who showed that Ru_3 single-cluster catalysts (SCCs) embedded on CN matrices can be synthesised through pyrolysis, a common technique to synthesise graphene-based single-atom catalysts. The catalyst was reported to be 23 times more efficient than the single-atom catalyst (SAC) analogue in alcohol oxidation. However, even with the access to XANES data, the microscopic details of the active sites remain elusive. The initial task for the computational project was therefore again to identify the structure of substrate surrounding the cluster. We analyse the electronic structures of different defect models on N-graphene and compare the cluster-substrate interactions. The stability of each models is considered, as is the influence of binding of oxo ligands. Finally, a feasible reaction mechanism of alcohol oxidation with oxo species is proposed. This project highlights how very small details in the chosen model can change the results of calculations profoundly, and also the ongoing challenges in modelling SCCs due to their significantly increased complexity of structures compared to SACs.

The final project concerns the thermal effects on the solid-state reactions of clusters in the Zintl phase Ba_3Ge_4, where naked clusters $[Ge_4]^{6-}$ are surrounded and stabilised by cations. This project was published in the journal *Physical Chemistry Chemical Physics* [1]. The title compound has been shown to undergo a structural phase transition in which half of the clusters are polymerised to form a 1-dimensional network, $^{1}_{\infty}[Ge_4]^{6-}$. These structural modifications are reminiscent of bond-stretching isomerism, a subject which has been a long-standing interest for many computational groups. By calculating phonon frequencies and employing the quasi-harmonic approximation for Gibbs energy, we identify the origin of entropy difference between two phases. We also predict that the existence of new phase p-tetragonal, where all of clusters are polymerised in two directions, may also be accessible under specific thermodynamic conditions, although theis has not yet been realised. The Ba_3Ge_4 phase is also compared to the isomorphic compound Ba_3Si_4 where no such polymerisation has been observed. We also examine another reported example of bond-stretch isomerism in Zintl phases, in this case the calcium doped Sr_3Sn_5. Our calculations suggest that the experimental report may in fact be an artifact in this case. Overall, the studies reported in this chapter highlight how temperature and pressure can influence the environments of supporting ions, and drive significant rearrangements of the covalent bonds within and between clusters.

As the chemistry in each of the three main projects is rather different, a detailed introduction to the relevant literature will be presented separated at the start of each chapter. However, it is useful to consider here some of the important computational and theoretical models that underpin this thesis, and are relevant to all three chapters.

1.3 Theoretical models

1.3.1 Frontier orbital theory

Frontier orbital theory [14–16], first formulated by Kenichi Fukui and Roald Hoffmann, provides a fundamental tool for interpreting the electronic structure and reactivity of metal clusters, in particular those that are saturated with ligands such as transition metal carbonyl clusters. By analysing orbitals available for bonding on metal centres, the isolobal analogy can be used to predict the structures of metal-metal bonds in a cluster. To illustrate this important concept, first consider the molecular orbital diagram of an octahedral transition metal complex ML_6 in Figure 1.4a. In this diagram, the symmetry-adapted linear combinations of σ lone pair electrons on ligands interact with orbitals on the metal centre to form 6 metal-ligand bonds. The t_{2g}, or equivalently d_{xy}, d_{xz} and d_{yz} orbitals on metal centre, are non-bonding unless there are π-symmetry orbitals on the ligands. If the metal centre has 6 metal electrons, this will lead to a stable closed-shell and electron-precise 18-electron configuration. Removing one ligand from this complex will lead to a ML_5 configuration, the molecular orbital diagram for which is shown in Figure 1.4b. In this case, 5 metal orbitals interact strongly with ligands and one low-lying orbital a_1 is now left untouched, on top of t_{2g} (or e and b_2) orbitals. The shape of this 'frontier' orbital (Figure 1.4c) points outward from the metal centre, and can be considered to be analogous to the singly-occupied orbital of a CH_3 radical. Therefore, we can treat the ML_5 complex like a building block in organic chemistry that forms a bond by overlapping with frontier orbitals on other fragments. In a similar manner, additional frontier orbitals can be generated by removing 1 or 2 more ligands from this ML_5 fragment, resulting into 2 and 3 frontier orbitals on ML_4 and ML_3 fragments respectively (Figure 1.4d, 1.4e). By combining multiple ML_n fragments, it is then possible to construct a larger cluster, as depicted in Figure 1.5.

The principle of combining frontier orbitals of different fragments leads directly to polyhedral skeletal electron pair theory which has been applied extensively to the prediction of the geometries of transition metal clusters [17–19]. Wade's rules [17] were formulated based on the observation that an n-vertex polyhedron cluster requires

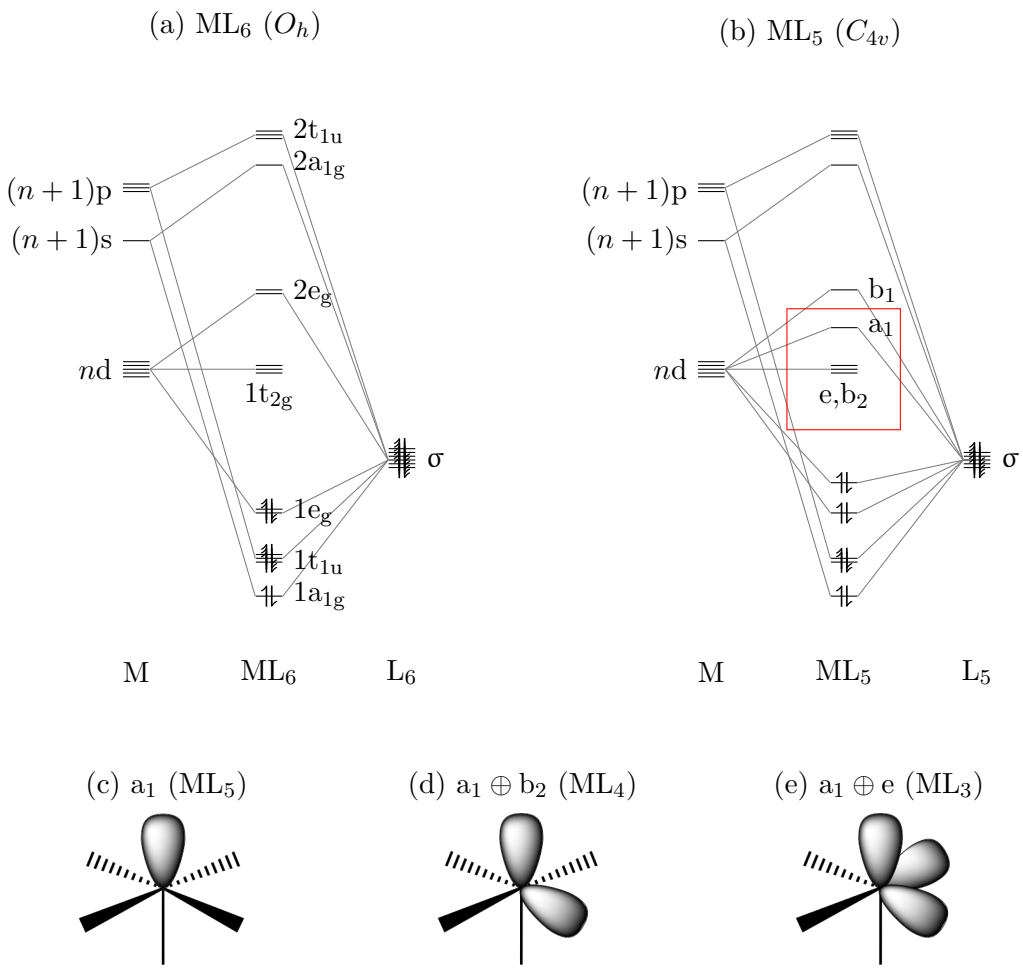


Figure 1.4: Molecular orbital diagram of (a) ML_6 and (b) ML_5 complexes and the frontier orbitals of (c) ML_5, (d) ML_4 and (e) ML_3 fragments. The red box in (b) indicate a list of frontier orbitals that are available to form covalent bonds with ligands or other fragments. The symbols in (c,d,e) indicate the irreps of hybridised frontier orbitals.

(a) tetrahedrane (b) $Ir_4(CO)_{12}$ (c) $Fe_4C(CO)_{13}$

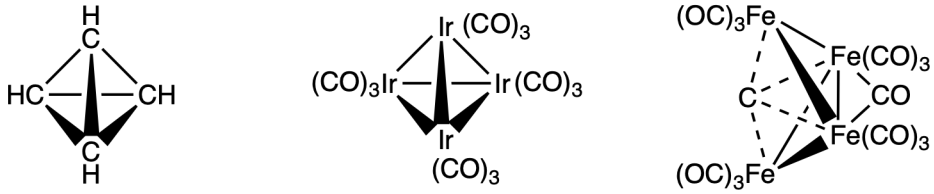


Figure 1.5: Comparison of structures of tetrahedrane, $Ir_4(CO)_{12}$ and $Fe_4C(CO)_{13}$. The $Ir(CO)_3$ fragment is isolobal to CH fragment. Both tetrahedrane have $Ir_4(CO)_{12}$ have 6 skeletal electron pairs and adapts a tetrahedron structure, while $Fe_4C(CO)_{13}$ has one additional electron pairs and adapts the butterfly shape.

$(n+1)$ 'skeletal electron pairs' to fill all molecular orbitals with bonding characters, for n greater than 4, (a tetrahedral cluster can accept 6 skeletal electron pairs, $(n+2)$, to fill all bonding molecular orbitals). Clusters that satisfy this skeletal electron pair counting are known to adapt *closo*-type structures, deltahedron shaped with all faces being equal. One or two additional electron pairs to these skeletal electron counts will cause opening up of the cluster, resulting in so-called *nido* and *arachno* species respectively. The number of pairs would be equal to the number of skeletal electron pairs in parental *closo* n-vertex species. Hence, *nido* and *arachno*-species adapt the parental *closo* structure with one or two vertices missing respectively. For example, both tetrahedrane and $Ir_4(CO)_{12}$ in Figure 1.5a and 1.5b contain 6 skeletal electron pairs and adapt *nido*-structure, while $Fe_4C(CO)_{13}$ in Figure 1.5c contains 7 skeletal electron pairs, adapting a butterfly shape. Wade's rules have been successful in predicting the structures of clusters, and also extended to predict structures of clusters with condensed polyhedrons [19].

More generally, frontier orbital theory has also been vital in describing the electronic structures of organometallic clusters and bridging the fields of inorganic and organic chemistry. Arguments based on frontier orbital theory can also be extended to the surfaces of metals in the context of chemisorption, as for example in the Blyholder model [20] and the cluster-surface analogy, although their validity remains a topic of debate [21].

1.3.2 First-principles models

Whilst the qualitative models described in the previous paragraph are useful for giving insights, accurate calculations of energies are essential for useful predictions

of the thermodynamic stability and reactivity of molecules, and their dependence on temperature and other environmental effects. To make quantitative arguments on the catalytic activities of clusters, the solution of the Schrödinger equation in some approximate form that takes into account all effects including those from the local atomic environments is unavoidable. The theoretical tools and *ansätze* used in this thesis will be discussed in detail in the next chapter. The energetic predictions from first-principles modelling can be converted into predictions on the performance of catalysis, including turn-over frequencies (TOF), using kinetic theory.

Gas-phase experimental data have been much valued by computational chemists, particularly in the early literature, because they can be directly compared to theoretical investigations without investing heavily in modelling the surrounding environment. With the help of the developments of modern electronic structure theories such as density functional theory (DFT) and the rapid advances in the computer hardware, accurate calculations of energies for medium-size systems are possible and indeed routine today. However, running calculations for liquid-solid interfaces remains a substantial challenge. In such circumstances, it is common to employ 'computational' models to reduce the size of the problem in order to suppress the computational efforts. There are mainly three categories of models in first-principles computation of heterogeneous catalysts:

1. cluster approaches

2. embedded cluster approaches

3. periodic supercell approaches [22].

Amongst these, the cluster approach is usually the cheapest computational model, and it links directly to the important concept of the 'cluster-surface analogy'. In this perspective, molecular clusters that resemble a portion of the surface of a solid are constructed with a size that is large enough to take account of the effects of the extended system. While the cluster models are capable of capturing characteristics of covalent bonding, long-range electrostatic interactions are necessarily missing. In the embedded cluster approach, classical point charges or interatomic potentials are added on top of the cluster model to mimic these long-range interactions caused by the ionic lattice. To model the surface of metallic systems effectively, a periodic supercell approach is usually needed to capture the properties of the delocalised electrons. With a supercell that is large enough to separate the periodic images of molecules

(adsorbates), the infinite periodicity of crystal structures is retained. A further complication is the presence of defects: under the experimental conditions, the surface of the solid is far from pristine, and the presence of defects and other adatoms might play important roles in catalysis. Modelling of defects at realistic concentrations would require a very large unit cell, with obvious cost implications. Whatever approach is taken to construct the computational model, it is crucial to evaluate the convergence of properties with respect to the number of atoms or the size of the supercell.

It is also important to note that the electronic structure method employed *per se* such as DFT can significantly influence the outcomes of calculations with first-principles models. Indeed, most of the approximations in quantum chemistry are developed by simplifying the physical processes in quantum mechanics, which can be justified with reasons. However, these approximations can be *ad hoc* in nature and one has to be careful in choosing an appropriate method for the system of interest. In the next chapter, more details on the physics of electronic structures, and their relation to various quantum chemistry methods will be discussed.

Chapter 2

Theory and Methodology

In this chapter, we present the fundamental principles of quantum theory essential in chemistry [23], and the practical computational methods for electronic structures [24, 25] employed in this thesis. We start by exploring the complexities of many-body problems in quantum mechanics. Then, in the following sections, various methods for approximating wavefunctions and the role of *ansätze* for electronic wavefunctions will be discussed. Additionally, several common numerical techniques such as basis sets employed in practical calculation will be presented. Finally, we provide a brief introduction to the theory of phonons [26, 27], the vibrational dynamics, and thermodynamics of solids. This chapter sets the stage for the subsequent results presented in this thesis.

2.1 Many-body problems in quantum mechanics

2.1.1 Many-body wavefunctions

A wavefunction for a single particle can be described as a vector in a Hilbert space,

$$|\psi\rangle \in \mathcal{H}. \tag{2.1.1}$$

For example, a single particle in the three-dimensional real space is described by a complex-valued wavefunction, $\psi(\mathbf{r})$, where $\mathbf{r}$ denotes the position of that particle. The many-body wavefunction for N distinguishable particles $|\Psi\rangle$ then is described by a tensor product of N Hilbert spaces,

$$\mathcal{H}^{(N)} = \bigotimes_{i=1}^{N} \mathcal{H} \tag{2.1.2}$$

and the real-space representation of the wavefunction is described by a function of N coordinates, $\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_N)$ where $\mathbf{r}_i$ denotes the position of ith particle.

Consider the case of a wavefunction for N identical particles, given by $\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ where $\mathbf{x} = \{\mathbf{r}, \omega\}$ is a spatial-spin coordinate. Since particles are indistinguishable in quantum mechanics, these two wavefunctions represent the same quantum state,

$$\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_N) = e^{i\theta}\Psi(\mathbf{x}_1, \ldots, \mathbf{x}_i, \ldots, \mathbf{x}_j, \ldots, \mathbf{x}_N). \tag{2.1.3}$$

in which the coordinates of particles i and j are exchanged. In the three-dimensional space, the phase factor $e^{i\theta}$ happens to only be either +1 or -1. Particles whose wavefunction is symmetric under an exchange of two particles are called bosons (particles with integer spin), and particles whose wavefunction is antisymmetric under an exchange of two particles are called fermions (particles with half-odd-integer spin).

2.1.2 The many-body Schrödinger equation

According to the postulates of quantum mechanics, a non-relativistic quantum system is governed by the time-dependent Schrödinger equation,

$$i\hbar\frac{\mathrm{d}}{\mathrm{d}t}\left|\Psi(t)\right\rangle = \hat{H}\left|\Psi(t)\right\rangle \tag{2.1.4}$$

where i is an imaginary unit, t is time, and $\hat{H}$ is the Hamiltonian operator. When the Hamiltonian is independent of time, a solution to the equation 2.1.4 can be separated into spatial and time components, $\left|\Psi(t)\right\rangle = e^{-iEt/\hbar}\left|\Psi\right\rangle$, where the energy, E, is an eigenvalue in the time-independent Schödinger equation,

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle \tag{2.1.5}$$

and the quantum state that satisfies this condition is called a stationary state. In the majority of chemical problems, we are interested in finding the solutions to equation 2.1.5 as its eigenstates that define the properties, such as the energy and densities, of many-body systems, although finding the exact solutions is not possible for most cases.

Any observable quantity associated with the eigenstate can be calculated with the expectation value of an operator, given as

$$\langle\hat{O}\rangle = \frac{\langle\Psi|\hat{O}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{2.1.6}$$

The energies of wavefunctions that are not necessarily exact eigenstates of the equation 2.1.5 can also be measured using the expectation value of the Hamiltonian,

$$E = \langle\hat{H}\rangle = \frac{\langle\Psi|\hat{H}|\Psi\rangle}{\langle\Psi|\Psi\rangle} \tag{2.1.7}$$

which offers an upper bound to the exact ground-state energy. Hence, the eigenstates of the Hamiltonian are located at the minimum or saddle points of this expression of energy functional, if one considers Ψ as a variable function of some parameters. By using the method of Lagrange multipliers, these eigenstates can be found by varying the wavefunction and solving the variational equation

$$\frac{\delta}{\delta\Psi}\Big[\langle\Psi|\hat{H}|\Psi\rangle - E\left(\langle\Psi|\Psi\rangle - 1\right)\Big] = 0 \tag{2.1.8}$$

where the constraint of normality of the wavefunction is applied. Finding the ground eigenstate of a given Hamiltonian using such an approach is called the variational principle.

The Hellmann-Feynman theorem rewrites an expression for the derivative of the total energy and can be derived by using the variational principle. Suppose that one is interested in finding the derivative of total energy with respect to a parameter λ (e.g., nuclear positions, electron charge, etc.). Starting from equation 2.1.7, the general expression is given as

$$\begin{aligned}\frac{\partial E}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\langle\Psi|\hat{H}|\Psi\rangle \\ &= \left\langle\frac{\partial\Psi}{\partial\lambda}\middle|\hat{H}\middle|\Psi\right\rangle + \left\langle\Psi\middle|\frac{\partial\hat{H}}{\partial\lambda}\middle|\Psi\right\rangle + \left\langle\Psi\middle|\hat{H}\middle|\frac{\partial\Psi}{\partial\lambda}\right\rangle.\end{aligned} \tag{2.1.9}$$

If Ψ is the exact ground-state solution located at the stationary point of the energy functional, then the terms with any variations in the wavefunctions should vanish by definition. Therefore, given that the Hamiltonian has an explicit dependence of λ, only the term $\left\langle\Psi\middle|\frac{\partial\hat{H}}{\partial\lambda}\middle|\Psi\right\rangle$ survives in the expression.

2.1.3 The Born-Oppenheimer approximation

Consider a molecular system composed of atomic nuclei and electrons. A complete form of the non-relativistic Hamiltonian for the molecule, in atomic units, is given by

$$\hat{H}_{\text{molecule}} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{r_{AB}} \tag{2.1.10}$$

where the first and second terms are the kinetic energy operators for electrons and atomic nuclei respectively, the third term describes the attractive Coulomb potential between an electron and an atomic nucleus, and the fourth and fifth terms describe the repulsive Coulomb potential among electrons, and nuclei respectively. Since the mass of atomic nuclei is much larger than that of electrons (i.e. $M_A \gg 1$), it is an

appropriate approximation to neglect the second term of (2.1.10) and consider the last term as a constant since the positions of atomic nuclei are constant in the time scale of electron dynamics. Hence, we can now introduce the electronic Hamiltonian, which is given by

$$\hat{H}_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{2.1.11}$$

with an electronic time-independent Schröndinger equation

$$\hat{H}_{\text{elec}} \left|\Psi_{\text{elec}}\right\rangle = E_{\text{elec}} \left|\Psi_{\text{elec}}\right\rangle \tag{2.1.12}$$

where E_{elec} and $|\Psi_{\text{elec}}\rangle$ are the electronic energy and wavefunction respectively. The total energy for a molecule within the Born-Oppenheimer approximation is hence

$$E_{\text{total}} = E_{\text{elec}} + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{r_{AB}} \tag{2.1.13}$$

which is just the Coulomb potential energy of the nuclei added to the electronic energy. Once the electronic Schrödinger equation is solved, the nuclear Hamiltonian can be expressed with the average potential generated by electrons, given as

$$\begin{aligned}\hat{H}_{\text{nucl}} &= -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{r_{AB}} + \left\langle -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \right\rangle \\ &= -\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2 + \hat{V}_{\text{nucl}}(\mathbf{R})\end{aligned} \tag{2.1.14}$$

with the nuclear Schrödinger equation,

$$\hat{H}_{\text{nucl}} \left|\Psi_{\text{nucl}}\right\rangle = E \left|\Psi_{\text{nucl}}\right\rangle . \tag{2.1.15}$$

The total wavefunction within the BO approximation is hence,

$$\Psi_{\text{molecule}}(\mathbf{x};\mathbf{R}) = \Psi_{\text{elec}}(\mathbf{x};\mathbf{R})\Psi_{\text{nucl}}(\mathbf{R}). \tag{2.1.16}$$

Note, however, that this is not a separation of variables since the electronic wavefunction is still dependent on the positions of atomic nuclei. Since the works in this thesis use the Born-Oppenheimer approximation, the notion of electronic wavefunction will be dropped in the next sections.

2.2 The Hartree-Fock theory

2.2.1 The Hartree product *ansatz*

To solve the electronic Schrödinger equation 2.1.12, we attempt to apply the separation of variables (*i.e.* spatial-spin coordinates $\mathbf{x}_i$ for each electron) to the many-body electronic wavefunction, that is, the *ansatz* of the wavefunction is simply a product of single-particle wavefunctions, namely spin-orbitals $\chi_i(\mathbf{x}_i)$, and has the form

$$\Psi(\{\mathbf{x}\}) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2)\cdots\chi_N(\mathbf{x}_N) \tag{2.2.1}$$

which is known as the Hartree product [28]. Using this wavefunction to obtain the expectation value of the Hamiltonian in 2.1.11 will give,

$$\langle\Psi|\,\hat{H}\,|\Psi\rangle = \\ \sum_{i=1}^{N}\int \mathrm{d}\mathbf{x}_i\chi_i^*(\mathbf{x}_i)\left(-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}}\right)\chi_i(\mathbf{x}_i) + \sum_{i=1}^{N}\sum_{j>i}^{N}\int \mathrm{d}\mathbf{x}_i\mathrm{d}\mathbf{x}_j\frac{|\chi_i(\mathbf{x}_i)|^2|\chi_j(\mathbf{x}_j)|^2}{r_{ij}} \tag{2.2.2}$$

The application of variational principle and Lagrangian multipliers to this equation will lead to the Hartree equation for a single-particle wavefunction, which is given by,

$$\left[-\frac{1}{2}\nabla_i^2 + \hat{v}_{\text{eff}}\right]\chi_i(\mathbf{x}_i) = \varepsilon_i\chi_i(\mathbf{x}_i) \tag{2.2.3}$$

where

$$\hat{v}_{\text{eff}} = -\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{j\neq i}^{N}\int \mathrm{d}\mathbf{x}_j\frac{|\chi_j(\mathbf{x}_j)|^2}{r_{ij}} \tag{2.2.4}$$

is the effective potential experienced by electron i. Again, in the left hand side of this equation, the first term is the kinetic energy operator for one particle, and the second term is the sum of the external potential generated by atomic nuclei and the Coulomb potential caused by the average positions of other electrons in the system. The eigenvalue of this equation represents the contribution of electron i to the total energy. While many-body effects other than averaged Coulomb potentials stemming from other electrons are entirely neglected, the independent particle picture in this approximation can be somewhat useful in the formulations of higher level of theory. We will see later in this chapter that the similarity of this equation can be found in the Hartree-Fock and the Kohn-Sham equations (2.2.17 and 2.4.17), only differing in the expression of the effective potential $\hat{v}_{\text{eff}}$ for electrons.

2.2.2 The Slater determinant *ansatz*

It is worth noting that a wavefunction with the form of equation 2.2.1 becomes problematic if we account for the Pauli exclusion principle, which requires antisymmetry of the fermionic wavefunctions with respect to the exchange of two identical particles (equation 2.1.3). The only solution of the Hartree product that satisfy this condition is zero; therefore, the Hartree product cannot be considered as a 'quantum mechanically' correct form for many-electron systems. However, it is still possible improve the *ansatz* unpon applying antisymmetrisation to the Hartree product.

Consider a Hartree product for two electrons:

$$\Psi^{\mathrm{HP}}(\mathbf{x}_1, \mathbf{x}_2) = \chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2). \tag{2.2.5}$$

Since the electrons are indistinguishable from each other, this is physically equivalent to the Hartree product,

$$\tilde{\Psi}^{\mathrm{HP}}(\mathbf{x}_1, \mathbf{x}_2) = \chi_1(\mathbf{x}_2)\chi_2(\mathbf{x}_1) \tag{2.2.6}$$

where the single-particle spin-orbitals are now occupied by electrons with different indices. We can therefore construct a new wavefunction by taking a linear combination of these two products,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \left(\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_1(\mathbf{x}_2)\chi_2(\mathbf{x}_1)\right) \tag{2.2.7}$$

where the factor $1/\sqrt{2}$ is a normalisation constant. It is straightforward to see that this wavefunction is antisymmetric with respect to the exchange of the two particles.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = -\Psi(\mathbf{x}_2, \mathbf{x}_1) \tag{2.2.8}$$

Notice that the equation 2.2.7 can be rewritten as a determinant,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) \end{vmatrix} \tag{2.2.9}$$

and we can generalise this equation for the expression for N-electron wavefunctions by writing:

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\mathbf{x}_1) & \chi_2(\mathbf{x}_1) & \cdots & \chi_N(\mathbf{x}_1) \\ \chi_1(\mathbf{x}_2) & \chi_2(\mathbf{x}_2) & \cdots & \chi_N(\mathbf{x}_2) \\ \vdots & \vdots & & \vdots \\ \chi_1(\mathbf{x}_N) & \chi_2(\mathbf{x}_N) & \cdots & \chi_N(\mathbf{x}_N) \end{vmatrix} \tag{2.2.10}$$

where the rows of the determinant are labelled by the index of electrons and the columns are labelled by the spin-orbitals. In this determinant, the exchange of two

electrons is equivalent to the exchange of two rows, which changes the sign of the determinant and naturally incorporates the antisymmetric principle. Hence, we have completed the 'antisymmetrisation' of the Hartree product $\mathcal{A}\Psi^{\mathrm{HP}}$ and such form of wavefunctions is also known as the Slater determinant [29] or the Hartree-Fock *ansatz*.

Now inserting this *ansatz* to the variational energy expression 2.2.1 will give

$$\begin{aligned}\langle\Psi|\,\hat{H}\,|\Psi\rangle &= \sum_i \langle\chi_i|\,\hat{h}\,|\chi_i\rangle + \frac{1}{2}\sum_{ij}\langle\chi_i\chi_j|\chi_i\chi_j\rangle - \langle\chi_i\chi_j|\chi_j\chi_i\rangle \\ &\equiv \sum_i \langle\chi_i|\,\hat{h}\,|\chi_i\rangle + \frac{1}{2}\sum_{ij}\langle\chi_i\chi_j|\,|\chi_i\chi_j\rangle\end{aligned} \tag{2.2.11}$$

where the notations

$$\langle\chi_i|\,\hat{h}\,|\chi_j\rangle = \int dx\mathbf{x}_1\chi_i^*(\mathbf{x}_1)\left(-\frac{1}{2}\nabla_1^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{1A}}\right)\chi_j(\mathbf{x}_1) \tag{2.2.12}$$

and

$$\langle\chi_i\chi_j|\chi_k\chi_l\rangle = \int \mathrm{d}\mathbf{x}_1\mathrm{d}\mathbf{x}_2\chi_i^*(\mathbf{x}_1)\chi_j^*(\mathbf{x}_2)\frac{1}{r_{12}}\chi_k(\mathbf{x}_1)\chi_l(\mathbf{x}_2) \tag{2.2.13}$$

are used for brevity. Comparing 2.2.2 and 2.2.11, one notices that the expression of the HF energy comes with an additional term involving two-electron integrals, which is referred to as the exchange energy. This exchange term arises from the antisymmetry of the Slater determinant and does not have a classical interpretation. The effect of lowering the energy of the system is consistent with the view that Coulomb repulsion between electrons are being reduced as the Pauli exclusion principle forbids two electrons to be in the same coordinates. Applying the variational principle and the technique of Lagrange multipliers, we obtain

$$\left[-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{j\neq i}^{N}\int \mathrm{d}\mathbf{x}_j\frac{|\chi_j(\mathbf{x}_j)|^2}{r_{ij}} + \sum_{j\neq i}^{N}\int \mathrm{d}\mathbf{x}_j\frac{\chi_j^*(\mathbf{x}_j)\chi_i(\mathbf{x}_j)}{r_{ij}}\right]\chi_i(\mathbf{x}_i) = \varepsilon_i\chi_i(\mathbf{x}_i) \tag{2.2.14}$$

and, with the introduction of new notation, this can be reduced to

$$\left[\hat{h}(\mathbf{x}_1) + \sum_{j\neq i}\hat{J}_j(\mathbf{x}_1) - \sum_{j\neq i}\hat{K}_j(\mathbf{x}_1)\right]\chi_i(\mathbf{x}_i) = \varepsilon_i\chi_i(\mathbf{x}_i) \tag{2.2.15}$$

where $\hat{J}$ and $\hat{K}$ are referred to as coulomb and exchange potentials respectively. Notice that, since

$$\left[\hat{J}_i(\mathbf{x}_1) - \hat{K}_i(\mathbf{x}_1)\right]\chi_i(\mathbf{x}_i) = 0, \tag{2.2.16}$$

the Hartree-Fock equation in the simplest form can be written as

$$\hat{f}(\mathbf{x}_1)\chi_i(\mathbf{x}_i) = \varepsilon_i\chi_i(\mathbf{x}_i) \tag{2.2.17}$$

where

$$\begin{aligned}\hat{f}(\mathbf{x}_1) &= \hat{h}(\mathbf{x}_1) + \sum_j \hat{J}_j(\mathbf{x}_1) - \hat{K}_j(\mathbf{x}_1) \\ &\equiv \hat{h}(\mathbf{x}_1) + \hat{v}^{\mathrm{HF}}(\mathbf{x}_1)\end{aligned} \tag{2.2.18}$$

is called the Fock operator. The HF equation is reminiscent of the Hartree equation 2.2.3 and a form of independent-particle approximation.

If we assume that spin-orbitals $\chi_i(\mathbf{x}_i)$ are separable into a spatial and a spin part, *i.e.*

$$\chi_i(\mathbf{x}) = \psi_i(\mathbf{r})\sigma_i(\omega) \tag{2.2.19}$$

where the spin part $\sigma(\omega)$ can be either spin up $\alpha(\omega)$ or spin down $\beta(\omega)$, using these in the Slater determinant will lead to either a restricted HF (RHF) or unrestricted HF (UHF) wavefunction. In the RHF wavefunction, the spatial part of spin orbitals is restricted be the same for a pair of up and down spin orbitals, whereas in the UHF wavefunction, no such restriction is imposed. It is worth noting that, in both cases, the exchange integral part of the Hartree-Fock energy (equation 2.2.11) between two spin orbitals with opposite spins vanishes due to the orthogonality of the spin wavefunctions. An important consequence is that there can be only exchange interactions between electrons with parallel spin.

2.3 Coupled-cluster theory

2.3.1 Form of exact wavefunctions

Here we first derive the true form of an exact wavefunction for identical fermions. Recall that any function of a single variable can be expanded in a complete set of basis functions,

$$\Phi(\mathbf{x}_1) = \sum_{i=1} a_i\chi_i(\mathbf{x}_1) \tag{2.3.1}$$

where a_i is a coefficient. In the case of a function of two variables, the function can be expanded as

$$\begin{aligned}\Phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{i=1} a_i(\mathbf{x}_2)\chi_i(\mathbf{x}_1) \\ &= \sum_{ij} b_{ij}\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\end{aligned} \tag{2.3.2}$$

where in the first line the variable x_2 is being fixed, and in the second line the coefficients of the first expansion are expanded in the same basis. To antisymmetrise the function, we require the coefficients to satisfy the conditions $b_{ij} = -b_{ji}$ and $b_{ii} = 0$, resulting in:

$$\begin{aligned}\mathcal{A}\Phi(\mathbf{x}_1, \mathbf{x}_2) &= \sum_{i=1} \sum_{j>i} b_{ij} \left[\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)\right] \\ &= \sum_{j>i} b_{ij}\sqrt{2} \left|\chi_i\chi_j\right\rangle\end{aligned} \tag{2.3.3}$$

where $|\chi_i\chi_j\rangle$ is a Slater determinant. In other words, a fermionic wavefunction can be exactly written as a linear combination of Slater determinants constructed with a complete set of basis functions. We can extend the argument to wavefunctions of any N-electron system, and, by using the HF wavefunction as a reference, the exact form can be written as

$$|\Phi\rangle = c_0 |\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{a<b,r<s} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{a<b<c,r<s<t} c_{abc}^{rst} \left|\Psi_{abc}^{rst}\right\rangle + \cdots \tag{2.3.4}$$

where $|\Psi_0\rangle$ represents the HF wavefunction, $|\Psi_a^r\rangle$ represents a singly excited Slater determinant in which one electron in an occupied spin orbital χ_a is promoted to a virtual spin orbital χ_r, and similarly for a doubly excited Slater determinant $|\Psi_{ab}^{rs}\rangle$, etc., and $\sum_{a<b,r<s}$ means the summation of all a and b greater than a, as well as all r and s greater than r. The exact ground and excited state energies of this wavefunction can be obtained by diagonalising the Hamiltonian matrix with the matrix element $\langle\Psi_i| \hat{H} |\Psi_j\rangle$ constructed with a set of Slater determinants (Figure 2.1). Such a procedure for calculating energies is called configuration interaction (CI), since it involves configurations of spin orbitals. The lowest energy obtained in this way gives the exact ground-state energy to the non-relativistic Schrödinger equation within the Born-Oppenheimer approximation. The difference between this exact energy and the HF limit energy is called the correlation energy. In a finite basis set, the number of Slater determinants required in the wavefunction is given as $\binom{2K}{N}$ where K is the number of basis functions. Therefore, the number of the matrix elements to be calculated rapidly becomes formidable as the size of the system increases and full CI calculations are only possible for very small molecules. To mitigate the computational effort, one may truncate the exact form to wavefunction to include only up to a certain tuples of excited configurations, such as doubly exited configuration, or singly and doubly excited configurations only, (termed DCI and SDCI respectively).

$$\mathbf{H} = \begin{bmatrix} \langle\Psi_0|\hat{H}|\Psi_0\rangle & 0 & \langle\Psi_0|\hat{H}|D\rangle & 0 & 0 & \\ & \langle S|\hat{H}|S\rangle & \langle S|\hat{H}|D\rangle & \langle S|\hat{H}|T\rangle & 0 & \\ & & \langle D|\hat{H}|D\rangle & \langle D|\hat{H}|T\rangle & \langle D|\hat{H}|Q\rangle & \cdots \\ & & & \langle T|\hat{H}|T\rangle & \langle T|\hat{H}|Q\rangle & \\ & & & & \langle Q|\hat{H}|Q\rangle & \\ & & \vdots & & & \ddots \end{bmatrix} \tag{2.3.5}$$

Figure 2.1: Structure of the full CI Hamiltonian block matrix, where $|S\rangle$ represents a set of configurations with single excitation, $|D\rangle$ represents a set of configurations with double excitations, and so on. Since the matrix is hermitian, only the upper triangle part is shown. As a consequence of the Slater-Condon rules [30, 31], the matrix elements between Slater determinants which are differ in more than two spin orbitals are zero. The matrix elements between $|\Psi_0\rangle$ and $|S\rangle$ are also zero if a Hartree-Fock wavefunction is used as a reference, due to the Brillouin theorem.

However, the applications of such methods are often limited in chemistry due to the lack of size consistency.

To illustrate this, consider a wavefunction of two noninteracting and identical molecules A and B separated by a large distance. Physically, the wavefunction of two molecules should satisfy the condition

$$|\Phi_{AB}\rangle = \mathcal{A}\,|\Phi_A\rangle\,|\Phi_B\rangle \tag{2.3.6}$$

and the energy of two molecules should be just twice the energy of one molecule alone

$$E_{AB} = E_A + E_B. \tag{2.3.7}$$

If a truncated CI wavefunction (let say DCI) is used to describe to the wavefunction of the individual molecule $\left|\Phi_{A/B}\right\rangle$, the description of the wavefunction of two molecules combined $|\Phi_{AB}\rangle$ requires up to twice the tuple excitation considered in the individual molecules (quadruple excitation), resulting in an inconsistent form of the wavefunction. Hence, truncated CI wavefunctions are not strictly separable and the energy obtained within these are not size-extensive.

Alternatively, the CI wavefunctions can be also expressed in the second quantisation formalism, given as

$$|\Phi\rangle = c_0\,|\Psi_0\rangle + \sum_{ra} c_a^r \hat{\tau}_a^r\,|\Psi_0\rangle + \sum_{a<b,r<s} c_{ab}^{rs}\hat{\tau}_{ab}^{rs}\,|\Psi_0\rangle + \sum_{a<b<c,r<s<t} c_{abc}^{rst}\hat{\tau}_{abc}^{rst}\,|\Psi_0\rangle + \cdots \tag{2.3.8}$$

where the operator $\hat{\tau}_a^r = \hat{a}_r^\dagger \hat{a}_a$ is a single excitation operator, $\hat{\tau}_{ab}^{rs} = \hat{a}_r^\dagger \hat{a}_s^\dagger \hat{a}_b \hat{a}_a$ is a double excitation operator, and so on, and the fermionic creational and annihilation operators have the following relations,

$$\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{i,j}, \quad \{\hat{a}_i, \hat{a}_j\} = 0, \quad \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0 \tag{2.3.9}$$

where the brackets denote anticommutators.

2.3.2 The coupled-cluster *ansatz*

We will introduce an alternative form of the correlated wavefunction akin to the CI wavefunctions. The coupled-cluster (CC) *ansatz* is given by

$$|\Phi\rangle = e^{\hat{T}} |\Psi_0\rangle = \left(\sum_n \frac{1}{n!} \hat{T}^n \right) |\Psi_0\rangle \tag{2.3.10}$$

where

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots + \hat{T}_N \tag{2.3.11}$$

is called the cluster operator and each term

$$\hat{T}_n = \sum_{i=1} t_i \hat{\tau}_i \tag{2.3.12}$$

where a cluster amplitude t_i describes the probability of the excitation process $\hat{\tau}_i$ which the same effect of exciting n electrons in equation 2.3.8. If all excitations are included in the cluster operator $\hat{T}$, the CC wavefunction becomes identical to the full CI wavefunction. With the CC *ansatz*, the Schrödinger equation can be therefore written by using the HF wavefunction as a reference,

$$\hat{H} e^{\hat{T}} |\Psi_0\rangle = E e^{\hat{T}} |\Psi_0\rangle \tag{2.3.13}$$

In principle, the ground state energy can be obtained variationally in the same manner as it is done for CI calculations. However, this will result in procedures that are not computationally tractable due to the non-truncatable series expansion of the energy functional (equation 2.1.7). Nevertheless, an alternative approach to calculate the energy using CC wavefunctions exists. Consider the Schrödinger equation 2.3.13, and projecting the equation onto $\langle \Psi_{a\cdots}^{r\cdots}| \, e^{-\hat{T}}$ will yield the amplitude equation

$$\langle \Psi_{a\cdots}^{r\cdots} | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Psi_0 \rangle = 0 \tag{2.3.14}$$

and the cluster amplitude of each excitation process t_i is found such that this equation is satisfied. Once all amplitudes are obtained, the coupled-cluster energy can be calculated by projecting the equation 2.3.13 onto $\langle\Psi_0|\, e^{-\hat{T}}$,

$$\langle\Psi_0|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Psi_0\rangle = E. \tag{2.3.15}$$

Notice that in these projected equations, the combination of cluster operators and the Hamiltonian can be thought as a similarity-transformed Hamiltonian $\hat{\bar{H}} = e^{-\hat{T}}\hat{H}e^{\hat{T}}$, which can now be expanded as

$$\hat{\bar{H}} = \hat{H} + \Big[\hat{H},\hat{T}\Big] + \frac{1}{2!}\Big[\Big[\hat{H},\hat{T}\Big],\hat{T}\Big] + \frac{1}{3!}\Big[\Big[\Big[\hat{H},\hat{T}\Big],\hat{T}\Big],\hat{T}\Big] + \frac{1}{4!}\Big[\Big[\Big[\Big[\hat{H},\hat{T}\Big],\hat{T}\Big],\hat{T}\Big],\hat{T}\Big] \tag{2.3.16}$$

by using the Baker-Campbell-Hausdorff expression. The expansion terminates after the fourth nested commutators as the Hamiltonian contains up to four creation / annihilation operators. Therefore, despite the exponential form of the coupled-cluster *ansatz*, the projected equations will only have a finite number of terms without any approximations. It is worth noting that the similarity transformed Hamiltonian is no longer hermitian, therefore the variational principle is not applicable and the coupled-cluster energy can be lower than the exact ground-state energy. However, in contrast to the truncated CI wavefunction, the truncation of the cluster operator to arbitrary order of excitations does not effect the size extesivity of the wavefunction. To illustrate this, it can be seen that the condition in equation 2.3.6 is satisfied,

$$\begin{aligned}
\big|\Phi^{AB}_{CC}\big\rangle &= \mathcal{A}\{\big|\Phi^{A}_{CC}\big\rangle \times \big|\Phi^{B}_{CC}\big\rangle\} \\
&= \mathcal{A}\{e^{\hat{T}^A}\big|\Psi^{A}_{0}\big\rangle \times e^{\hat{T}^B}\big|\Psi^{B}_{0}\big\rangle\} \\
&= e^{\hat{T}^A}e^{\hat{T}^B}\big|\Psi^{AB}_{0}\big\rangle \\
&= e^{\hat{T}^{AB}}\big|\Psi^{AB}_{0}\big\rangle
\end{aligned} \tag{2.3.17}$$

where the excitations of electrons from molecule A to B (and vice versa) can be neglected and $[\hat{T}^A,\hat{T}^B] = 0$ since two molecules are non-interacting. Hence, truncated coupled-cluster wavefunctions are size extensive.

In practice, the cluster operator including only single and double excitation is feasible for obtaining the CC wavefunctions of small molecules (termed CCSD). CCSD energies, however, are often not accurate enough for achieving chemical accuracy (error within 1 kcal mol^{-1}) for many molecules. While the inclusion of triplet excitation in the cluster operator (termed CCSDT) makes the computation of amplitudes prohibitively expensive, the energy contribution from the triplet excitation can be estimated through perturbation theory and the corrective energy term can be added on

top of the CCSD energy (termed CCSD(T)). The CCSD(T) method is often regarded as a good compromise between accuracy and computational cost, and considered as the gold standard in the chemistry community.

2.3.3 Local correlation methods

Unfortunately, due to the polynomial scaling and memory requirements, straightforward applications of the CCSD(T) method are limited to very small molecules even with the advances of computer hardware. However, it is worth recognising that the scale of electron interactions should decrease as their separation increases (as we saw in the demonstration of non-interacting molecules earlier). The motivation for the local correlation methods is to further approximate the existing methods, by exploiting locality of electrons in real space, to make the computation feasible for laBook Boxes - Corrugated to fit A4 and adjustable to book thickness, self sealing, protects corners for safe delivery.rger systems. Recall that the exact wavefunction can be written as equation 2.3.4 using the HF wavefunctions as a reference. Inserting the wavefunction into the Schrödinger equation and projecting onto $\langle\Psi_0|$, we arrive at an alternative expression for the correlation energy, given as

$$E_{\text{corr}} = \sum_{a<b,r<s} c_{ab}^{rs} \langle ab| \, |rs\rangle = \sum_{a<b} \varepsilon_{ab} \tag{2.3.18}$$

where c_{ab}^{rs} is the coefficient of a doubly excited configuration in the intermediate normalised full CI wavefunction, and ε_{ab} is the pair-correlation energy. One must expect this pair correlation energy to decay as the orbitals a and b are further apart in distance. Since the expression for the correlation energy is invariant to the unitary transformation of HF orbitals (or canonical molecular orbitals), localisation of MOs can be utilised to select a list of orbitals that are strongly contributing to the leading terms of the pair correlation energies. For example, in the formulation of local natural orbital CCSD(T) (LNO-CCSD(T)) [32–35], the traditional CCSD correlation energy

$$E_{\text{CCSD}} = \sum_{ar} t_a^r f_{ar} + \sum_{a<b,r<s} \left(t_{ab}^{rs} + t_a^r t_b^s\right) \langle ab| \, |rs\rangle \tag{2.3.19}$$

is expressed in terms of each LMO's contribution to the correlation energy,

$$\begin{aligned} E_{\text{CCSD}} &= \sum_{a'} \delta E_{a'} \\ &= \sum_{a'} \left[\sum_{r} t_a^r f_{ra} + \sum_{a<b,r<s} \left(t_{ab}^{rs} + t_a^r t_b^s\right) \langle ab| \, |rs\rangle \right] \end{aligned} \tag{2.3.20}$$

where $\delta E_{a'}$ is the energy contribution of a'th localised occupied MO. In the first approximation, we can restrict the computation of pair correlation energies $\varepsilon_{a'b'}$ with the CCSD equation to only pairs that might be significant. This can be determined based on estimations from multipole approximations of MP2 energies, which can be easily calculated. The distant pair correlation energies that were not considered in the CCSD equation are evaluated at the MP2 level, and added to the total energy later. In the second approximation, the domain for virtual orbitals r and s is restricted to localised virtual orbitals that surround the pair of LMOs a' and b'. This is achieved by constructing projected atomic orbitals (PAOs) centred at atoms surrounding each LMO, and orthogonalising PAOs on the union of atoms belonging to the pair of LMOs. The domain of PAO-centred atoms is chosen so that it recovers 99.9 % of the molecular orbitals by default. With these two leading approximations, the number of operations in the summation of equation 2.3.20 is significantly suppressed, ensuring the linear scaling of correlation methods with respect to the size of (non-metallic) systems.

In addition to LNO-CCSD(T), there are other local approximations to CCSD(T) available such as DLPNO-CCSD(T_0/T_1) [36, 37] in Orca [38], and PNO-CCSD(T) [39, 40] in Turbomole [41, 42] and Molpro [43], mostly differing in the method of truncating the expression of equation 2.3.19. All of these implementations can effectively calculate accurate energies while suppressing the computational scaling of the canonical CCSD(T). However, a wide range of benchmarking studies have shown a consistent superiority of LNO-CCSD(T) compared to other implementaions in thermochemistry and energy barriers of organometallic reactions [44, 45], interconversion energies of extended porphyrins [46], non-covalent interaction energies in molecular complexes [47, 48] and conformational energies of alkane chains [49]. Hence, in this thesis, LNO-CCSD(T) will be used to calculate the accurate energies for reactions of hydrido-ruthenium carbonyl clusters in Chapter 3.

2.4 Density functional theory

There have been several attempts to calculate the electronic ground-state energy solely based on the density of electrons. An early example can be found in the Thomas-Fermi-Dirac (TFD) model, which gives an approximation to an electronic

ground-state energy, given by

$$\begin{aligned} E_{\mathrm{TFD}}[\rho] &= T_{\mathrm{TF}}[\rho] + E_{\mathrm{ext}}[\rho] + E_{\mathrm{x}}[\rho] + E_{\mathrm{Hartree}}[\rho] \\ &= C_1 \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})^{\frac{5}{3}} + \int \mathrm{d}\mathbf{r} V_{\mathrm{ext}}(\mathbf{r})\rho(\mathbf{r}) + C_2 \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})^{\frac{4}{3}} + \frac{1}{2}\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \end{aligned} \tag{2.4.1}$$

where the first term is an approximation to the kinetic energy, the second term is the energy due to interactions with the external potential, the third term is the Slater exchange and the fourth term is the classical Hartree energy, and C_1 and C_2 are empirical parameters. The ground-state energy and density are then obtained by minimising the functional with varying density. However, the accuracy of the TFD model is often very limited, due to several accounts, including assuming uniformity of the electron gas (resulting in inaccuracies in the kinetic and exchange energy terms) and self-interactions in the Hartree term.

2.4.1 The Hohenberg-Kohn theorems

A rigorous and exact formulation of the DFT begins with the Hohenberg-Kohn (HK) theorems [50] which make two important statements as given in the following:

I: The external potential $V_{\mathrm{ext}}(\mathbf{r})$, given that it is not constant, is uniquely determined by the ground state particle density $\rho_0(\mathbf{r})$ for any many-body systems.

proof Let us consider two different Hamiltonians $\hat{H}$ and $\hat{H}'$ which are only different in the external potentials $V_{\mathrm{ext}}(\mathbf{r})$ and $V'_{\mathrm{ext}}(\mathbf{r})$ with two different corresponding ground-state wavefunctions Ψ and Ψ'. We assume the two ground-state wavefunctions to have the same density $n_0(\mathbf{r})$. By following the variational principle, it can be stated that

$$\begin{aligned} E_0 < \langle\Psi'|\hat{H}|\Psi'\rangle &= \langle\Psi'|\hat{H}'|\Psi'\rangle - \langle\Psi'|\hat{H}-\hat{H}'|\Psi'\rangle \\ &= E'_0 + \int \rho(\mathbf{r})\left[V_{\mathrm{ext}}(\mathbf{r}) - V'_{\mathrm{ext}}(\mathbf{r})\right]\mathrm{d}\mathbf{r} \end{aligned} \tag{2.4.2}$$

given that the ground state of the Hamiltonian is non-degenerate. Similarly, for the ground-state energy of $\hat{H}'$ is given as

$$\begin{aligned} E'_0 < \langle\Psi|\hat{H}'|\Psi\rangle &= \langle\Psi|\hat{H}|\Psi\rangle - \langle\Psi|\hat{H}'-\hat{H}|\Psi\rangle \\ &= E_0 + \int \rho(\mathbf{r})\left[V'_{\mathrm{ext}}(\mathbf{r}) - V_{\mathrm{ext}}(\mathbf{r})\right]\mathrm{d}\mathbf{r} \end{aligned} \tag{2.4.3}$$

Adding equation 2.4.2 and 2.4.3, we will arrive at an inequality

$$E_0 + E'_0 < E'_0 + E_0$$

which is a contradiction.

II: The energy functional of any trial density $\tilde{\rho}$ for a given $V_{\text{ext}}(\mathbf{r})$ is always greater than or equal to the exact ground-state energy,

$$E_0 \leq E_V[\tilde{\rho}] \tag{2.4.4}$$

where $E_V[\tilde{\rho}]$ is the energy functional.

proof Since all properties of a wavefunction can be determined for a given $\rho(\mathrm{r})$, the total energy functional can be expressed as

$$E_{\text{HK}}[\rho] = F_{\text{HK}}[\rho] + \int \mathrm{d}\mathbf{r} V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \tag{2.4.5}$$

where the functional

$$F_{\text{HK}}[\rho] = T[\rho] + V_{ee}[\rho] \tag{2.4.6}$$

must be universal for all electron systems as it is a functional of the density only. Now notice that the first HK theorem assures that a density ρ uniquely determines its own $V_{\text{ext}}(\mathrm{r})$, the Hamiltonian $\hat{H}$, and the wavefunction Ψ.

$$E_0 = E_{\text{HK}}[\rho] = \langle\Psi|\hat{H}|\Psi\rangle \tag{2.4.7}$$

If we consider a different density $\tilde{\rho}$, which must belong to a different wavefunction $\tilde{\Psi}$, it immediately follows, from the variational principle, that

$$E_0 = E_V[\rho] = \langle\Psi|\hat{H}|\Psi\rangle < \langle\tilde{\Psi}|\hat{H}|\tilde{\Psi}\rangle = E_V[\tilde{\rho}]. \tag{2.4.8}$$

Hence the energy functional evaluated with the density of the exact ground-state wavefunction is always lower than those with any other densities.

2.4.2 The Kohn-Sham *ansatz*

A more practical implementation of DFT is led by the Kohn-Sham approach [51], which simplifies the many-body interacting system into to an auxiliary non-interacting system that is easier to solve. The effects of many-body interactions are approximated in the so-called exchange-correlation energy functional, and the limitation of the approach is only this approximation.

Two main assumptions are being made in the Kohn-Sham auxiliary system:

1. Noninteracting v-representability: The density of an exact ground-state wavefunction can be represented by the density of a ground-state auxiliary system of noninteracting particles.

2. The Hamiltonian of such an auxiliary system is chosen to have the form

$$\hat{H}_s = \sum_{i=1}^{N} \left(-\frac{1}{2}\nabla_i^2 \right) + \sum_{i=1}^{N} v_s(\mathbf{r_i}) \tag{2.4.9}$$

in which there are no electron-electron repulsion terms

The density of the auxiliary system is given by

$$\rho(\mathbf{r}) = \sum_{i=1}^{N} |\chi_i(\mathbf{r})|^2 \tag{2.4.10}$$

and the kinetic energy of independent particles is given by

$$T_s = -\frac{1}{2}\sum_{i=1}^{N} \langle \chi_i | \nabla^2 | \chi_i \rangle = \frac{1}{2}\sum_{i=1}^{N} \int \mathrm{d}r^3 |\nabla \chi_i(\mathbf{r})|^2 \tag{2.4.11}$$

The Hartree energy can be given as

$$E_{\mathrm{Hartree}}[\rho] = \frac{1}{2}\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}' \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \tag{2.4.12}$$

Now the Hohenberg-Kohn functional can be written as

$$E_{\mathrm{KS}}[n] = T_s[n] + \int \mathrm{d}\mathbf{r} V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) + E_{\mathrm{Hartree}}[n] + E_{\mathrm{xc}}[n] \tag{2.4.13}$$

and the exchange-correlation functional is defined as

$$\begin{aligned} E_{\mathrm{xc}}[\rho] &= F_{\mathrm{HK}}[\rho] - (T_s[\rho] + E_{\mathrm{Hartree}}[\rho]) \\ &= \langle \hat{T} \rangle - T_s[\rho] + \langle \hat{V}_{ee} \rangle - E_{\mathrm{Hartree}}[\rho] \end{aligned} \tag{2.4.14}$$

Therefore, the exchange-correlation energy can be interpreted as the sum of the difference in the kinetic energies of the exact wavefunction and the noninteracting system, and the nonclassical terms in the electron-electron interaction energy. To solve for the ground state of the Kohn-Sham system, the variational equation is given as

$$\frac{\delta E_{\mathrm{KS}}}{\delta \chi_i^*(\mathbf{r})} = \frac{\delta T_s[n]}{\delta \chi_i^*(\mathbf{r})} + \left[\frac{\delta E_{\mathrm{ext}}(\mathbf{r})}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{Hartree}}[n]}{\delta n(\mathbf{r})} + \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})} \right] \frac{\delta n(\mathbf{r})}{\delta \chi_i^*(\mathbf{r})} = 0 \tag{2.4.15}$$

$$\frac{\delta \rho(\mathbf{r})}{\delta \chi_i^*(\mathbf{r})} = \chi_i(\mathbf{r}); \quad \frac{\delta T_s}{\delta \chi_i^*(\mathbf{r})} = -\frac{1}{2}\nabla^2 \chi_i(\mathbf{r}) \tag{2.4.16}$$

lead to the Kohn-Sham equation

$$\left[-\frac{1}{2}\nabla^2 - V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{Hartree}}(\mathbf{r}) + V_{\mathrm{xc}}(\mathbf{r}) \right] \chi_i(\mathbf{r}) = \varepsilon_i \chi_i(\mathbf{r}) \tag{2.4.17}$$

where ε_i is the eigenvalue, $V_{\mathrm{Hartree}}(\mathbf{r}) = \frac{\delta E_{\mathrm{Hartree}}[n]}{\delta n(\mathbf{r})}$ is the Hartree potential, $V_{\mathrm{xc}}(\mathbf{r}) = \frac{\delta E_{\mathrm{xc}}[n]}{\delta n(\mathbf{r})}$ is the exchange-correlation potential.

2.4.3 The exchange-correlation functional

The general and basic expression for the exchange-correlation functional energy is given as

$$E_{\mathrm{xc}} = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\epsilon_{\mathrm{xc}}([\rho], \mathbf{r}) \tag{2.4.18}$$

where $\epsilon_{\mathrm{xc}}([\rho], \mathbf{r})$ is a functional of ρ that gives an energy per electron at $\mathbf{r}$. The functional $\epsilon_{\mathrm{xc}}([\rho], \mathbf{r})$ can be understood as the electrostatic potential energies of the interactions between an electron and an exchange-correlation hole,

$$\epsilon_{\mathrm{xc}}^{\mathrm{int}}([\rho], \mathbf{r}) = \frac{1}{2}\int \mathrm{d}\mathbf{r}' \frac{\rho_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.4.19}$$

where $\rho_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')$ is the hole density at $\mathbf{r}'$ for a given position of the electron $\mathbf{r}$.

A more rigorous expression for the exchange-correlation energy in the Kohn-Sham approach can be formulated by using the adiabatic connection, as proposed by Harris [52]. Notice that the HK functional (2.4.13) can be rewritten as

$$\begin{aligned} E_{\mathrm{xc}}[\rho] &= \langle \hat{T} \rangle - T_s[\rho] + \langle \hat{V}_{ee} \rangle - E_{\mathrm{Hartree}}[\rho] \\ &= F_1[\rho] - F_0[\rho] - E_{\mathrm{Hartree}}[\rho] \\ &= \int_0^1 \mathrm{d}\lambda \frac{\partial F_\lambda[\rho]}{\partial \lambda} - E_{\mathrm{Hartree}}[\rho] \end{aligned} \tag{2.4.20}$$

where $F_1[\rho] = \langle \hat{T} \rangle + \langle \hat{V}_{ee} \rangle$ is a functional for the fully-interacting system and $F_0[\rho] = T_s[\rho]$ is the functional for the noninteracting case. These two functionals can be adiabatically connected by using the coupling constant λ, which varies the value of the electronic charge from 0 to 1. Hence an alternative expression for the exchange-correlation energy can be given as

$$E_{\mathrm{xc}}[\rho] = \frac{1}{2}\int \mathrm{d}\mathbf{r}\rho(\mathbf{r}) \int \mathrm{d}\mathbf{r}' \frac{\bar{\rho}_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.4.21}$$

where

$$\bar{\rho}_{\mathrm{xc}}(\mathbf{r}, \mathbf{r}') = \int_0^1 \mathrm{d}\lambda \rho_{\mathrm{xc}}^{\lambda}(\mathbf{r}, \mathbf{r}'). \tag{2.4.22}$$

is the average of the exchange-correlation hole density with respect to the value of λ. Therefore, the exchange-correlation energy can be understood as the interpolation from the exchange energy ($\lambda = 0$) to the fully-correlated energy at a fixed electronic density.

While the Kohn-Sham formulation of DFT itself is an exact theory, the exact expression for a universal exchange-correlation functional is not yet known. Nevertheless, approximate exchange-correlation functionals based on physical models and/or parameterisations proved useful in practice and can be categorised into several levels.

1. Local (spin) density approximations
   In local (spin) density approximations (L(S)DA), the exchange-correlation energy is calculated by using the exchange-correlation energy density of a homogeneous electron gas, with an explicit dependence on the density at a given point only. The functional has the general form,

$$E_{\mathrm{xc}}^{\mathrm{LSDA}}(\rho^{\uparrow},\rho^{\downarrow}) = \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\epsilon_{\mathrm{xc}}^{\mathrm{hom}}(\rho^{\uparrow}(\mathbf{r}),\rho^{\downarrow}(\mathbf{r})) \tag{2.4.23}$$

   where $\rho^{\uparrow/\downarrow}$ is the electron density with spin up/down. Admittedly real atoms and molecules are far from a homogeneous electron gas and LDA fails to provide accurate exchange-correlation hole densities. However, LDA can still provide a reasonable estimate to the exchange-correlation energy as this only depends on the spherical average of hole densities (equation 2.4.21).

2. Generalised gradient approximations
   A systematic improvement to LDA can be achieved through building a functional that explicitly depends on both density and its gradients, thereby accounting for the fact the electron density in a molecule is not uniformly distributed. In the generalised gradient approximation (GGA), the general expression for the exchange-correlation energy is given as,

$$\begin{aligned} E_{\mathrm{xc}}^{\mathrm{GGA}}(\rho^{\uparrow},\rho^{\downarrow},\left|\nabla\rho^{\uparrow}\right|,\left|\nabla\rho^{\downarrow}\right|) &= \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\epsilon_{\mathrm{xc}}(\rho^{\uparrow},\rho^{\downarrow},\left|\nabla\rho^{\uparrow}\right|,\left|\nabla\rho^{\downarrow}\right|) \\ &= \int \mathrm{d}\mathbf{r}\rho(\mathbf{r})\epsilon_{\mathrm{x}}^{\mathrm{hom}}(\rho)F_{\mathrm{xc}}(\rho^{\uparrow},\rho^{\downarrow},\left|\nabla\rho^{\uparrow}\right|,\left|\nabla\rho^{\downarrow}\right|) \end{aligned} \tag{2.4.24}$$

   where $\epsilon_{\mathrm{x}}^{\mathrm{hom}}(\rho)$ is the exchange energy of homogeneous electron gas and F_{xc} is a parameterised dimensionless function. A further systematic improvement on GGA can be achieved upon taking into account the Laplacian of the electron density $\nabla^2\rho(\mathbf{r})$, or the kinetic energy density $\tau(\mathbf{r})$ and such functionals are referred as the meta-GGA functionals.

3. Hybrid functionals:
   Notice that in the adiabatic connection formula (2.4.20), the exchange energy of the noninteracting case $\lambda = 0$ can be exactly obtained with the Kohn-Sham orbitals by using the expression in the Hartree-Fock approximation (assuming the noniteracting v-representability). If we set the energy at the case $\lambda = 1$ to the LDA or GGA exchange-correlation energy and assume that the energy varies

linearly with λ, we will arrive at the original "half-and-half" hybrid functional proposed by Becke [53].

$$E_{\mathrm{xc}}^{\mathrm{H\&H}} = \frac{1}{2}\left(E^{\mathrm{EXX}} + E_{\mathrm{xc}}^{\mathrm{LSDA}}\right) \tag{2.4.25}$$

A more practical hybrid functional based on the observation of non-linear relation between the energy and λ was proposed by Perdew *et al.* [54] and given as

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = \frac{1}{4}E^{\mathrm{EXX}} + \frac{3}{4}E_{\mathrm{x}}^{\mathrm{PBE}} + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{2.4.26}$$

Alternatively, the ratio of the exact exchange energy and the exchange density functional can be empirically parameterised to fit data sets. For example, the B3LYP functional [55] is given as

$$E_{\mathrm{xc}}^{\mathrm{B3LYP}} = E_{\mathrm{xc}}^{\mathrm{LSDA}} + c_0\left(E^{\mathrm{EXX}} - E_{\mathrm{x}}^{\mathrm{LSDA}}\right) + c_1\left(E_{\mathrm{x}}^{\mathrm{B88}} - E_{\mathrm{x}}^{\mathrm{LSDA}}\right) + c_2\left(E_{\mathrm{c}}^{\mathrm{LYP}} - E_{\mathrm{c}}^{\mathrm{LSDA}}\right) \tag{2.4.27}$$

where $E_{\mathrm{x}}^{\mathrm{B88}}$ and $E_{\mathrm{c}}^{\mathrm{LYP}}$ are the Becke B88 exchange and LYP correlation functionals respectively, and c_i are empirical parameters.

4. Double hybrid functionals:
   In a similar manner to hybrid functionals, the correlation energy can be mixed with the post-HF correlation energy (mostly the second-order pertubative correction) in double hybrid functionals. The double hybrid functional PBE0-DH [56], which is based on PBE0 hybrid functional, is given as

$$E_{\mathrm{xc}}^{\mathrm{PBE0-DH}} = \frac{1}{2}\left(E^{\mathrm{EXX}} + E_{\mathrm{x}}^{\mathrm{PBE}}\right) + \frac{1}{2}\left(\frac{7}{4}E_{\mathrm{c}}^{\mathrm{PBE}} + \frac{1}{4}E^{\mathrm{MP2}}\right) \tag{2.4.28}$$

   Similar to hybrid functionals, the mixing parameters for the correlation energy can be determined based on the argument using the adiabatic connection formula or empirical fitting to data sets. While the double hybrid functionals are capable of describing long-range correlation effects, their applications are limited due to their higher computational effort scaling with respect to the system size. Moreover, perturbation energies may diverge for metallic systems or systems with small band gaps.

5. van der Waals functionals:
   Since the calculation of the global correlation energy from the aforementioned double hybrid functional can be formidably extensive, alternative methods to account for long-range correlation energy are desirable. Dion and coworkers [57]

have proposed a form of the exchange-correlation density functional which is given by

$$E_{\mathrm{xc}}^{\mathrm{vdW}}[\rho] = E_x^{\mathrm{GGA}} + E_c^{\mathrm{LDA}} + E_c^{\mathrm{nl}} \tag{2.4.29}$$

where E_c^{nl} is the non-local correlation energy and has the general form

$$E_c^{\mathrm{nl}}[\rho] = \frac{1}{2}\int \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r}'\rho(\mathbf{r})\phi(\mathbf{r},\mathbf{r}')\rho(\mathbf{r}') \tag{2.4.30}$$

where $\phi(\mathbf{r},\mathbf{r}')$ is the vdW kernel. Popular approximations for the long-range energy functional and its kernel are based on the random-phase approximation (RPA), which has been shown to provide the correct description of dispersion interactions, which decay with $1/R^6$.

2.5 Basis sets

2.5.1 The Roothaan-Hall equation

The special case of the Hartree-Fock equation when spin-orbitals are restricted can be written as,

$$\hat{f}(\mathbf{r}_1)\psi_i(\mathbf{r}_1) = \varepsilon_i\psi_i(\mathbf{r}_1) \tag{2.5.1}$$

where ψ_i is a spatial orbital in real space. In principle, the solution to this equation can be numerically found by solving the integro-differential equation directly. However, the applications of such procedures are limited to very simple systems such as individual atoms. For molecules, an alternative approach is need for feasible calculations. Notice that, if we expand the orbital in a finite set of basis functions ϕ_μ

$$\psi_i = \sum_{i=1}^{K} C_{\mu i}\phi_\mu \tag{2.5.2}$$

which are not necessarily orthogonal, the Hartree-Fock equation 2.5.1 becomes

$$\hat{f}(\mathbf{r}_1)\sum_\nu C_{\nu i}\phi_\nu = \varepsilon_i\sum_\nu C_{\nu i}\phi_\nu. \tag{2.5.3}$$

Projecting the equation onto $\langle\phi_\mu|$, gives

$$\begin{aligned}\sum_\nu C_{\nu i}\int \mathrm{d}\mathbf{r}_1\phi_\mu^*\hat{f}(\mathbf{r}_1)\phi_\nu &= \varepsilon_i\sum_\nu C_{\nu i}\int \mathrm{d}\mathbf{r}_1\phi_\mu^*\phi_\nu \\ \sum_\nu C_{\nu i}f_{\mu\nu} &= \varepsilon_i\sum_\nu C_{\nu i}S_{\mu\nu}\end{aligned} \tag{2.5.4}$$

where in the second line, $f_{\mu\nu}$ is a matrix element of the Fock operator, and $S_{\mu\nu}$ is an inner product (or overlap) between two basis functions ϕ_μ and ϕ_ν. Therefore, in the matrix notation, the Hartree-Fock equation can be rewritten as a generalised eigenvalue equation with a finite basis set,

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{2.5.5}$$

where $\mathbf{C}$ is the matrix of coefficients, and $\boldsymbol{\varepsilon} = \mathrm{diag}(\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_K)$ is the diagonal matrix of eigenvalues. This expression is also known as the Roothaan-Hall equation. It is important to point out that the elements of the Fock matrix also depend on the coefficients of expansions through the Coulomb and exchange components; hence, the Roothaan-Hall equations are non-linear and have to be solved iteratively.

There is many sets of mathematical functions can be chosen for the basis set. In practical calculations, there are two primary categories of basis sets: atomic orbitals and plane waves. These will be discussed in the next sections.

2.5.2 Atomic orbital basis sets

The natural choice for basis sets for the purpose of calculating electronic properties of molecules would be the atomic orbitals centred on each atom. Moreover, using the linear combinations of atomic orbitals (LCAO) to describe the molecular orbitals gives a more chemically intuitive insight to chemical bonding.

The Slater-type orbitals (STOs) are proposed as the basis sets for early molecular calculations and, in spherical coordinates, they are given as

$$\phi^{STO}_{nlm}(r, \theta, \phi) = N r^{n-1} e^{-\zeta r} Y_{lm}(\theta, \phi) \tag{2.5.6}$$

where N is a renomalisation constant, n, l and m are quantum numbers, and $Y_{lm}(\theta, \phi)$ is the real spherical harmonic function. The parameter ζ is a positive number that defines the shape or diffuseness of the orbitals. STOs resemble some electronic wavefunctions of hydrogen-like atoms (1s, 2p, etc.) but lack radial nodes – to describe nodes, a linear combination of STOs must be used. While STOs are capable of describing several important characteristics of atomic/molecular orbitals efficiently, the evaluations of 1 or 2 electron integrals with STOs are somewhat cumbersome and too numerically demanding for many practical calculations.

Alternatively, Gaussian-type orbitals (GTOs) have been proposed as a replacement for STOs and commonly have the form,

$$\phi^{GTO}_{abc}(x, y, z) = N e^{-\alpha r^2} x^a y^b z^c \tag{2.5.7}$$

where α is the Gaussian exponent, and a, b and c are non-negative integer numbers. Notice that there are two major modifications to the expression in equation 2.5.6: the radial part of the STO is replaced with a Gaussian function and the spherical harmonic part is replaced with polynomial functions in the Cartesian coordinates. To illustrate how GTOs are preferable over STOs in numerical tasks, consider a four-centre two-electron integral,

$$\langle \phi_A \phi_B | \phi_C \phi_D \rangle = \int \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \phi_A^*(\mathbf{r}_1) \phi_B^*(\mathbf{r}_2) \frac{1}{r_{12}} \phi_C(\mathbf{r}_1) \phi_D(\mathbf{r}_2) \tag{2.5.8}$$

where ϕ_A is a basis function centred on the atom A, ϕ_B is a basis function centred on the atom B, and so on. If the basis functions in the integral are simple $1s$ Gaussian basis functions, by using the Gaussian product rule which simplifies the expressions of the product into one Gaussian function centred at a third point P,

$$e^{-\alpha_A r_A^2} e^{-\alpha_B r_B^2} = K_{AB} e^{-\alpha_P r_P^2} \tag{2.5.9}$$

the integral 2.5.8 can be now simplified into a two-centred integral,

$$\langle \phi_A \phi_B | \phi_C \phi_D \rangle = K_{AC} K_{BD} \int \mathrm{d}\mathbf{r}_1 \mathrm{d}\mathbf{r}_2 \phi_P(\mathbf{r}_1) \frac{1}{r_{12}} \phi_Q(\mathbf{r}_2) \tag{2.5.10}$$

which can be evaluated analytically. Hence, the numerical tasks on two-electron integrals, which are extensively used in HF and correlated methods, can be greatly reduced with the usage of GTOs.

However, due to the intrinsic shapes of Gaussian functions, GTOs are not as accurate as STOs in describing orbitals. In particular, GTOs are not capable of satisfying Kato's cusp condition [58], which states that the derivative of the wavefunction around an atomic nucleus must be given by,

$$\left[\frac{\partial \Psi}{\partial r_i} \right]_{r_i = 0} = -Z \Psi(r_i = 0) \tag{2.5.11}$$

where Z is the atomic number of the nucleus. While the derivatives of STOs around the nuclei are negative numbers, those of GTOs are equal to zero. Moreover, it can also be seen that GTOs decay much quicker than STOs as they move away from nuclei. To mitigate these limitations of GTOs, one may use multiple Gaussian functions in order to capture the characteristics of STOs. For example, in the STO-3G basis set, a basis function is constructed as a contraction of three GTOs with different Gaussian exponents, given as

$$\phi_{abc}^{\mathrm{STO-3G}}(x, y, z) = N \sum_{i=1}^{3} c_i e^{-\alpha_i r^2} x^a y^b z^c \tag{2.5.12}$$

This way, the shape of the orbitals around nuclei can be improved. More practical basis sets are constructed upon introductions of several other techniques, such as split valence, and additions of polarisation and diffuse functions. Ideally, the energy must be converged with respect to the size of the basis set used; however, this is practically not feasible, especially for correlation energies. There are several families of basis sets used in modern electronic structure calculations, including def2 basis set family [59] which is optimised for properties, and correlation-consistent [60] basis set which is developed specifically for post-HF calculations.

2.5.3 Plane wave basis sets

While it is possible to use atomic orbital basis sets in calculations for periodic systems, plane wave basis sets are often preferred due to their applicability in describing the periodicity and translational symmetry of crystals. Under the Born–von Karman boundary conditions, a wavefunction in a crystal can be normalised and have a periodic boundary condition at a large volume Ω, hence any state can be expanded in the complex Fourier series,

$$\psi_\nu(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{q}} c_{\mathbf{q}\nu} e^{i\mathbf{q}\cdot\mathbf{r}} = \sum_{\mathbf{q}} c_{\mathbf{q}\nu} \left|\phi_{\mathbf{q}}\right\rangle \tag{2.5.13}$$

where $|\phi_{\mathbf{q}}\rangle$ is a plane wave associated with $\mathbf{q}$, which satisfies the orthonormality relation

$$\langle\phi_{\mathbf{q}'}|\phi_{\mathbf{q}}\rangle = \frac{1}{\Omega}\int_\Omega \mathrm{d}\mathbf{r} e^{-i\mathbf{q}'\cdot\mathbf{r}} e^{i\mathbf{q}\cdot\mathbf{r}} = \delta_{\mathbf{q},\mathbf{q}'} \tag{2.5.14}$$

Therefore, the Schrödinger equation in Fourier space can be obtained as,

$$\sum_{\mathbf{q}} c_{\mathbf{q}\nu} \langle\phi_{\mathbf{q}'}|\hat{H}|\phi_{\mathbf{q}}\rangle = \varepsilon_\nu \sum_{\mathbf{q}} c_{\mathbf{q}\nu} \langle\phi_{\mathbf{q}'}|\phi_{\mathbf{q}}\rangle = \varepsilon_\nu c_{\mathbf{q}'\nu} \tag{2.5.15}$$

The matrix element of the kinetic energy can be easily calculated as,

$$\langle\phi_{\mathbf{q}'}|-\frac{1}{2}\nabla^2|\phi_{\mathbf{q}}\rangle = \frac{1}{2}q^2\delta_{\mathbf{q},\mathbf{q}'} \tag{2.5.16}$$

whereas the potential $V(r)$ (assuming that the effective potential for electrons is periodic with crystal) can be expressed in the Fourier series,

$$V(\mathbf{r}) = \sum_m V(\mathbf{G}_m) e^{i\mathbf{G}_m\cdot\mathbf{r}} \tag{2.5.17}$$

where $\mathbf{G}_m$ are reciprocal lattice vectors, and

$$V(\mathbf{G}) = \frac{1}{\Omega_{\text{cell}}}\int_{\Omega_{\text{cell}}} \mathrm{d}\mathbf{r} V(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} \tag{2.5.18}$$

where Ω_{cell} is the primitive cell volume. Hence, the matrix element of the effective potential can be given as,

$$\langle \phi_{\mathbf{q}'} | \hat{V} | \phi_{\mathbf{q}} \rangle = \sum_{m} V(\mathbf{G}_m) \delta_{\mathbf{q}'-\mathbf{q},\mathbf{G}_m} \tag{2.5.19}$$

which is only nonzero if the difference in $\mathbf{q}$ and $\mathbf{q}'$ is equal to a given $\mathbf{G}_m$. If we recognise that $\mathbf{q}$ and $\mathbf{G}_m$ can be related as $\mathbf{q} = \mathbf{k} + \mathbf{G}_m$, the Schrödinger equation can be finally written for a given $\mathbf{k}$,

$$\sum_{m'} H_{m,m'}(\mathbf{k}) c_{\nu,m'}(\mathbf{k}) = \varepsilon_{\nu}(\mathbf{k}) c_{\nu,m}(\mathbf{k}) \tag{2.5.20}$$

where the matrix element of Hamiltonian is given by

$$\begin{aligned} H_{m,m'}(\mathbf{k}) &= \left\langle \phi_{\mathbf{k}+\mathbf{G}_m} \middle| \hat{H} \middle| \phi_{\mathbf{k}+\mathbf{G}_{m'}} \right\rangle \\ &= \frac{1}{2} |\mathbf{k} + \mathbf{G}_m|^2 \delta_{m,m'} + V(\mathbf{G}_m - \mathbf{G}_{m'}) \end{aligned} \tag{2.5.21}$$

The eigenfunction of this Hamiltonian for a given $\mathbf{k}$ has the form,

$$\psi_{\nu,\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{\Omega}} \sum_{m} c_{\nu,m} e^{i(\mathbf{k}+\mathbf{G}_m)\cdot\mathbf{r}} = \frac{1}{\sqrt{N_{\text{cell}}}} u_{\nu,\mathbf{k}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{2.5.22}$$

where N_{cell} is the number of cells in the volume Ω. This is also known as a Bloch state.

In practical calculations, the truncation of the plane wave basis set is accomplished by setting a 'cutoff' energy to the kinetic energy part of the plane wave (equation 2.5.21),

$$\frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 < E_{\text{cutoff}} \tag{2.5.23}$$

in order to mitigate the effects of the primitive cell size on the quality of basis sets (as opposed to naively setting a number of plane waves). For properties that require the integration over $\mathbf{k}$-space such as total energies, charge densities, a finite $\mathbf{k}$ sampling in the first Brillouin zone (BZ) is performed with (usually but not always) a uniform sampling mesh.

2.5.4 Projector augmented wave formalism

It is worth noting that the wavefunctions close to the nuclei oscillate much quicker than wavefunctions in the valence region, as orbitals are required to be orthogonal to each other. To describe these rapid oscillations, the number of plane waves required in the Fourier expansions becomes impractically high to run any calculations. However, these oscillations of orbitals in the near-core regions play little role in chemical

bonding and certainly less than than those in the 'valence' regions. In the pseudopotential approximations, the descriptions of core electrons that do not play significant roles in chemical bonds are replaced with the effective potentials generated by the shielded nuclear charges and the repulsive forces caused by the Pauli exclusion principle. This effectively reduces the size of many-body problems in the calculations, as well as smoothens the near-core oscillations of valence orbitals as the requirement for orthogonality with core-electrons is eliminated.

The projector augmented wave (PAW) formalism is a specific implementation of norm-conserving pseudopotential methods that allows a recovery of physical properties of all-electron wavefunctions from smooth wavefunctions. In the PAW formalism, it is assumed that there exists a linear transformation of the smooth wavefunction to all-electron wavefunctions, which is a sum of contributions centred at atoms enclosed by 'augmented regions', and unity elsewhere. Within the spherical augmentation region, the smooth wavefunction can be expanded in spherical harmonics,

$$|\tilde{\psi}\rangle = \sum_m c_m |\tilde{\varphi}_m\rangle \tag{2.5.24}$$

and similarly for the all-electron wavefunction,

$$|\psi\rangle = \sum_m c_m |\varphi_m\rangle = \hat{\mathcal{T}} |\tilde{\psi}\rangle \tag{2.5.25}$$

which can be thought as a linear transform $\hat{\mathcal{T}}$ of the smooth wavefunction. Therefore, the all-electron wavefunction in space with augmentation regions centred at $\mathbf{R}$ can be explicitly written as

$$\begin{aligned} |\psi\rangle &= (1 + \sum_{\mathbf{R}} \hat{\mathcal{T}}_{\mathbf{R}}) |\tilde{\psi}\rangle \\ &= |\tilde{\psi}\rangle + \sum_{\mathbf{R}m} (|\varphi_{\mathbf{R}m}\rangle - |\tilde{\varphi}_{\mathbf{R}m}\rangle) \langle \tilde{p}_{\mathbf{R}m} | \tilde{\psi}\rangle \end{aligned} \tag{2.5.26}$$

where $\langle \tilde{p}_i|$ is a projector that is biorthogonal to the smooth wavefunction in each sphere,

$$\langle \tilde{p}_i | \tilde{\varphi}_j \rangle = \delta_{ij} \tag{2.5.27}$$

hence the coefficients of the expansions can be obtained.

It is worth mentioning that the PAW formalism also allows the calculation of an expectation value of an operator directly from the smooth wavefunction. Notice that the expression for an expectation value of an operator $\hat{A}$ acting on an all-electron

wavefunction can be written as a transformed operator $\hat{\tilde{A}}$ acting on a smooth wavefunction,

$$\langle\psi|\hat{A}|\psi\rangle = \langle\tilde{\psi}|\hat{\mathcal{T}}^{\dagger}\hat{A}\hat{\mathcal{T}}|\tilde{\psi}\rangle = \langle\tilde{\psi}|\hat{\tilde{A}}|\tilde{\psi}\rangle\,. \tag{2.5.28}$$

Hence, for any local operator acting within the augmented regions around an atom, the expression for operators acting on the smooth wavefunction can be given as,

$$\hat{\tilde{A}} = \hat{A} + \sum_{mm'} |\tilde{p}_m\rangle \Big(\langle\varphi_m|\hat{A}|\varphi_{m'}\rangle - \langle\tilde{\varphi}_m|\hat{A}|\tilde{\varphi}_{m'}\rangle \Big) \langle\tilde{p}_{m'}|\,. \tag{2.5.29}$$

2.6 Self-consistent procedures and geometry optimisations

2.6.1 Achieving self-consistency

As it was mentioned, the Roothaan-Hall equation (2.5.5) needs to be solved iteratively because the matrix elements of the Fock matrix depend on the matrix elements of the coefficient matrix. The pragmatic standard procedures for achieving the 'self-consistency' of the Hartree-Fock calculation are illustrated in figure 2.2. In the begging of the procedures, the external potential is specified and the choice of the basis set is made. From this information, initial guess of orbitals is made through methods such as superpositions of atomic densities or diagonalisation of the Hamiltonian without the two-electron term. Then, the Fock matrix elements can be readily calculated with coefficients of initial orbitals and precalculated two-electron integrals of the basis set. The diagonalisation of the Fock matrix will yield a new matrix of coefficients, which is again used to calculate matrix elements of a new Fock matrix. These procedures are repeated until the convergence criteria such as differences in energy and coefficients are satisfied, giving the self-consistent solution to the Roothaan-Hall equation. Once the self-consistency is achieved, output quantities such as the total energy and forces are calculated, and the procedure is stopped. To achieve self-consistency to Kohn-Sham equations, similar approaches can be taken except the expression of the Fock matrix is modified.

There are alternative approaches available to accelerate the iterative procedures of the Roothaan-Hall method. These methods often relies on information of orbital gradients and Hessians, and have been successful in converging both Hartree-Fock and Kohn-Sham wavefunctions. Here, the algorithms of some other numerical methods will be discussed in the later subsection.

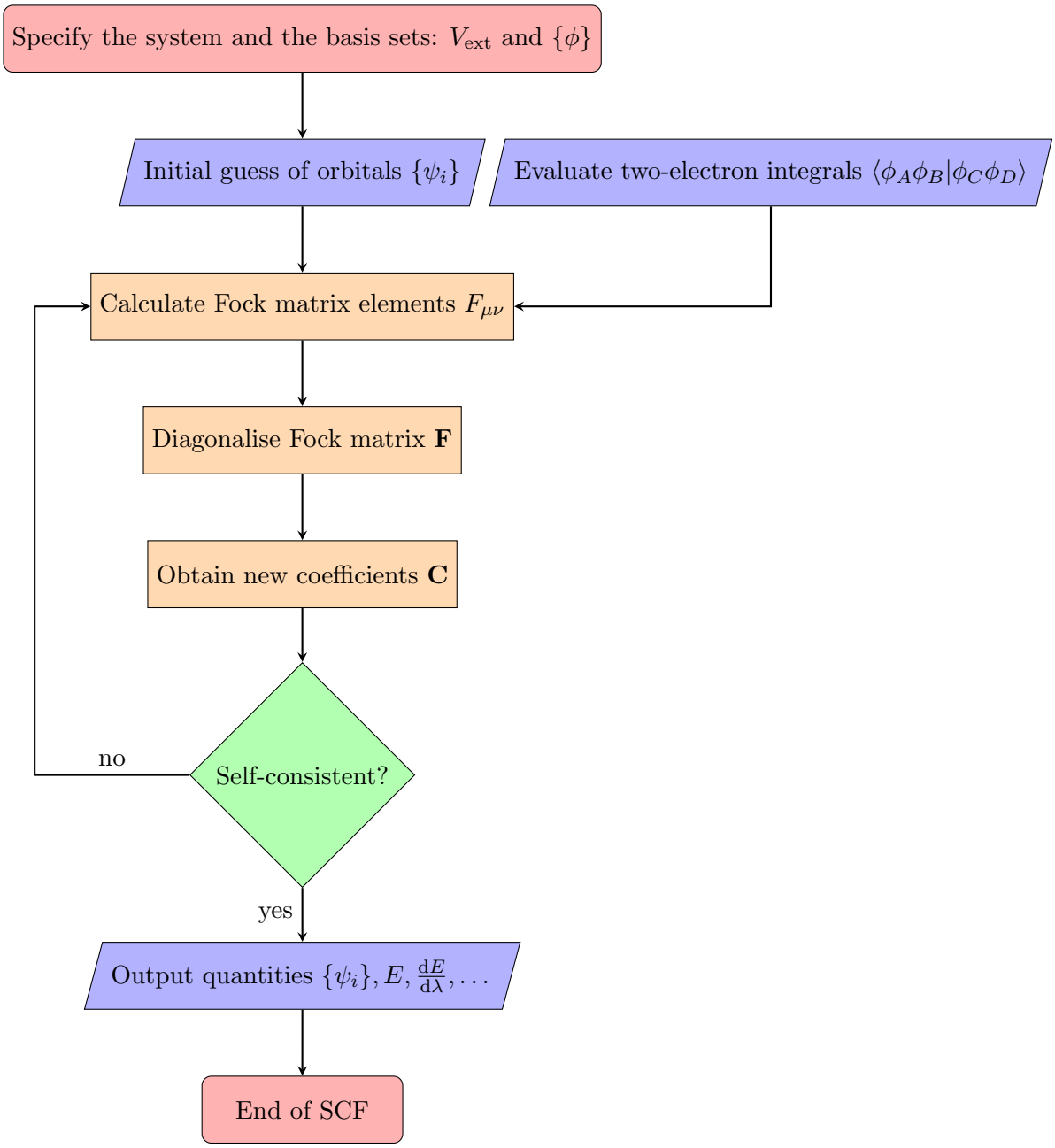


Figure 2.2: A simplified schematic of the self-consistent procedures for a Roothaan-Hall equation.

2.6.2 Molecular geometry optimisation

Once the electronic energy and wavefunction are obtained from the SCF cycle, one may wish to use the information available to optimise the geometry of the molecule. This optimisation process also has to be performed iteratively, since the energy of the molecule is not explicit in atomic coordinates. Here, the general consideration for geometry optimisations will be introduced.

Consider the potential energy surface V of a molecular system described as a function of nuclear coordinates $\mathbf{R}$ within the Born-Oppenheimer approximation. If the positions of nuclei are moved to $\mathbf{R}'$ with a displacement $\mathbf{u} = \mathbf{R}' - \mathbf{R}$, the Taylor expansion for the potential energy surface about $\mathbf{R}$ in a matrix form is given as

$$V(\mathbf{R}') = V(\mathbf{R}) + \mathbf{u}^\top \mathbf{f}(\mathbf{R}) + \frac{1}{2}\mathbf{u}^\top \mathbf{H}(\mathbf{R})\mathbf{u} + \cdots \tag{2.6.1}$$

where $f_i = \partial V(\mathbf{R})/\partial R_i$ is the gradient and $H_{ij} = \partial^2 V(\mathbf{R})/\partial R_i \partial R_j$ is the Hessian. At a stationary point $\mathbf{R}^0$, the gradient of the potential energy surface should be equal to zero $\mathbf{f}(\mathbf{R^0}) = \mathbf{0}$ by definition, and equation 2.6.1 becomes

$$V(\mathbf{R}') = V(\mathbf{R}^0) + \frac{1}{2}\mathbf{u}^\top \mathbf{H}(\mathbf{R}^0)\mathbf{u} \tag{2.6.2}$$

where third and higher-order terms are dropped. Similarly, the expression of the gradients is given as

$$\mathbf{f}(\mathbf{R}') = \mathbf{f}(\mathbf{R}) + \mathbf{H}(\mathbf{R})\mathbf{u} \tag{2.6.3}$$

and setting $\mathbf{R}' = \mathbf{R}^0$ gives

$$\mathbf{f}(\mathbf{R}) = -\mathbf{H}(\mathbf{R})\mathbf{u} \tag{2.6.4}$$

Therefore, to seek the coordinates of a stationary point $\mathbf{R}^0$ from any point $\mathbf{R}$ close enough that energy expression is quadratic, the displacement from that point can be given as

$$\mathbf{u} = -\mathbf{H}^{-1}(\mathbf{R})\mathbf{f}(\mathbf{R}) \tag{2.6.5}$$

given that the inversion of the Hessian matrix exists. In the same way, an estimation of the potential energy can be also obtained from

$$V(\mathbf{R}^0) = V(\mathbf{R}) - \frac{1}{2}\mathbf{f}^\top(\mathbf{R})\mathbf{H}^{-1}(\mathbf{R})\mathbf{f}(\mathbf{R}). \tag{2.6.6}$$

These equations provides the basis of efficient procedures for finding extrema on the potential energy surfaces not explicit in nuclear coordinates of molecules.

Needless to say, in practice, potential energy surfaces of molecules are far from being harmonic and molecular geometries have to be found iteratively. Moreover,

while forces can be easily calculated by using the Hellman-Feynman theorem (equation 2.1.9) given that the basis set is complete, the direct calculation of second-order properties are not trivial. There are various numerical approaches available to optimise geometries depending on the availability of information such as the gradients and Hessian, and these will be introduced in the next subsection.

2.6.3 Optimisation techniques

Both self-consistent field procedures in electronic structures and finding the geometries of molecules result in a minimisation (optimisation) problem of an energy functional defined by a set of parameters such as coefficients of orbitals or atomic coordinates. There are wide ranges of techniques available in the realm of numerical analysis, and commonly used techniques in, for an example, geometry optimisations are given in the following.

1. **Steepest decent and conjugate gradient methods**
   When only the gradients are available, the best direction of displacements from a point to reach the stationary point is given as

$$d_i^{(k)} = -\left[\frac{\partial E}{\partial R_i}\right]_{R_i=R_i^{(k)}}, \tag{2.6.7}$$

   which is the steepest-decent direction. Therefore, a better new point in the $k+1$th step

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} + \alpha^{(k+1)}\mathbf{d}^{(k)} \tag{2.6.8}$$

   can be found such that the step size $\alpha^{(k+1)}$ along that direction minimises the energy functional (line search). This procedure is continued until the gradient and the step size become negligibly small. The steepest-decent method can be inefficient when the point is close to the stationary point, where the gradients are relatively small.

   An improvement to the steepest-decent method can made by modifying the direction of the displacement, as suggested in the conjugate-gradient method, given as

$$\mathbf{d}^{(k+1)} = \mathbf{f}^{(k+1)} + \gamma^{(k+1)}\mathbf{d}^{(k)} \tag{2.6.9}$$

   In the Fletcher–Reeves version of the conjugate-gradient method, $\gamma^{(k+1)}$ is calculated as

$$\gamma^{(k+1)} = \frac{\mathbf{f}^{(k+1)} \cdot \mathbf{f}^{(k+1)}}{\mathbf{f}^{(k)} \cdot \mathbf{f}^{(k)}}. \tag{2.6.10}$$

In this way, the conjugate gradient method takes into an account of the history of previous steps in minimisation process.

2. **Newton-Raphson and quasi-Newton methods**
   When a full Hessian is available, the best displacement on a quadratic surface can be given by equation 2.6.5. Therefore, in the Newton-Raphson method, iterative scheme can be proposed to optimise the geometry,

$$\mathbf{R}^{(k+1)} = \mathbf{R}^{(k)} - \mathbf{H}^{(k)^{-1}}\mathbf{f}^{(k)} \tag{2.6.11}$$

   given that the Hessian is invertible. However, the calculation of second derivatives of energy is not trivial compared to the calculation of first derivatives, and, in often case, requires numerical differentiation with many electronic structure methods.

   In quasi-Newton methods, the Hessian matrix is replaced with an approximation, and the need of explicit calculations of second derivatives is eliminated. In each step of the optimisation process, an update to the approximated Hessian is made by using the gradients at previous steps. There are various popular update schemes used in chemistry codes, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) and the direct inversion iterative sub-space (DIIS) methods.

In many computational packages, both gradient and Hessian-based methods are available for SCF procedures and geometry optimisations. However, their efficiency in convergence and ability to escape local minima depend on the type of the job and the molecular/solid system.

2.7 Theory of lattice dynamics

While, in the previous sections, we have focused on methodologies to approximate the electronic wavefunctions and energies, a significant portion of the finite-temperature effects such as change in the rate of reactions and structural phase-transitions, in both the molecular and solid-state systems, derive from the dynamic movements of atomic nuclei. Therefore, one needs to solve for the equations of motion for the entire system in order to take account of thermal conditions. The Born-Oppenheimer approximation is still useful in finding the positions and trajectories of atomic nuclei and in many cases they can be treated classically since their mass is much heavier in contrast to electrons (although some light atoms such as hydrogen may not be correctly modelled with classical mechanics). While solving the equations of motion can

be readily achieved for molecules within the harmonic approximations, the periodicity of crystals introduces additional complexity. Here, we will explore the fundamental principles of phonons in solids.

2.7.1 Force constants and dynamical matrices

Recall that in the Born-Oppenheimer approximation, since the movement of electrons is much faster than nuclei, the nuclear Schrödinger equations can be written independent from the electronic coordinates (equation 2.1.16). The potential energy surface is therefore a function of only nuclear coordinates. Now, consider the Taylor expansion of the potential energy surface in a periodic system, with respect to the displacements from the equilibrium position $\mathbf{u}_{l\kappa} = \mathbf{R}_{l\kappa} - \mathbf{R}^0_{l\kappa}$,

$$\begin{aligned} V = V_0 &+ \sum_{l\kappa\alpha} \Phi_{l\kappa\alpha} u_{l\kappa\alpha} \\ &+ \frac{1}{2} \sum_{l\kappa\alpha, l'\kappa'\alpha'} \Phi_{l\kappa\alpha, l'\kappa'\alpha'} u_{l\kappa\alpha} u_{l'\kappa'\alpha'} \\ &+ \frac{1}{3!} \sum_{l\kappa\alpha, l'\kappa'\alpha', l''\kappa''\alpha''} \Phi_{l\kappa\alpha, l'\kappa'\alpha', l''\kappa''\alpha''} u_{l\kappa\alpha} u_{l'\kappa'\alpha'} u_{l''\kappa''\alpha''} + \cdots \end{aligned} \tag{2.7.1}$$

where l, κ and α represent the indices for unit cells, atoms (in the cell) and Cartesian axes respectively, and the notations $\Phi_{l\kappa\alpha} = \frac{\partial V}{\partial u_{l\kappa\alpha}}$, $\Phi_{l\kappa\alpha, l'\kappa'\alpha'} = \frac{\partial^2 V}{\partial u_{l\kappa\alpha} \partial u_{l'\kappa'\alpha'}}$, etc. are used and called force constant tensors. When the system is in the equilibrium positions, the net total force acting on each atom should be equal to zero by definition, and hence the first term vanishes. We could further simplify this equation by also dropping the third and higher terms and setting $V_0 = 0$, and we obtain the expression for a harmonic potential energy surface, which is given by

$$V^{\text{harm}} = \frac{1}{2} \sum_{l\kappa\alpha, l'\kappa'\alpha'} \Phi_{l\kappa\alpha, l'\kappa'\alpha'} u_{l\kappa\alpha} u_{l'\kappa'\alpha'}. \tag{2.7.2}$$

We can now write the expression for the classical Hamiltonian for atomic nuclei under the harmonic potentials

$$\begin{aligned} H^{\text{harm}} &= T + V^{\text{harm}} \\ &= \sum_{l\kappa\alpha} \frac{1}{2} m_\kappa \left(\frac{d u_{l\kappa\alpha}}{dt} \right)^2 + \frac{1}{2} \sum_{l\kappa\alpha, l'\kappa'\alpha'} u_{l\kappa\alpha} \Phi_{l\kappa\alpha, l'\kappa'\alpha'} u_{l'\kappa'\alpha'} \\ &= \sum_{l\kappa\alpha} \frac{1}{2} \left(\frac{d \tilde{u}_{l\kappa\alpha}}{dt} \right)^2 + \frac{1}{2} \sum_{l\kappa\alpha, l'\kappa'\alpha'} \tilde{u}_{l\kappa\alpha} \tilde{\Phi}_{l\kappa\alpha, l'\kappa'\alpha'} \tilde{u}_{l'\kappa'\alpha'} \end{aligned} \tag{2.7.3}$$

where mass-reduced displacements $\tilde{u}_{l\kappa\alpha} = \sqrt{m_\kappa} u_{l\kappa\alpha}$ are used for brevity in the last line. Note that the mass-reduced force constant matrix $\tilde{\Phi}_{l\kappa\alpha,l'\kappa'\alpha'}$ is symmetric and can be diagonalised as $\tilde{\Phi} = \mathrm{U}\Omega^2\mathrm{U}^\top$, in which Ω is a diagonal matrix composed of ω_ξ and U is the orthogonal matrix of the corresponding eigenvectors $w_{l\kappa\alpha}(\xi)$. Therefore we can now rewrite the Hamitonian (equation 2.7.3) in the matrix notation

$$\begin{aligned} H^{\mathrm{harm}} &= \frac{1}{2}\dot{\tilde{\mathbf{u}}}^\top \dot{\tilde{\mathbf{u}}} + \frac{1}{2}\tilde{\mathbf{u}}^\top \tilde{\Phi}\tilde{\mathbf{u}} \\ &= \frac{1}{2}(\mathrm{U}^\top\dot{\tilde{\mathbf{u}}})^\top(\mathrm{U}^\top\dot{\tilde{\mathbf{u}}}) + (\mathrm{U}^\top\tilde{\mathbf{u}})^\top\Omega^2(\mathrm{U}^\top\tilde{\mathbf{u}}) \\ &= \frac{1}{2}\dot{\mathbf{Q}}^\top\dot{\mathbf{Q}} + \frac{1}{2}\mathbf{Q}^\top\Omega^2\mathbf{Q} \end{aligned} \tag{2.7.4}$$

where $\tilde{\mathbf{u}}$ is a column vector of $\tilde{u}_{l\kappa\alpha}$, $\mathbf{Q} = \mathrm{U}^\top\tilde{\mathbf{u}}$ gives the normal coordinates, and the Newtonian notation of the time-derivative is used.

Alternatively, the eigenvalue problem of the force-constant matrix can be given as

$$\sum_{l'\kappa'\alpha'} \tilde{\Phi}_{l\kappa\alpha,l'\kappa'\alpha'} w_{l'\kappa'\alpha'}(\xi) = \omega_\xi^2 w_{l\kappa\alpha}(\xi) \tag{2.7.5}$$

which is also the equation of motion. By using the Bloch theorem, we guess the form of the eigenvector for a given $\mathbf{q}$ as

$$w_{l\kappa\alpha,\mathbf{q}}(\xi) = \frac{1}{\sqrt{N}} W_{\kappa\alpha}(\xi) e^{i\mathbf{q}\cdot\mathbf{R}^0_{l\kappa}} \tag{2.7.6}$$

where $W_{\kappa\alpha}(\xi)$ is a periodic function and N ensures the eigenvector $w_{l\kappa\alpha,\mathbf{q}}(\xi)$ is normalised to unity. Inserting the equation 2.7.6 into the equation 2.7.5, we obtain a new eigenvalue equation

$$\sum_{\kappa'\alpha'} D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) W_{\kappa'\alpha'}(\mathbf{q}\nu) = \omega^2_{\mathbf{q}\nu} W_{\kappa\alpha}(\mathbf{q}\nu), \tag{2.7.7}$$

where the normal mode index ξ is replaced with the phonon band index ν at a given $\mathbf{q}$ and

$$\begin{aligned} D_{\kappa\alpha,\kappa'\alpha'}(\mathbf{q}) &= \sum_{ll'} \frac{e^{-i\mathbf{q}\cdot\mathbf{R}^0_{l\kappa}}}{\sqrt{N}} \tilde{\Phi}_{l\kappa\alpha,l'\kappa'\alpha'} \frac{e^{i\mathbf{q}\cdot\mathbf{R}^0_{l'\kappa'}}}{\sqrt{N}} \\ &= \frac{1}{N\sqrt{m_\kappa m_{\kappa'}}} \sum_{ll'} \Phi_{l\kappa\alpha,l'\kappa'\alpha'} e^{i\mathbf{q}\cdot(\mathbf{R}^0_{l'\kappa'} - \mathbf{R}^0_{l\kappa})} \\ &= \frac{1}{\sqrt{m_\kappa m_{\kappa'}}} \sum_{l'} \Phi_{0\kappa\alpha,l'\kappa'\alpha'} e^{i\mathbf{q}\cdot(\mathbf{R}^0_{l'\kappa'} - \mathbf{R}^0_{0\kappa})} \end{aligned} \tag{2.7.8}$$

is called the dynamical matrix. In the last line of the equation 2.7.8, the translational symmetry of the crystal is applied. As a result of the presence of the phase factors,

the dynamical matrix is not symmetric but hermitian and the diagonalisation of the matrix

$$\mathrm{D}(\mathbf{q}) = \mathrm{U}(\mathbf{q})\Omega^2(\mathbf{q})\mathrm{U}^\dagger(\mathbf{q}) \tag{2.7.9}$$

will only give real eigenvalues. If all of the eigenvalues in the matrix Ω^2 for any given $\mathbf{q}$ are positive, the crystal is said to be *dynamically* stable, while if it contains negative eigenvalues, then the crystal will undergo a structural transition by breaking the crystal symmetry.

2.7.2 Quantum-mechanical treatment of phonons

The classical Hamiltonian of the normal coordinates in equation 2.7.3 can be rewritten in terms of quantum operators:

$$\hat{H}^{\mathrm{harm}} = \frac{1}{2}\sum_{\mathbf{q}\nu}\left(\hat{\Pi}^\dagger_{\mathbf{q}\nu}\hat{\Pi}_{\mathbf{q}\nu} + \omega^2_{\mathbf{q}\nu}\hat{Q}^\dagger_{\mathbf{q}\nu}\hat{Q}_{\mathbf{q}\nu}\right) \tag{2.7.10}$$

where $\hat{Q}_{\mathbf{q}\nu}$ and $\hat{\Pi}_{\mathbf{q}\nu}$ are the operators corresponding to the normal coordinate and the corresponding momentum variable respectively. Note, however, that unlike the conventional $\hat{p}$ and $\hat{x}$ operators, the operators $\hat{\Pi}_{\mathbf{q}\nu}$ and $\hat{Q}_{\mathbf{q}\nu}$ are not hermitian and their adjoints are given as

$$\hat{Q}^\dagger_{\mathbf{q}\nu} = \hat{Q}_{-\mathbf{q}\nu}, \quad \hat{\Pi}^\dagger_{\mathbf{q}\nu} = \hat{\Pi}_{-\mathbf{q}\nu} \tag{2.7.11}$$

Nevertheless, if we define a new set of operators

$$\begin{aligned}\hat{Q}_{\mathbf{q}\nu} &= \sqrt{\frac{\hbar}{2\omega_{\mathbf{q}\nu}}}\left(a_{\mathbf{q}\nu} + a^\dagger_{-\mathbf{q}\nu}\right)\\ \hat{\Pi}_{\mathbf{q}\nu} &= -i\sqrt{\frac{\hbar\omega_{\mathbf{q}\nu}}{2}}\left(a_{\mathbf{q}\nu} - a^\dagger_{-\mathbf{q}\nu}\right)\end{aligned} \tag{2.7.12}$$

and, the Hamiltonian (equation 2.7.10) can be hence rewritten into

$$\hat{H}^{\mathrm{harm}} = \sum_{\mathbf{q}\nu}\hbar\omega_{\mathbf{q}\nu}\left(a^\dagger_{\mathbf{q}\nu}a_{\mathbf{q}\nu} + \frac{1}{2}\right) \tag{2.7.13}$$

which is simply the ladder operator formalism for multiple quantum harmonic oscillators. The creation and annihilation operators have the following commutator relations,

$$[\hat{a}_{\mathbf{q}\nu}, \hat{a}^\dagger_{\mathbf{q}'\nu'}] = \delta(\mathbf{q}-\mathbf{q}')\delta_{\nu\nu'}, \quad [\hat{a}_\nu(\mathbf{q}), \hat{a}_{\nu'}(\mathbf{q}')] = 0, \quad [\hat{a}^\dagger_\nu(\mathbf{q}), \hat{a}^\dagger_{\nu'}(\mathbf{q}')] = 0 \tag{2.7.14}$$

and by using these operators, the wavefunction of phonons can be represented in the second quantisation formalism,

$$|n_1, n_2, \cdots, n_j, \cdots\rangle = \frac{1}{\sqrt{\prod_i n_i!}} \left(\hat{a}_1^\dagger\right)^{n_1} \left(\hat{a}_2^\dagger\right)^{n_2} \cdots \left(\hat{a}_j^\dagger\right)^{n_j} \cdots |0\rangle \tag{2.7.15}$$

This expression of the wavefunction is symmetric (e.g. does not change the sign of the wavefunction) under the exchange of two particles, implying that phonons are *bosons*.

Since phonons are bosons, unlike electrons which are fermions, more than one particle can occupy the same mode, and the occupation number of each mode is governed by the Bose-Einstein distribution, given as

$$\langle \hat{n}_{\mathbf{q}\nu} \rangle = \langle \hat{a}_{\mathbf{q}\nu}^\dagger \hat{a}_{\mathbf{q}\nu} \rangle = \frac{1}{\exp\left(\hbar\omega_{\mathbf{q}\nu}/k_\mathrm{B}T\right) - 1} \tag{2.7.16}$$

where T is the temperature and k_B is the Boltzmann constant. The statistical average harmonic phonon energy is therefore given as

$$\langle \hat{H}^\mathrm{harm} \rangle = \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu} \left(\langle \hat{n}_{\mathbf{q}\nu} \rangle + \frac{1}{2} \right). \tag{2.7.17}$$

2.7.3 Thermodynamics of phonon gas

Since the frequencies of harmonic phonons are independent from each other (e.g. non-interacting), the partition functions for harmonic phonons can be given as a product of the partition function of each state,

$$Z = \prod_{\mathbf{q}\nu} \frac{\exp\left(-\hbar\omega_{\mathbf{q}\nu}/2k_\mathrm{B}T\right)}{1 - \exp\left(-\hbar\omega_{\mathbf{q}\nu}/k_\mathrm{B}T\right)}. \tag{2.7.18}$$

From this partition function, the Helmholtz free energy can be calculated as

$$\begin{aligned} F &= -k_\mathrm{B}T \ln Z \\ &= \frac{1}{2} \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu} + k_\mathrm{B}T \sum_{\mathbf{q}\nu} \ln\left[1 - \exp(-\hbar\omega_{\mathbf{q}\nu}/k_\mathrm{B}T)\right] \end{aligned} \tag{2.7.19}$$

and the phonon contribution to the entropy is

$$\begin{aligned} S &= -\frac{\partial F}{\partial T} \\ &= \frac{1}{2T} \sum_{\mathbf{q}\nu} \hbar\omega_{\mathbf{q}\nu} \coth((\hbar\omega_{\mathbf{q}\nu})/2k_\mathrm{B}T) - k_\mathrm{B} \sum_{\mathbf{q}\nu} \ln\left[2\sinh(\hbar\omega_{\mathbf{q}\nu}/2k_\mathrm{B}T)\right]. \end{aligned} \tag{2.7.20}$$

Similarly, by using the Maxwell's relation, the phonon pressure can be obtained by partially differentiating the Helmholtz free energy with respect to the volume at a constant temperature

$$\begin{aligned} P_{\mathrm{ph}} &= -\left(\frac{\partial F}{\partial V}\right)_T \\ &= -\frac{\partial}{\partial V}\left[\frac{1}{2}\sum_{\mathbf{q}\nu}\hbar\omega_{\mathbf{q}\nu}\right] - \sum_{\mathbf{q}\nu}\left(\hbar\frac{\partial\omega_{\mathbf{q}\nu}}{\partial V}\right)\frac{1}{\exp(-\hbar\omega_{\mathbf{q}\nu}/k_{\mathrm{B}}T)-1}. \end{aligned} \tag{2.7.21}$$

Assuming that the frequencies of phonon modes do not change, the temperature dependency of phonon gas pressure does not appear in the first term and only appears in the second term through the Bose-Einstein distribution. However, in the harmonic approximation, the phonon frequencies do not have an explicit dependency on the volume and all terms with factor $\frac{\partial\omega_{\mathbf{q}\nu}}{\partial V}$ become zero. This is an indication that the thermal expansion of solids, a common phenomenon, cannot be explained with harmonic phonons, and that they are essentially a product of anharmonicity. The dependency of volume on phonon frequencies can be quantitatively calculated as the Grüneisen mode parameters, which is given as

$$\gamma_{\mathbf{q}\nu} = -\frac{V}{\omega_{\mathbf{q}\nu}}\frac{\partial\omega_{\mathbf{q}\nu}}{\partial V} = -\frac{\partial\ln\omega_{\mathbf{q}\nu}}{\partial\ln V} \tag{2.7.22}$$

and the factor of -1 is to ensure that the values are usually positive, as expansions of the cells usually decreases phonon frequencies due to reduced interactions between atoms.

2.7.4 Quasi-harmonic approximations

Explicit considerations of anharmonic effects on phonon frequencies are much more expensive than the harmonic calculations, as evaluations of third and higher order force constants are required. Nevertheless, it is still possible to incorporate anharmonic effects 'numerically' within the framework of harmonic approximations. In the quasi-harmonic approximation (QHA), frequencies of harmonic phonons at different constraints of volumes are calculated, allowing the effects of anhamonicity on the phonon frequencies to be incorporated. For example, the mode Grüneisen parameters (equation 2.7.22) can be obtained by replacing the derivative with a numerical derivative. Phonon thermal properties, such as the Helmholtz free energy and entropy, can be also calculated as a function of both temperature and volume in the similar approach. Once calculated, we can then perform the Legendre transformation

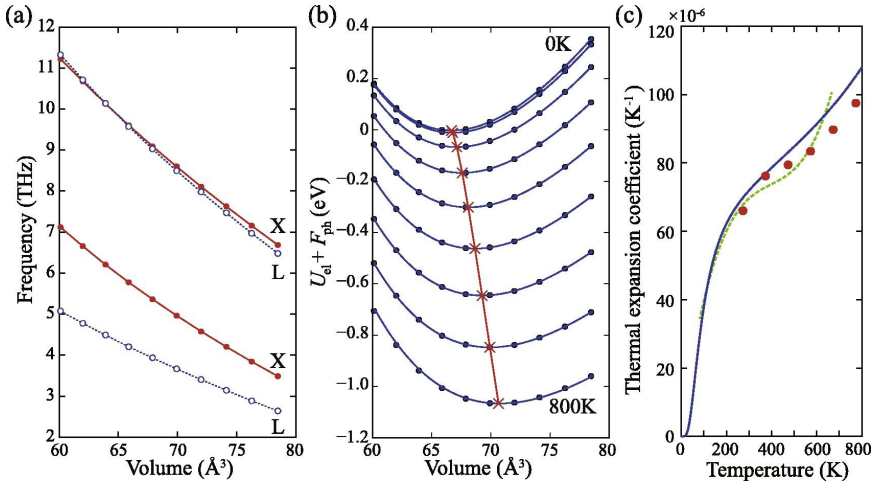


Figure 2.3: Illustrations of quasi-harmonic approximation applied to Al. Subplot (a) shows the volume dependencies of phonon modes at a given reciprocal point, (b) shows the temperature dependencies of free energy-volume curves (blue) and the change in the equilibrium volumes (red), and (c) shows the thermal expansion coefficients calculated within the QHA (blue) and experimentally observed (red circles and green lines). These figures are obtained from reference [61], licensed under CC-BY-4.0.

to convert the Helmholtz free energy, which is a function of T and V, into the Gibbs free energy, a function of T and P. Specifically in this case, the transformation finds the volume such that it minimises the Helmholtz free energy plus PV,

$$G(T,P) = \min_V \left[U_0(V) + F_{\text{ph}}(T,V) + PV\right] \tag{2.7.23}$$

where U_0 is the lattice internal energy of the crystal, F_{ph} is the phonon contribution to the Helmholtz free energy and the P is the external pressure. An example of the application of the QHA is illustrated in the figure 2.3. This methodology will be used in Chapter 5.

2.8 Summary

In this section, I have presented several methodologies for calculating energies of electronic and nuclear wavefunctions within the (adiabatic) Born-Oppenheimer approximation. The energies of electronic wavefunctions are obtained from further approximations to the electronic Schrödinger equation, in particular, with the introduction of various ansatze to the electronic wavefunction. Ansatze that were discussed in this sections were Hartree product, Slater determinant, Kohn-Sham, configuration interaction and coupled-cluster ansatze. Each ansatz captures different physics of fermionic many-body wavefunctions The (free) energies of nuclear wavefunctions are

obtained by solving the equations of motion on the potential energy surface with the harmonic approximations. In lattice, this will lead to the introduction of bosonic particle phonons, and their partition functions are used to various macroscopic thermodynamic properties, including Helmholtz free energies. To calculate Gibbs free energies, the quasi-harmonic approximations are used to take into account of volume changes.

Chapter 3

Gas-phase reactivity of partially unsaturated ruthenium carbonyl clusters with alkenes

3.1 Introduction

Transition metal clusters can exhibit unique catalytically behaviour that is not typical of their mononuclear counterparts, or indeed of metal nanoparticles [62]. There are, in principle, many ways in which transition metal clusters might catalyse reactions, perhaps most obviously by fragmenting to generate smaller catalytic active species or by aggregating to form colloidal species, as happens, for example, in the hydrogenation of phenylacetylene by $PtRu_5$ carbonyl cluster [63]. However, there are also alternative mechanisms in which the clusters remain intact throughout the catalytic cycles, as proposed, for example, in the mechanism of alkyne transformations catalysed by $H_4Pt_3Ru_6(CO)_{21}$ [64]. The surface-cluster analogy emphasises the structural similarities between metal clusters and surfaces, and has been used extensively to understand the structure and reactivity of metal surfaces [65]. The cluster-surface analogy draws parallels between coordination sites on a discrete molecular cluster and those on an exposed metallic surface, and has a lengthy history dating back to early work from Muetterties [66], Lewis, Johnson[67] and others [68, 69]. This concept has proved to be remarkably useful in understanding the bonding of ligands such as CO, which typically show very similar coordination modes on both surfaces and clusters. Whilst the analogy is undeniably appealing, it is clear that there are also substantial differences between bulk and discrete molecular clusters, and the boundaries between clusters, nano-particles and fragments of bulk remain a matter of some debate. In the context of catalysis using molecular species, it is also a significant challenge to iden-

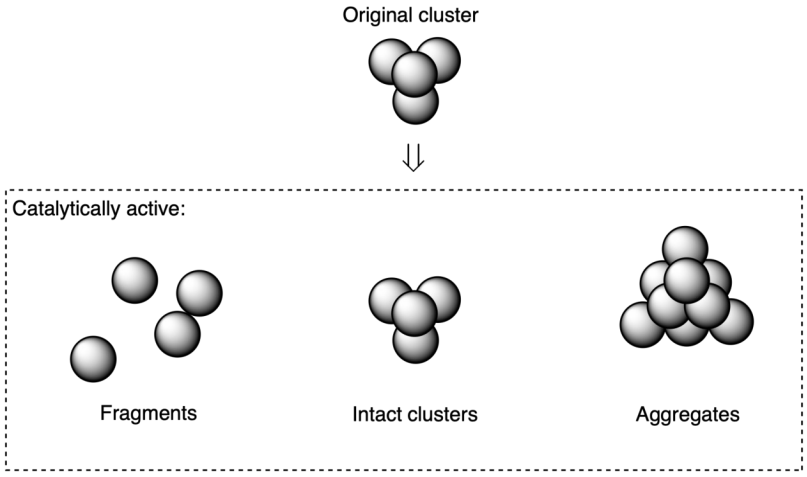


Figure 3.1: Examples of possible forms of catalytically active species from the original cluster.

tify the active species under turn-over conditions amongst the many small fragments that may be present, making it difficult to establish precise comparisons [70–73]. In more recent times, the attention of the catalysis community has turned to so-called double, triple and multi-atom catalysts (SACs, DACs, TACs etc) which are typically synthesised by the deposition of atomically dispersed fragments onto surfaces [74]. Metal-metal interactions are of potential significance in double/triple atom catalysts [75] as a result of their ability to moderate the energies of frontier orbitals as well as to admit bridging coordination sites. The precise role of metal-metal bonding in catalysis remains, however, a subject of active investigation [76].

The distinction between cluster catalysts and single/double/multi-atom catalysts depends to a large degree on the degree of coordinative unsaturation. In the former, the active species is typically considered to be a cluster that is near to saturation (in an electronic and coordinative sense), with only a small number of available sites for binding of reactant species. Single/double-atom catalysts, in contrast, are typically considered as 'naked' fragments, although this perspective neglects the important point that the surface support play a role as a 'ligand', occupying some of the available coordination sites of the atom or cluster. The theoretical tools appropriate to these limits are somewhat different - single configuration methods such as the ubiquitous DFT are most applicable to cases that approach electronic saturation, but the progressive loss of ligands typically introduces multi-configurational character that is best treated with techniques such as the Complete Active Space Self-Consistent Field (CASSCF). These wavefunction-based methods are, however, time consuming

and far from 'black-box', and applications to problems in catalysis are just beginning to be explored [77].

If we consider 'cluster catalysts' and double/triple/multi-atom catalysts to occupy two ends of a continuum defined by the degree of coordinative and electronic unsaturation, it is clear that intermediate cases may emerge where multiple vacant sites are present, but where the overall coordination environment more saturated than a typical multi-atom catalyst. Partially unsaturated clusters of this type are, by definition, unstable and their reactivity is therefore difficult to explore. However, recent developments in mass spectrometry, and in particular the emergence of energy-dependent electrospray ionisation mass spectrometry (EDESI-MS) offers an opportunity to probe reactivity of increasingly unsaturated clusters with a range of substrates [78–83]. By increasing the cone voltage between 0 and 200 V, the energy imparted to the ions by CID is increased, making it possible to observe the sequential dissociation of ligands and the subsequent reactivity of these partially coordinated species with substrate. The resultant 2-dimensional plots of m/z against cone voltage are rich in information (note that a conventional ESI-MS represents a 1-dimensional slice at a fixed cone voltage). They reveal not just the range of products possible, but also how their concentrations vary as a function of cone voltage.

Much of the work that has been reported has focused on metal carbonyl clusters, and in particular on the mixed carbonyl hydride, $[H_3Ru_4(CO)_{12}]^-$ [82]. This cluster, like its neutral parent $H_4Ru_4(CO)_{12}$, is known to catalyse reactions such as the water-gas shift reaction, but the anionic form is particularly amenable to ESI-MS studies by virtue of its negative charge. The data in Figure 3.3 (first reported in reference [82]) relate to the reaction with 1-hexene (C_6H_{12}), which is the focus of the computational analysis reported here. In the Figure, we see a series of peaks on the lower diagonal that correspond to successive loss of CO from the parent anion, $[H_3Ru_4(CO)_{12}]^-$. The second diagonal, displaced ∼25 eV to higher cone voltages, corresponds to binding of a single molecule of 1-hexene, the third diagonal to binding of two equivalents of 1-hexene *etc.* 1-hexene has the same mass as three equivalents of CO (84 in each case), so clusters aligned approximately vertically above each other differ by one molecule of 1-hexene and by three equivalents of CO. Low-intensity peaks on the second diagonal corresponding to the binding of 1-hexene to clusters with 11 and 10 CO ligands are apparent, but the first peak of significant intensity, centred at m/z 742, corresponds to the $[Ru_4(CO)_9C_6H_x]^-$ cluster, indicating that loss of three CO ligands is required for strong binding. A significant feature of these 2-dimensional spectra is the evidence for loss of H_2 under the prevailing reaction conditions. This is immediately apparent from

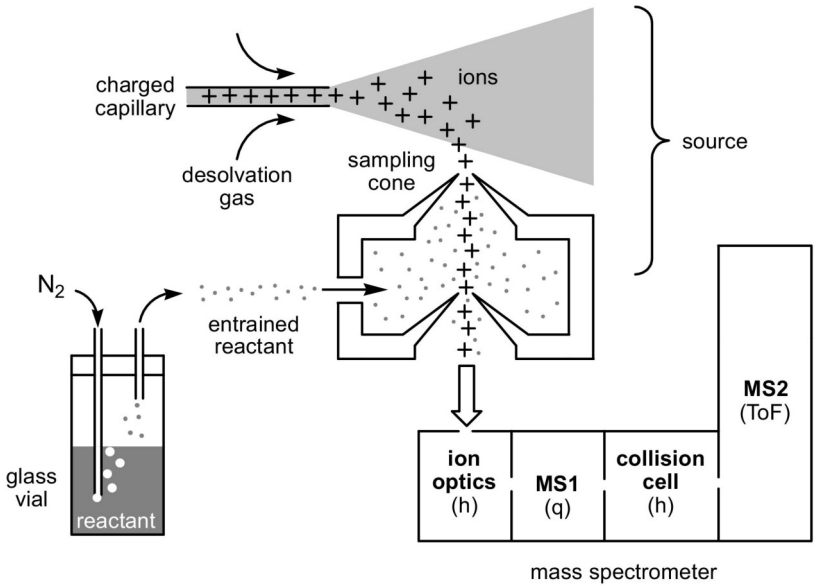


Figure 3.2: A schematics showing the basic procedures in the EDESI experiments. The [PPN][$H_3Ru_4(CO)_{12}$] salt was dissolved in dichloromethane, and pumped into the source housing. Ions ($[H_3Ru_4(CO)_{12}]^-$) are then exposed to the reactant gas (1-hexene), before being processed by the mass spectrometer. Source and desolvation gas temperatures were at 50 and 100 °C, respectively. Figure adapted with permission from Reference [82]. Copyright 2009 American Society for Mass Spectrometry.

the width of the peaks in Figure 3.3, but also from the small but distinct shift to lower m/z at successively higher cone voltages for peaks that are in approximate vertical alignment. As noted above, a simple substitution reaction would generate peaks for $[H_3Ru_4(CO)_m]^-$ and $[H_3Ru_4(CO)_{m-3}(C_6H_{12})]^-$ that would be in exact vertical alignment, and the shift to lower m/z therefore indicates that replacement of CO by 1-hexene is accompanied by loss of one or more molecules of H_2, either from the original hydride ligands of through activation of C-H bonds of the bound 1-hexene. As an example, the maximum intensity for the peak labelled as $[H_yRu_4(CO)_9(C_6H_x)]^-$ is found at 742 m/z, corresponding to $[Ru_4(CO)_9C_6H_{11}]^-$ and hence a loss of two equivalents of H_2 from the one-to-one adduct, $[H_3Ru_4(CO)_9(C_6H_{12})]^-$. As only three hydride ligands are present in the original cluster, the loss of two molecules of H_2 is indicative of C-H bond activation as well as H_2 reductive elimination. Indeed, McIndoe and co-workers have suggested that some of the highly CO-deficient clusters can generate benzene, which would require six C-H bond activation steps along with a C-C bond formation. As a result of this uncertainty in the precise number and location of the hydrogens (bound directly to the Ru_4 cluster or part of the hydrocarbon chain), we label the compositions in Figure 3.3 as $[H_yRu_4(CO)_9(C_6H_x)]^-$, where only the total

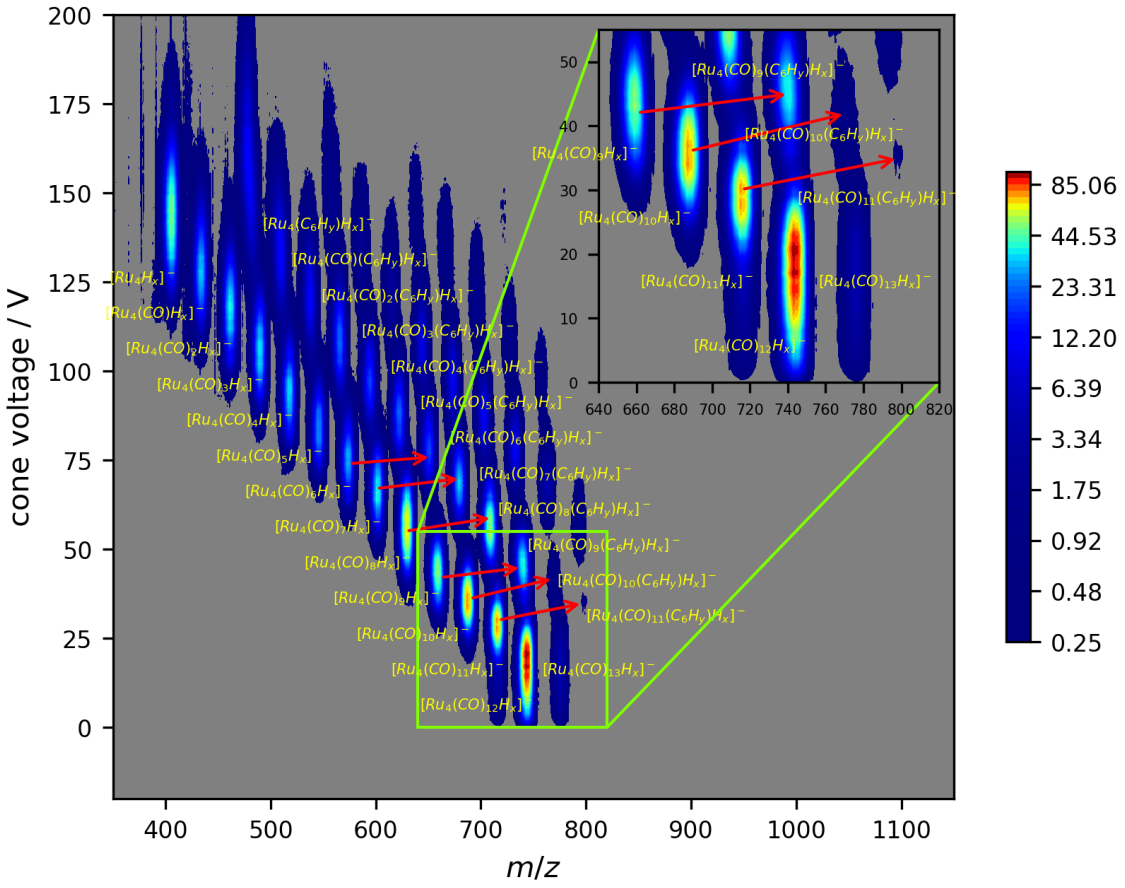


Figure 3.3: EDESI-MS spectrum of $[H_3Ru_4(CO)_{12}]^-$ with 1-hexene (original data reported in references [78, 80, 82–85]). Intensities are given in arbitrary units on a logarithmic scale. The red arrows identify pairs that are linked by the loss of three CO ligands and the binding of 1-hexene (for example $[H_xRu_4(CO)_{12}]^-$ and $[H_xRu_4(CO)_9(1\text{-hexene})]^-$). The zoomed low-voltage/high-mass region indicates the area of interest for this study.

hydrogen content, $x + y$, is accessible from the experiment.

As is the case with all mass spectrometric data, structural information can only be inferred from knowledge of structures of the reactants or closely-related compounds and, increasingly, by recourse to computation. DFT is the tool of choice in most studies of metal carbonyl clusters, offering as it does an attractive balance between accuracy and computational expedience. In this work, we use DFT to interrogate the data summarised in Figure 3.3, but we also use coupled cluster techniques (specifically the LNO-CCSD(T) method) as a benchmarking exercise. The local correlation methods have emerged in recent years as a viable method to reduce the scaling of coupled-cluster calculations, making clusters with four or more transition metal ions accessible to accurate computation. Our purpose is to survey the potential energy surfaces for candidate structures for the low-cone-voltage/high mass clusters highlighted in the inset in Figure 3.3. This region captures, at least potentially, all of the important processes including CO loss, binding of 1-hexene, activation of C-H bonds, formation of H-H bonds and loss of H_2. In particular, we are interested in investigating the following reaction:

$$[H_3Ru_4(CO)_m]^- + \text{1-hexene} \longrightarrow [H_3Ru_4(CO)_mC_6H_{12-2x}]^- + xH_2$$

where the amount of dehydrogenation x is ambiguous from the experiment. In the reaction proposals, we will identify the precise number of carbonyl ligands on the cluster and dehydrogenation steps for these reactions to occur. We also pay attention to the way in which the relative orientation of vacant sites on the cluster (on the same Ru center or on adjacent centers) influences the subsequent reactivity.

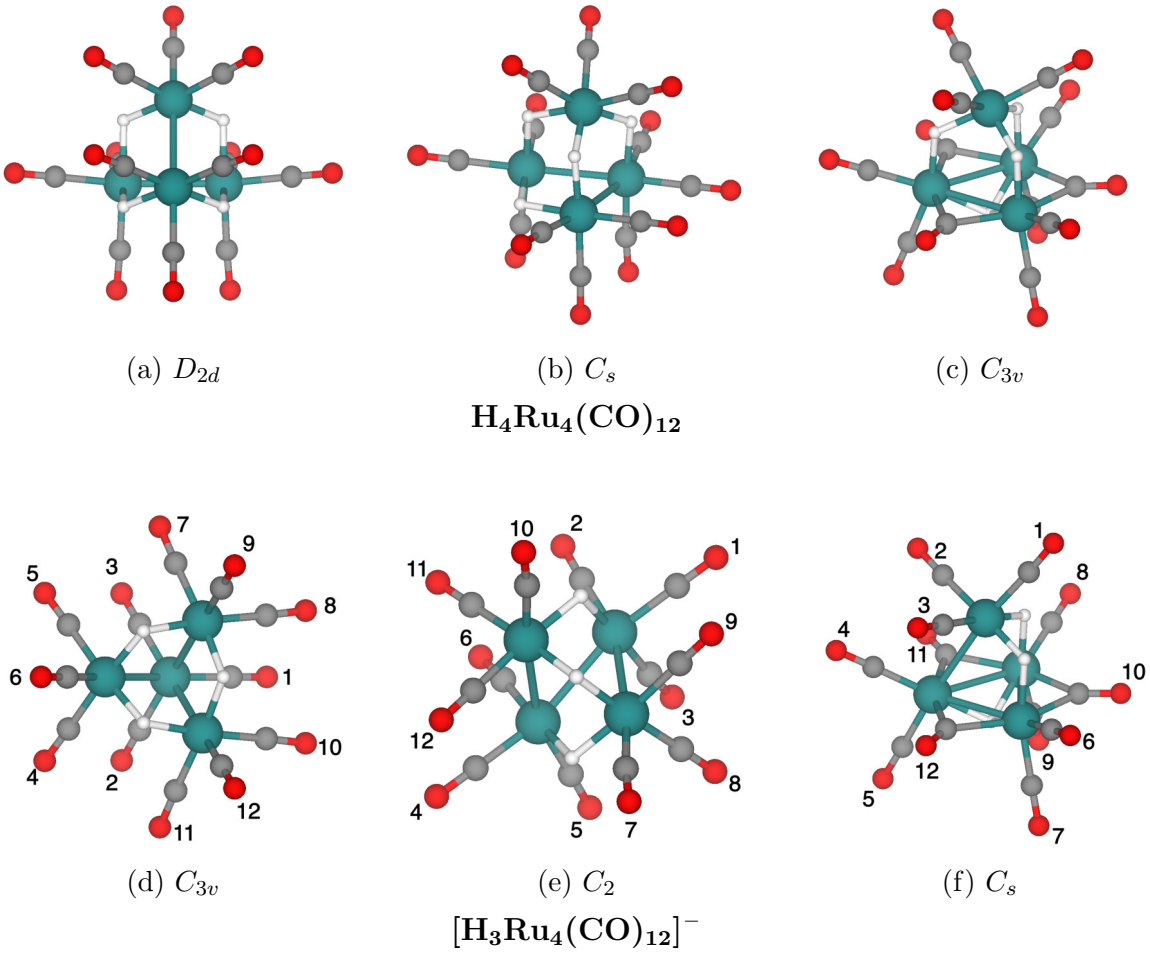
3.2 Computational Methods

Geometries for intermediate and transitional states of reaction mechanism were obtained by using DFT calculations as implemented in ADF 2021.104 [86]. All-electron Slater-type TZ2P orbital was used as a set of basis sets [87] and scalar ZORA correction was used [88–90]. Meta-GGA r^{2}SCAN functional [91] implemented in LibXC [92] along with D4 dispersion corrections parametrised by Ehlert *et al.* [93, 94] was employed as density functional approximation, as r^{2}SCAN-D4 has shown to be able to provide geometries as accurate as or better than PBE0-D4 can provide while retaining the computational cost much lower than hybrid functionals [94]. Spin-orbitals are unrestricted and 'Good' numerical quality was used throughout the calculation.

The LNO-CCSD(T) method [32–35, 95] as implemented in MRCC2022 [96] was employed to obtain coupled-cluster energies for an accurate potential energy surface for reactions. LNO-CCSD(T) has shown to capture static correlation effectively, and has been shown to provide accurate reaction energies in the area of organometallic catalysis [45, 46]. The truncation threshold set of 'Normal' was used and the augmented def2 basis set of def2-QZVPPD was used as the basis set for this calculation [59, 97, 98].

3.3 Structure of parent compounds: $H_4Ru_4(CO)_{12}$ and $[H_3Ru_4(CO)_{12}]^-$

Our analysis begins with a survey of the potential energy surfaces for the electronically saturated parent cluster $[H_3Ru_4(CO)_{12}]^-$ and its neutral protonated form, $H_4Ru_4(CO)_{12}$, where experimental data is available for comparison. $H_4Ru_4(CO)_{12}$ has been characterised crystallographically, and is known to adopt a D_{2d}-symmetric geometry with a tetrahedral Ru_4 core, twelve terminal CO ligands and hydrides bridging four of the six Ru-Ru edges. The presence of two short and four long Ru-Ru bonds (2.780 Å and 2.944 Å, respectively), is a clear indication of the position of the hydride ligands, with the longer separations being associated with the hydride-bridged edges. Several derivatives where one or more of the CO ligands is replaced by a phosphine have also been synthesised, some of which have a C_2-symmetric arrangement, such that the two short, unbridged, Ru-Ru bonds are adjacent to each other. Indeed an early extended Hückel study by Hoffmann, Mingos and co-workers noted that the hydrides were mobile, with low barriers to interconversion of different isomers.[99] Our survey of the potential energy surface at the BP86-D4/TZP level confirms that the D_{2d} cluster is the global minimum, with optimised Ru-Ru bond lengths of 2.80 Å (non-bridged) and 2.97 Å (bridged), within 0.03 Å of experiment in both cases. The C_s-symmetric isomer adopted by some of the phosphine-substituted clusters is only 4.7 kcal mol^{-1} higher in energy, but the unbridged and bridged Ru-Ru bond lengths of 2.79 Å and 2.94 Å average (av.), respectively, are very similar to their counterparts in the D_{2d}-symmetric variant. We have also identified a third, C_{3v}-symmetric, isomer with a face-capping (μ_3-) hydride, and three μ_2-CO ligands, which is only marginally less stable (6.5 kcal mol^{-1}) but is as yet unknown in the experimental literature (as crystals). Turning to the deprotonated form, $[H_3Ru_4(CO)_{12}]^-$, has been crystallised in two distinct isomers, again differing in the positions of the hydride bridges: in one, the three hydrides bridge the three edges of a triangular face, giving a C_{3v}-symmetric



(a) D_{2d} (b) C_s (c) C_{3v}

$H_4Ru_4(CO)_{12}$

(d) C_{3v} (e) C_2 (f) C_s

$[H_3Ru_4(CO)_{12}]^-$

Figure 3.4: Isomers of saturated (60-electron) clusters, $H_4Ru_4(CO)_{12}$ (top row) and $[H_3Ru_4(CO)_{12}]^-$ (bottom row). The numbering of the CO ligands in the latter case is relevant to studies of CO loss described in later sections.

Table 3.1: Energies of the three most stable isomers of $H_4Ru_4CO_{12}$ and $[H_3Ru_4CO_{12}]^-$. All values are given in kcal mol^{-1}, relative to the highest symmetry (D_{2d}- and C_{3v}) isomers of the two clusters.

	$H_4Ru_4(CO)_{12}$			$[H_3Ru_4(CO)_{12}]^-$		
	D_{2d}	C_s	C_{3v}	C_{3v}	C_2	C_s
BP86-D4	0	4.7	6.5	0	2.2	-1.6
PBE-D4	0	4.7	6.8	0	1.3	-2.2
r^{2}SCAN-D4	0	4.8	8.4	0	0.7	-2.3
M06-L	0	3.5	8.9	0	0.6	-1.3
PBE0-D4	0	5.3	10.8	0	1.3	-0.2
B3LYP-D4	0	5.1	12.3	0	2.9	3.7
LNO-CCSD(T)	0	4.9	13.7	0	0.5	0.8

geometry with three long Ru-H-Ru bonds (2.934 Å) and three shorter unbridged bonds (2.787 Å). The second, C_2-symmetric, isomer has hydrides on three adjacent edges, again giving three short (2.803 Å) and three longer (2.923 Å) bonds. Our BP86-D4-computed potential energy surface again confirms that both of these isomers are local minima, and that the former is 2.2 kcal mol^{-1} more stable than the latter. The DFT-optimised unbridged Ru-Ru bond lengths are 2.80 Å (av.) in both isomers, while the bridged counterparts are 2.93 Å (av.). We have also identified a third minimum, again with a μ_3- hydride, which is 1.6 kcal mol^{-1} more stable than the C_{3v}-symmetric alternative. This isomer has not been observed by experiment, and indeed there are no known structures of stable Ru/CO clusters with bridging CO ligands.

To explore the origins of this rather surprising result, we have recomputed the energies and fully relaxed the geometries of the three most stable isomers of each cluster using a range of density functionals, and also the LNO-CCSD(T) methodology (as a single point, using the geometries optimised with the r^{2}SCAN-D4 functional). The results, summarised in Table 3.1, indicate that different DFT functionals provide a consistent picture of the relative energies of the structures that have only terminal CO ligands (the D_{2d} and C_S isomers of $H_4Ru_4CO_{12}$ and the C_{3v} and C_2 isomers of $[H_3Ru_4CO_{12}]^-$, and also that these energies are broadly consistent with the LNO-CCSD(T) benchmark. However, all DFT functionals show a clear tendency to overstabilise the isomers with bridging carbonyl ligands and a face-capping hydride (C_{3v} isomer of $H_4Ru_4CO_{12}$ and C_S isomer of $[H_3Ru_4CO_{12}]^-$) when compared to the LNO-CCSD(T) benchmark. The sensitivity of the energies of different isomers of tetranuclear metal carbonyl clusters has been noted by Ding *et al.* [100]. The discrep-

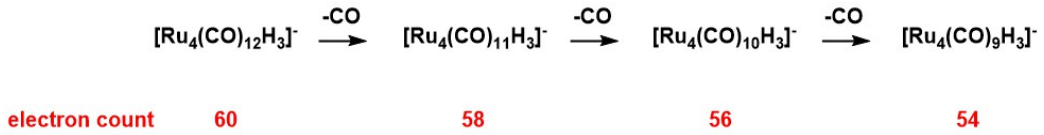


Figure 3.5: Electron counting in the clusters $[H_3Ru_4(CO)_n]^-$, n = 9-12.

ancy between the DFT and LNO-CCSD(T) benchmark is less marked for the hybrid functionals in the lower part of the Table (PBE0-D4, B3LYP-D4), so it appears that the introduction of Hartree-Fock exchange goes some way to reducing the apparent over-stabilisation of bridging modes by common GGA or meta-GGA functionals. It is also striking that the discrepancies between DFT and LNO-CCSD(T) are much less marked in the deprotonated form, $[H_3Ru_4CO_{12}]^-$, compared to its neutral counterpart, perhaps a result of stronger Ru $\rightarrow$ CO π^* backbonding in the former, which reduces electron-electron repulsion.

3.4 CO-loss and structures of $[H_3Ru_4(CO)_n]^-$

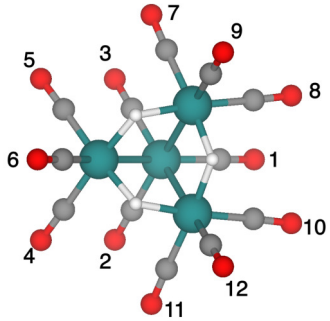
The primary motivation for this study is the reactivity of the CO-deficient clusters generated from $[H_3Ru_4(CO)_{12}]^-$ with an alkene, hexene. The EDESI-MS results summarised in the introduction confirm a complex pattern of reactivity, with CO loss being followed by binding of an alkene, followed, potentially, by C-H bond activation and/or loss of H_2. We defer the discussion of the reactions with 1-hexene to section 3.5.1, and focus first on the electronic and structural consequences of loss of CO. The parent cluster $[H_3Ru_4(CO)_{12}]^-$ and its protonated analogue $Ru_4H_4(CO)_{12}$ both have a total valence electron count of 60 and, as emphasised by Hoffmann and Mingos, the bonding can be understood in terms of four 15-electron $Ru(CO)_3^-$ fragments, each bonded to three others through localised Ru-Ru single bonds. The successive loss of CO depletes the valence electron count by two at each step, leading, after the loss of three ligands, to a rather electron-deficient 54-electron cluster $[H_3Ru_4(CO)_9]^-$ (the intense peak centred on coordinate (659,43) in Figure 3.3).

In contrast to their electronically saturated precursors, potential energy surfaces for unsaturated carbonyl clusters are typically rather flat, with numerous stationary points of similar energy that presents a substantial challenge to DFT. The geometries of unsaturated carbonyl dimers and trimers has has discussed extensively by King and co-workers, [101–104] and we have explored similar issues in a study of $Ru_3(CO)_x$, x = 6-12.[105] In all cases, a number of closely-spaced minima emerge, but we can identify a general pattern that multiple metal-metal bonds form to compensate for the

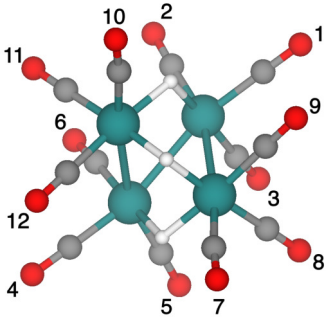
electronic unsaturation. Extend these approach to the rather more complex problem of the $[H_3Ru_4(CO)_x]^-$ clusters of interest in this paper represents a significant challenge, simply because of the dimensionsality of the potential energy surfaces $(3N-6)$. In recent years a number of algorithms have emerged to facilitate the identification of local and global minima, including the 'particle-swarm optimisation (PSO)' approach that we and others have applied to problems in cluster chemistry. The PSO approach, however, involves the generation of a very large number (typically thousands) of candidate structures which are pre-optimised using relatively small basis sets before performing higher-level calculations on the sub-set of most stable structures. It is not, however, obvious that such an approach is readily extendable to systems such as $[H_3Ru_4(CO)_x]^-$, simply because of time required to perform such a large number of optimisations, alongside concerns about the suitability of small basis sets in describing multiple metal-metal bonds. In this work, we therefore adopt an alternative strategy for potential-energy searching based on the systematic the removal of CO ligands from the various stable isomers of the saturated parent compound, $[H_3Ru_4(CO)_{12}]^-$. So for $[H_3Ru_4(CO)_{11}]^-$, for example, we take the optimised structures of the three most stable isomers of the parent dodeca-carbonyl $[H_3Ru_4(CO)_{12}]^-$ (Figure 3.4) and removed each of the 12 CO ligands to generate initial guess structures, from which a geometry optimisation was initialised. For $[H_3Ru_4(CO)_{10}]^-$, we remove each unique pair of ligands and for $[H_3Ru_4(CO)_9]^-$, we remove each unique triad.

3.4.1 Potential Energy Surface for $[H_3Ru_4(CO)_{11}]^-$

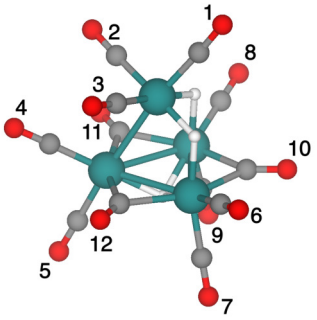
The relative energies of all local minima for $[H_3Ru_4(CO)_{11}]^-$, at the DFT and LNO-CCSD(T) levels (the latter at the r^{2}SCAN-D4-optimised geometries), are summarised in figure 3.6. The labelling of the isomer (1, 2, 3 *etc.*) indicates the index of carbonyl ligand removed from the parent structure as indicated in Figure 3.4. We make maximum use of symmetry by identifying that there are only three symmetry-distinct CO ligands in the C_{3v}-symmetric parent structure of $[Ru_4H_3(CO)_{12}]^-$ ([1,2,3], [4,5,7,8,10,11] and [6,9,12]), and therefore only three unique structures that can be obtained by removing a single ligand. In C_2 and C_s symmetry, there are 6 and 8 symmetry-distinct CO ligands, respectively, and hence 6 and 8 entries in the corresponding tables. Whilst this wealth of information is digestible in tabular form for the case of loss of a single CO, it becomes increasingly cumbersome for loss of two or three ligands, when the number of possible isomers becomes much larger. We have therefore sought a graphical way to represent the data in the form of heat maps (Figure 3.6,



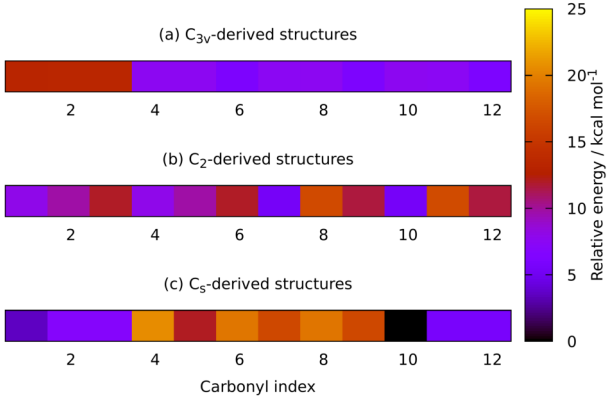
(a) Atom numbering in C_{3v} parent structure



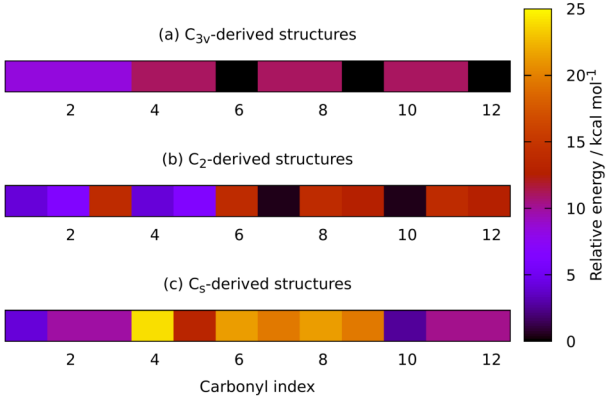
(b) Atom numbering in C_2 parent structure



(c) Atom numbering in C_s parent structure



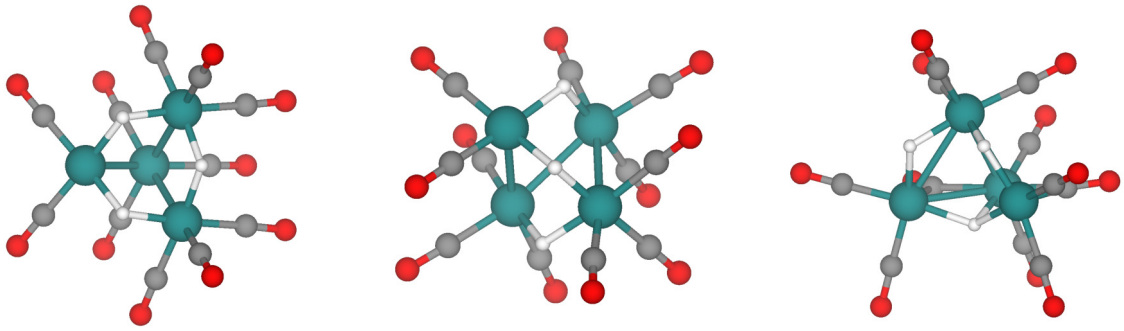
(d) Heat map for CO loss: BP86†



(e) Heat map for CO loss: LNO-CCSD(T)

Figure 3.6: Relative energies of $[H_3Ru_4(CO)_{11}]^-$ calculated within BP86†, r^{2}SCAN-D4 and LNO-CCSD(T). The index of the isomer indicates the index of the carbonyl ligand that was removed from the parent compound in figure 3.4

	C_{3v}			C_2						C_s							
	1	4	**6**	1	2	3	**7**	8	9	1	2	4	5	6	7	**10**	11
BP86†	7.8	1.5	**0**	1.9	3.7	5.9	**-0.6**	10.6	5.5	-2.6	0.8	14.5	6.0	13.3	10.4	**-6.0**	-0.4
r^{2}SCAN-D4	8.8	3.7	**0**	2.9	5.5	8.4	**0.5**	10.2	7.2	0.1	3.0	18.1	9.4	15.8	13.6	**-2.6**	4.0
LNO-CCSD(T)	8.3	11.0	**0**	4.0	6.4	13.8	**0.4**	13.9	12.5	4.1	9.8	23.9	13.1	21.3	19.5	**2.6**	10.2



(a) C_{3v}-6 (b) C_2-7 (c) C_s-10

Figure 3.7: Geometries of low-energy isomers of $[H_3Ru_4(CO)_{11}]^-$

the value of which will become clear in subsequent sections). Figure 3.6d summarises the relative energies of the 17 isomers identified in the table of Figure 3.6 at the BP86 level, from which the most stable structure (C_s-10) is immediately identifiable as the black square indicating zero energy, with $C_{3v}-6$ and C_2-7 only marginally higher in energy. Optimised structures of these three isomers of $[Ru_4H_3(CO)_{11}]^-$ are shown in Figure 3.7. The symmetries of the parent clusters are also evident in the heat-map: C_{3v} $1, 2, 3$ all have the same colour, as do C_{3v} $4, 5, 7, 8, 10, 11$ and C_{3v} $6, 9, 12$.

The second row in the table in Figure 3.6 summarises the energies computed with the r^{2}SCAN-D4, and shows broadly similar trends: the most stable structures are independent of functional choice. The final row corresponds to single-point energies computed at the LNO-CCSD(T) level using the geometries from the r^{2}SCAN-D4 optimisations. These energies are also illustrated in the heat-map on the right hand side of Figure 3.6. Here, we see some striking differences compared to the DFT results: specifically, structures with μ_2 bridging CO ligands (including C_s-10) are relatively destabilised compared to those with only terminal ligands by ~ 8 kcal mol^{-1}, precisely as we saw in the survey of the saturated 60-electron clusters. The result is a change in the identity of the global minimum, which is $C_{3v}-6$ at the LNO-CCSD(T) level. In light of the striking discrepancy between the BP86/r^{2}SCAN-D4 and LNO-CCSD(T) results, we have surveyed the relative energies of the three most stable structures, C_s-10, C_2-7 and $C_{3v}-6$ with a wider range of functionals (Table 3.2). The over-stabilisation of the C_s-10 isomer appears to be a very general feature of density functional theory, and even with the B3LYP-D4 functional, which does predict C_s-10 to be less stable than $C_{3v}-6$, the difference is still too small. Within the series of clusters with only terminal carbonyls, however, the relative energies of different isomers appears to be relatively independent of functional, and all values

Table 3.2: Relative energies of different isomers of $[H_3Ru_4(CO)_{11}]^-$

	$[H_3Ru_4(CO)_{11}]^-$		
	C_{3v}-6	C_2-7	C_s-10
BP86-D4	0	2.5	-3.6
PBE-D4	0	1.5	-4.3
r^{2}SCAN-D4	0	0.5	-3.1
M06-L	0	0.3	0.1
PBE0-D4	0	1.6	-2.2
B3LYP-D4	0	3.3	0.7
LNO-CCSD(T)	0	0.4	2.6

are broadly consistent with the LNO-CCSD(T) results. As a result, we persist with the BP86 functional for our initial survey of the potential energy surfaces of more unsaturated clusters.

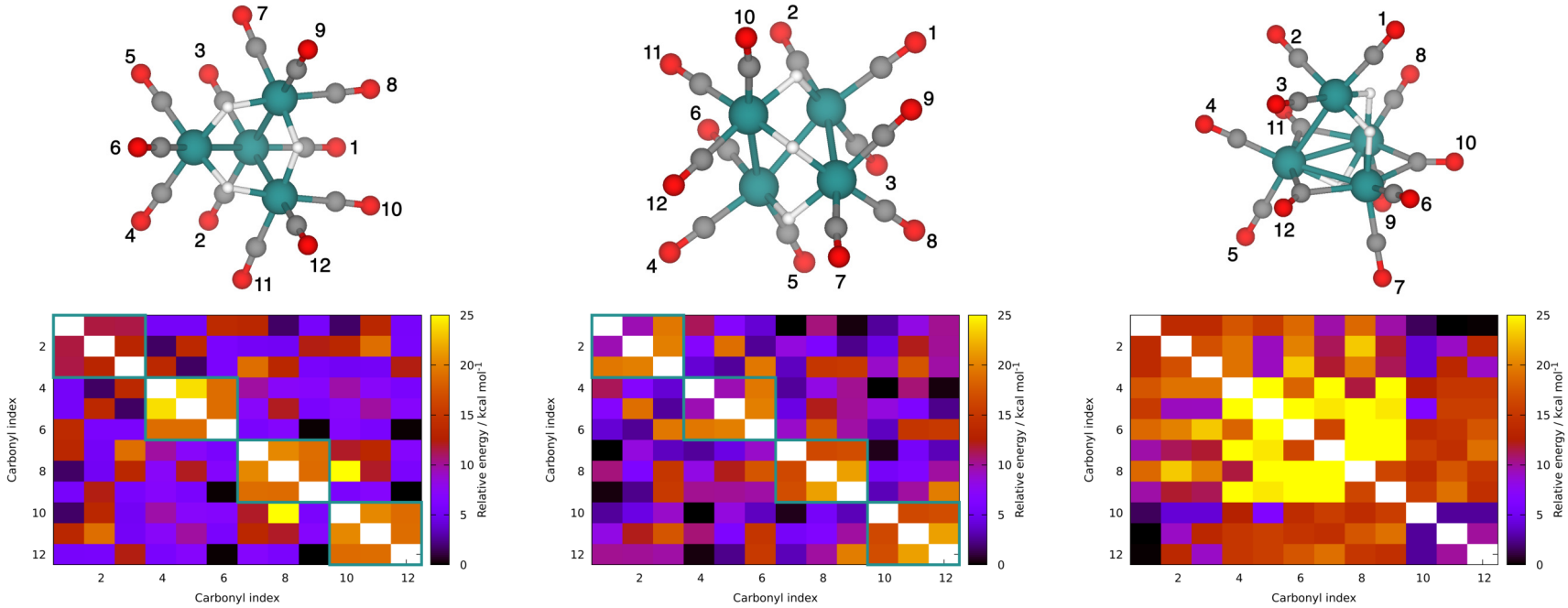
The heat maps for $[H_3Ru_4(CO)_{11}]^-$ have identified three near-degenerate isomers, C_s-10, C_2-7 and $C_{3v}-6$, that differ in the positions of the hydride and CO ligands around an approximately tetrahedral Ru_4 core. There are, in addition, a number of alternative structures that are only marginally less stable, and in such circumstances it is not possible to draw conclusions from any individual isomer, given that a dynamic equilibrium will likely exist between several different atomic arrangements. Nevertheless, we can identify structural patterns that link the lowest-lying isomer, irrespective of their symmetries or of the arrangement of the three hydride ligands around the tetrahedral core. In the two all-terminal isomers identified in Figure 3.7, $C_{3v}-6$ and C_2-7, we can identify a single very short, unbridged, Ru-Ru bond (2.67 Å and 2.69 Å, respectively) between the unsaturated Ru centre (*i.e.* the one carrying only two CO ligands) and one of the others. The remaining bond lengths remain close to the typical values of 2.80 Å and 2.94 Å for unbridged and bridged Ru-Ru single bonds observed in the saturated clusters. The contraction of a single Ru-Ru bond is characteristic of the development of multiple Ru-Ru bonding of the type noted in the CO-deficient Ru_3 analogues,[105] and it is largely independent of the precise position of the ligands around the Ru_4 core.

3.4.2 Potential Energy Surface for $[H_3Ru_4(CO)_{10}]^-$

For $[Ru_4H_3(CO)_{10}]^-$, generated by the removal of two CO ligands from the saturated 60-electron precursor, the potential energy surface is more complex again. In order to achieve an unbiased survey, we have again adopted a systematic approach by

initiating geometry optimisations starting from structures where removing all possible combinations of two CO ligands are removed from the parent $[Ru_4H_3(CO)_{12}]^-$ structures (C_{3v}, C_2 and C_s in Figure 3.8). This process generates $\binom{12}{2} = 66$ possible starting structures, the energies of which (at the BP86† level) are shown in the heat maps in Figure 3.8. In these Figures, the rows and columns denote the numbers of the carbonyl ligands that have been removed from the parent structure, while the color indicates the energy (values in all three maps are given relative to the most stable isomer generated from the C_{3v}-symmetric parent). Thus, for example, the black color in the 6th column, 9th row (the 6,9 element) of the first map indicates that removing the CO ligands labelled 6 and 9 in the parent structure leads to a very stable structure. The diagonal squares are empty (white) because it is obviously impossible to remove the same ligand twice. The 3×3 blocks along the lead diagonals (outlined in green) group together CO ligands on the same Ru centre: CO ligands 1, 2 and 3 are on the same Ru centre, 4, 5 and 6 are on another *etc.*. The fact that all off-diagonal elements within these diagonal blocks are relatively high in energy (red/orange colours) is an immediate indication that removing two CO ligands from the same Ru centre is energetically unfavourable compared to removing two COs from different centres. The maps are necessarily symmetric about the lead diagonal because the i,j and j,i elements are identical. Additional symmetries reflect the intrinsic symmetry of the parent clusters: for the C_{3v}-symmetric parent, ligands 6, 9 and 12 are related by 3-fold rotational symmetry, and so the (6,9), (9,12) and (6,12) elements are identical by symmetry. Likewise, the (1,7) and (4,10) elements of the C_2 matrix are symmetry equivalent, as are (1,9) and (4,12). The structure of the C_s heat-map is somewhat different from the other two because the structure has both bridging and terminal CO's, and so we cannot order the CO's easily according to which Ru they are bound to. The bridging COs are labelled 10, 11 and 12, and the blue/purple colours in the bottom right 3×3 diagonal block are an immediate indication that these bridging CO's are removed more readily than those in terminal positions.

The information in the heat maps can be condensed down to a summary of the energies of the nine most stable structures in the table in Figure 3.8: the corresponding geometries are shown in Figure 3.9. In the structures derived from the C_{3v}-symmetric parent, the most stable isomer is C_{3v}-(6,9), where the (6,9) pair of carbonyl ligands is removed (note that the (6,12) and (9,12) isomers are symmetry equivalent). The energies of the other eight structures lie within $\sim$15 kcal mol^{-1} of this most stable isomer. In a number of cases, the removal of two ligands causes one or more of the



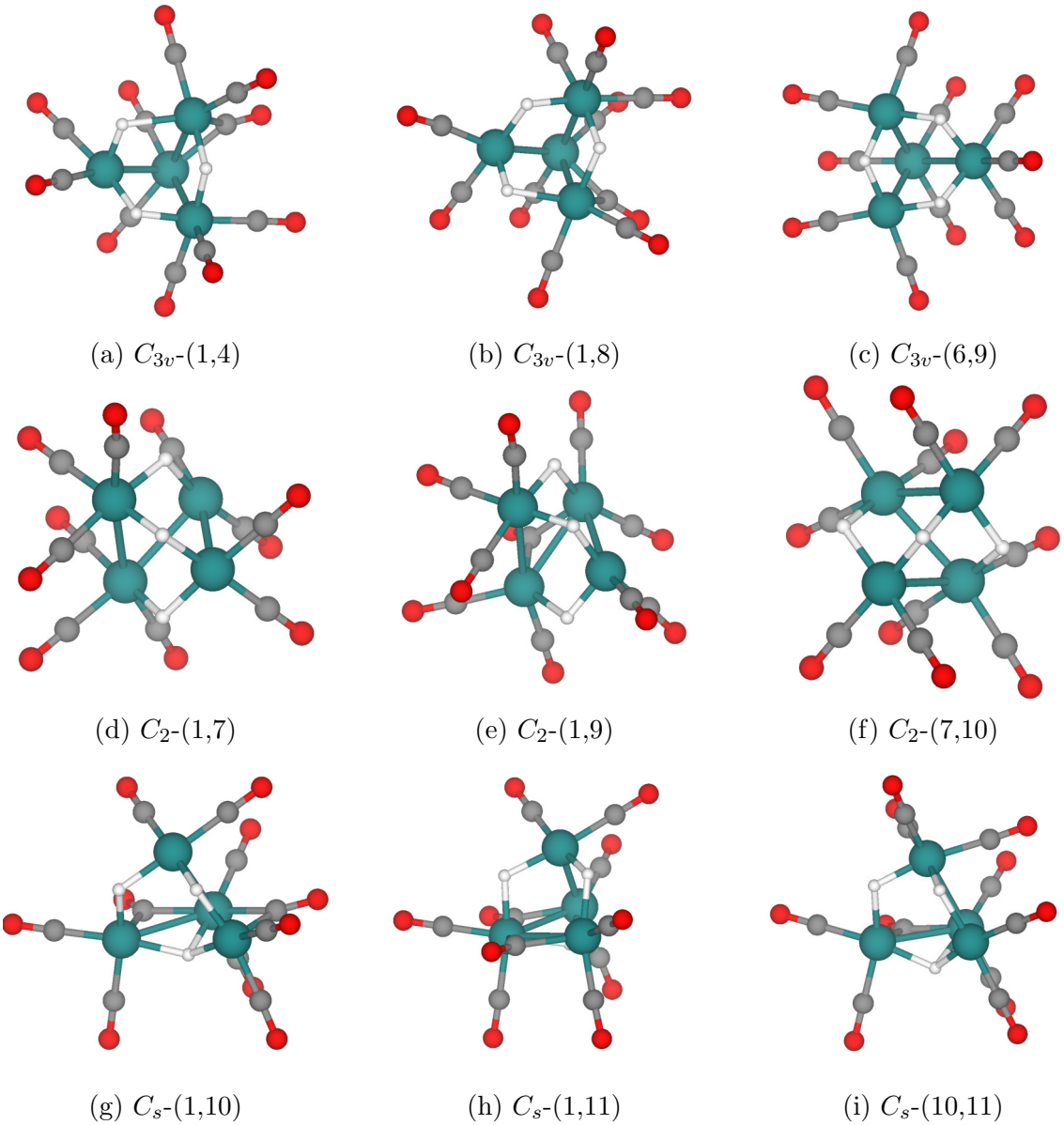
(a) Heat map for the C_{3v} parent (BP86)

(b) Heat map for the C_2 parent (BP86)

(c) Heat map for the C_s parent (BP86)

Figure 3.8: Relative energies of $[H_3Ru_4(CO)_{10}]^-$ calculated within BP86†, r^{2}SCAN-D4 and LNO-CCSD(T). Heat maps show the relative energies of all isomers of $[H_3Ru_4(CO)_{10}]^-$ derived by removing different pairs of carbonyl ligands removed from the (a) C_{3v}, (b) C_2 and (c) C_s parent structures after the geometry optimisations with BP86†. The energies are reported relative to the isomer with the lowest energy in the same matrix. The indices for carbonyl ligands for each parent structure are shown in Figure 3.4. The relative energies of the nine most stable isomers in the heat map are summarised in tabular form below, for the BP86† and r^{2}SCAN-D4 functionals, and also with LNO-CCSD(T).

	C_{3v}			C_2			C_s		
	(1,4)	(1,8)	**(6,9)**	**(1,7)**	(1,9)	**(7,10)**	(1,10)	(1,11)	(10,11)
BP86†	5.3	1.6	**0**	**0.7**	0.9	**1.1**	-0.4	-1.9	0.7
r^{2}SCAN-D4	8.7	5.3	**0**	**1.6**	4.5	**2.1**	2.2	1.5	3.8
LNO-CCSD(T)	14.1	13.5	**0**	**1.0**	11.7	**1.8**	7.9	8.5	7.2



(a) C_{3v}-(1,4) (b) C_{3v}-(1,8) (c) C_{3v}-(6,9)

(d) C_2-(1,7) (e) C_2-(1,9) (f) C_2-(7,10)

(g) C_s-(1,10) (h) C_s-(1,11) (i) C_s-(10,11)

Figure 3.9: Structures of the nine lowest energy isomers of $[H_3Ru_4(CO)_{10}]^-$ identified in the heat maps in the table in Figure 3.8

remaining CO ligands to shift from a terminal to a bridging position when the structure is allowed to relax. C_{3v}-(1,4), C_{3v}-(1,8) and C_2-(1,9) are cases in point, and we see the now familiar destabilisation of these structures by the LNO-CCSD(T) method compared to DFT. At the LNO-CCSD(T) level, only three structures remain at low energy: C_{3v}-(6,9), C_2-(1,7) and C_2-(7,10) (and their symmetry-related analogues). The optimised structural parameters for these three most stable isomers, also shown in Figure 3.9, show a continuation of the contraction of the unbridged Ru-Ru bonds identified for $[H_3Ru_4(CO)_{11}]^-$. In the C_{3v}-(6,9) isomer, we now find two short unbridged Ru-Ru bond lengths (2.67 Å) between the apical $Ru(CO)_3$ fragment and each of the the unsaturated $Ru(CO)_2$ centres, while the other unbridged Ru-Ru bond remain at 2.82 Å, a value that we now recognise as being typical of an unbridged single bond. The two unsaturated centres are linked *via* a hydride-bridged bond, and this is somewhat longer at 2.87 Å, only marginally contracted relative to equivalent bonds in the more saturated clusters. In isomer C_2-(1,7), the two unsaturated Ru centres are linked *via* an Ru-Ru bond, and this is now extremely short, at 2.62 Å. The remaining Ru-Ru bond lengths remain close to the typical single-bond values of 2.8 Å and 2.93 Å for unbridged and hydride-bridged, respectively. Finally, in C_2-(7,10), we identify a similar pattern to that in C_{3v}-(6,9), where we have two moderately short Ru-Ru bond lengths of 2.70 Å, with the remaining bonds relatively unchanged from values typical of the saturated counterparts.

The emerging structural picture based on the $[H_3Ru_4(CO)_{12}]^- \rightarrow [H_3Ru_4(CO)_{11}]^- \rightarrow [H_3Ru_4(CO)_{10}]^-$ series is summarised in Scheme 3.10. In the saturated parent clusters (top row), the three short, unbridged, bonds are identified by the bold lines. The loss of a CO ligand from an individual Ru centre results in an enhancement of the Ru-Ru bond trans to the unsaturated site, which was previously weakened by the σ-donation from carbonyl ligands due to trans influence. These enhanced bonds are identified as red arrows, in Scheme 3.10. The trans influence of a carbonyl ligand causes the frontier orbital of Ru centre to hybridise more with the σ orbital on carbonyl, thereby reducing the bonding character of trans ligand/fragment, in this case, the Ru-Ru bond. The evolution from $C_{3v} \rightarrow C_{3v}$-6 $\rightarrow C_{3v}$-(6,9) therefore corresponds to the formation of one and two enhanced bonds from the apical $Ru(CO)_3$ fragment to the unsaturated centres in the basal plane. Equally, the evolution across the C_2 family, $C_2 \rightarrow C_2$-7 $\rightarrow C_2$-(1,7) corresponds to the evolution of a chain of Ru-Ru bonds with enhanced σ character. Critically, although the potential energy surfaces of the unsaturated clusters are clearly very flat, with multiple minima of similar

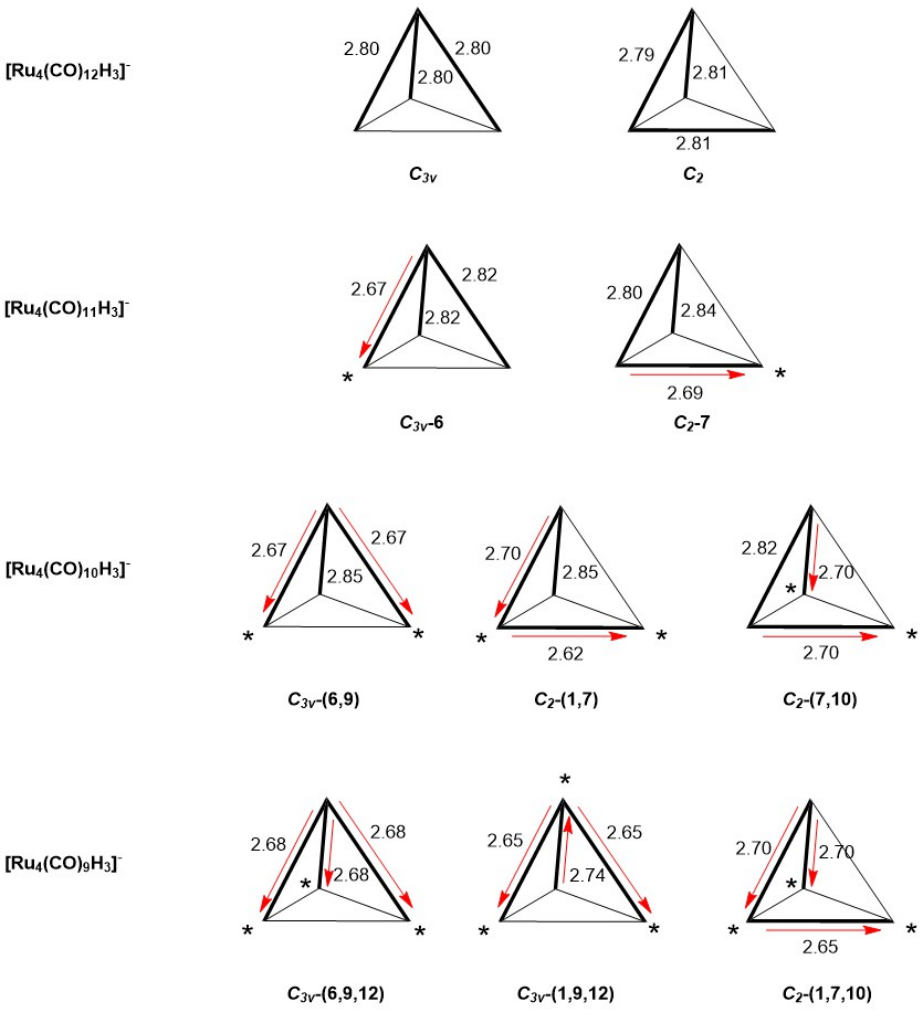
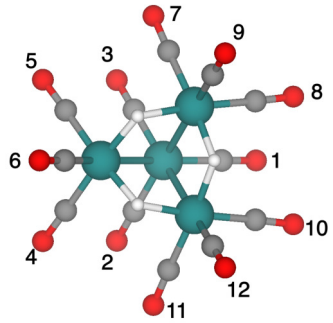


Figure 3.10: Evolution of the Ru-Ru bonds with progressive loss of CO. Unbridged Ru-Ru bonds are shown in bold, enhanced bonds with a red arrow and unsaturated Ru centres are identified with an asterisk.

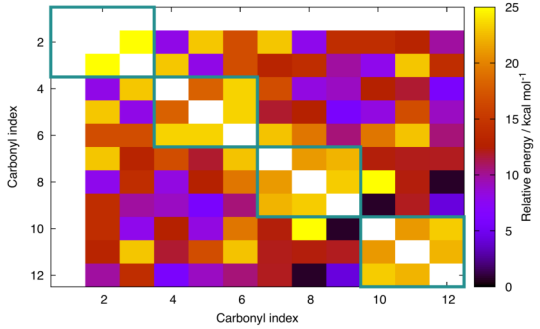
energies, we can identify patterns in the Ru-Ru bonding that are common to all low-lying isomers of a given composition, and are not dependent on the precise positions of the CO ligands.

3.4.3 Potential Energy Surface for $[H_3Ru_4(CO)_9]^-$

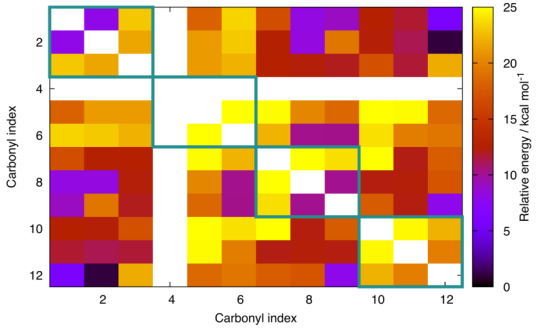
Moving on to the final cluster of interest here, $[H_3Ru_4(CO)_9]^-$, we can understand the various features that emerge as a natural extension to the trends identified in the first three rows of Scheme 3.10. Our systematic approach of removing each unique triad of CO's would generate $\binom{12}{3} = 220$ possibilities (without the consideration of symmetry) for each of the three parent structures. However, since structures of the less unsaturated clusters derived from the C_s isomer, with bridged carbonyls, have always proved to be relatively unstable at the LNO-CCSD(T) level, we do not consider this structure further here, leaving only the C_{3v} and C_2 families to consider. The heat maps for the loss of three CO ligands are intrinsically 4-dimensional (the four dimensions being the three ligand indices, i,j,k and the relative energy), which presents an immediate problem in presenting the data in Euclidean space. For the C_{3v}-symmetric parent, however, we note again that there are only three symmetry-independent types of CO ligand ([1,2,3], [4,5,7,8,10,11] and [6,9,12]). Thus, three separate slices through the heat map are sufficient to capture all of the relevant data. In Figure 3.11b, ligand 1 is removed, along with all possible other pairs. In Figure 3.11c, ligand 4 and two others are removed, while in Figure 3.11d, ligand 6 and two others are removed. The resultant slices can be interpreted in the same way as the heat maps for $[Ru_4H_3(CO)_{10}]^-$. In addition to the white squares along the diagonals, the ([1,4,6],j) columns and (j,[1,4,6]) rows are also blank because we cannot remove the same ligand twice. The black squares pick out the most stable structures, notably the (1,9,10) and (1,8,12) triads in Figure 3.11b. The high symmetry of the cluster means that the (2,4,12) element in Figure 3.11c and the (3,6,7) element in Figure 3.11d are in fact identical to (1,9,10) in Figure 3.11b, and (3,6,7) in Figure 3.11d is symmetry-equivalent to (1,8,12) in Figure 3.11b. The only new stable (black) structure that emerges from Figures 3.11c and 3.11d is therefore the (6,9,12) element where three symmetry-equivalent CO ligands around the μ_2-H capped face are removed, generating a C_{3v}-symmetric structure. By using symmetry in this way, we can reduce the 220 candidate structures that emanate from the C_{3v}-symmetric parent down to three distinct relatively stable isomers, (1,9,10), (1,8,12) and (6,9,12), the energies of which are summarised in the first three columns of the table in Figure 3.11. Repeating the process starting from the C_2-symmetric parent generates a further 220



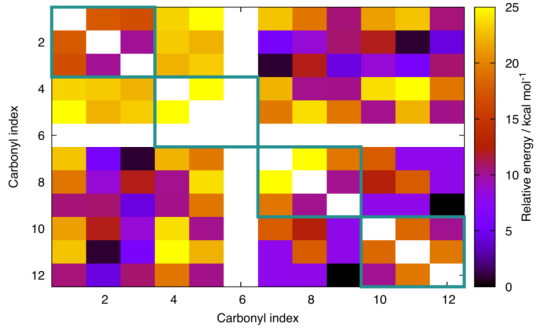
(a) Atom numbering in the C_{3v} parent



(b) C_{3v}-(1,j,k) heat map



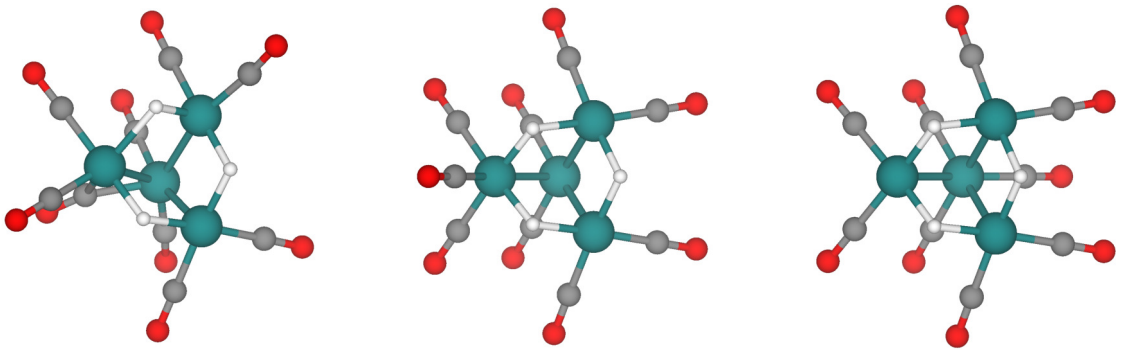
(c) C_{3v}-(4,j,k) heat map



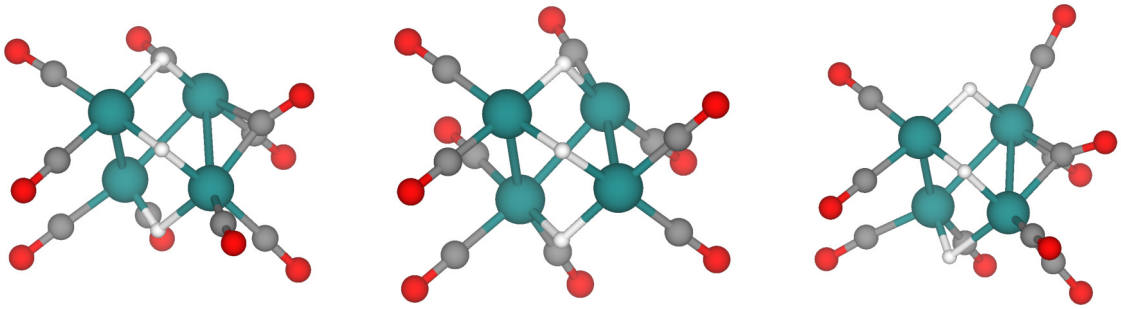
(d) C_{3v}-(6,j,k) heat map

Figure 3.11: Heat maps of the relative energies of all isomers of $[H_3Ru_4(CO)_9]^-$ derived by removing different pairs of carbonyl ligands removed from the C_{3v} parent structures after the geometry optimisations with BP86†. The energies are reported relative to the isomer with the lowest energy in the same matrix. The relative energies of the six most stable isomers in the heat map are summarised in tabular form below, for the BP86† and r^{2}SCAN-D4 functionals, and also with LNO-CCSD(T).

	C_{3v}			C_2		
	(1,8,12)	**(1,9,12)**	**(6,9,12)**	**(1,7,10)**	(1,6,10)	(2,6,10)
BP86†	0.6	**4.2**	**0**	**2.6**	2.5	2.9
r^{2}SCAN-D4	3.2	**4.1**	**0**	**2.8**	4.3	3.5
LNO-CCSD(T)	11.3	**3.3**	**0**	**2.9**	9.2	11.6



(a) C_{3v}-(1,8,12) (b) C_{3v}-(1,9,12) (c) C_{3v}-(6,9,12)



(d) C_2-(1,6,10) (e) C_2-(1,7,10) (f) C_2-(2,6,10)

Figure 3.12: Geometries of low-energy isomers of $[H_3Ru_4(CO)_9]^-$

initial structures. The lower symmetry (compared to C_{3v}) means that six different slices would be needed to capture all of the information in the heat-map. Rather than presenting all of this data graphically, we simply quote the relative energies of the three most stable isomers that emerge from this analysis, which are the (1,6,10), (1,7,10) and (2,6,10) triads, in the final three columns of the table in Figure 3.11 .

This analysis leads us to three possible candidates for the global minimum (at the LNO-CCSD(T) level), C_{3v}-(6,9,12), C_{3v}-(1,9,12) and C_2-(1,7,10), the most stable being the first of these at all levels of theory tested. The Ru-Ru bond lengths for these three clusters are collected in Figure 3.12, and also summarised in schematic form in the final row of Scheme 3.10. The removal of the third CO ligand continues the previously established pattern of Ru-Ru enhanced bond formation as a compensatory mechanism for the electron deficiency. Loss of a third CO to form C_{3v}-(6,9,12) completes a set of three equivalent enhanced Ru-Ru bonds from the apical $Ru(CO)_3$ unit to the basal plane, while in C_2-(1,7,10) we form a zig-zag chain of three enhanced Ru-Ru bonds. An average Ru-Ru bond order of 1.5 is therefore common to all of the low-lying isomers of $[H_3Ru_4(CO)_9]^-$, irrespective of the precise location of the CO ligands.

Based on the data summarised in Scheme 3.10, we can formulate a set of principles that govern the structural consequences of progressive CO loss:

1. It is more favourable to remove CO ligands from different Ru centres than to remove multiple ligands from the same metal centre.

2. The carbonyl ligands that are trans to unbridged M-M bonds are more likely to be removed than those trans to bonds bridged with hydride.

3. The CO ligands are removed such that the vacant sites lie *trans* to a linear CO-Ru-Ru unit.

4. The most labile CO ligands are those that are *cis* to one or more hydride ligand.

The first of these rules is largely self-explanatory: the successive loss of CO at a single centre localises the electron deficiency on a single centre, and this is less favourable than distributing the load over multiple centres. The second, the observation that it is the unbridged bonds that adapt, reflects the fact that direct overlap of Ru-based orbitals is more important here than in the bridged bonds. The preference for CO-Ru-Ru-X units (where X is a CO vacancy) over X-Ru-Ru-X is most likely due to the dative nature of the bonds formed on CO loss: these can be viewed as the donation of a pair of electrons from the unsaturated centre to the unsaturated member of the pair (the Ru atom carrying the vacancy). The ability of the saturated centre to push electron density towards the unsaturated one is enhanced by the presence of a ligand *trans* to the Ru-Ru bond. The final note regarding the presence of *cis* hydrides, is illustrated neatly by the structures of the two most stable isomers of $[H_3Ru_4(CO)_{11}]^-$, both of which have the vacancy *cis* to <u>two</u> hydride ligands. The strong trans influence of hydrides means that there are rarely found *trans* to each other or *trans* to metal-metal bonds, but their strong σ-donors character clearly helps to buffer the electron deficiency at the unsaturated centres.

To further analyse the interactions between the carbonyl ligands on different sites and the cluster, fragment analysis was performed and the decomposition of interaction energies for the C_2 parental anion is given in Table 3.3. As expected, carbonyl ligands trans to bridged hydride (3, 8 and 9) are more strongly interacting with the cluster, compared to those trans to unbridged Ru-Ru bonds (1, 2, and 7). At a glace, the source of this enhanced interaction can be found in the difference in the steric interaction energies, in particular from Pauli repulsion (but there is an opposite trend in the electrostatic energies). The reduced Pauli repulsion between the carbonyl ligand *trans* to the bridged hydride and the ruthenium atom may be traced back to the

Table 3.3: Fragment analysis of the C_2 parent anion $[H_3Ru_4(CO)_{12}]^-$ within r^{2}SCAN-D4. The molecular models are partitioned into a $[H_3Ru_4(CO)_{11}]^-$ cluster and a selected carbonyl ligand in each column. The orbital interaction energy is decomposed into contributions from each significant ETS-NOCV pair. The values of energies are given in kcal mol^{-1}.

	C_2-1	C_2-2	C_2-3	C_2-7	C_2-8	C_2-9
Steric interaction:	66.0	63.7	48.1	63.6	48.7	49.2
Pauli repulsion	212.5	208.6	179.2	204.6	177.6	178.0
Electrostatic interaction	-146.6	-144.9	-131.2	-141.0	-129.0	-128.8
Orbital interaction:	-117.6	-117.7	-112.3	-110.2	-109.1	-107.9
Ru $\leftarrow$ CO σ	-44.4	-44.6	-48.0	-42.1	-47.5	-45.8
Ru $\rightarrow$ CO π^* (1)	-30.6	-31.6	-28.4	-28.7	-29.3	-29.0
Ru $\rightarrow$ CO π^* (2)	-28.6	-27.7	-27.8	-26.9	-25.3	-24.7
4th polarisation	-12.7	-12.5	-6.8	-12.6	-6.5	-7.7
Dispersion:	-0.9	-1.0	-1.2	-1.1	-1.5	-1.4
Interaction energy	-52.55	-54.91	-65.40	-47.71	-61.92	-60.11

character of hydride-bridged Ru-Ru bonds; the metal-metal bonding orbital will have more character of hydrogen s orbitals and less character diffuse ruthenium frontier orbitals, thereby reducing the overlap between two occupied orbitals (the metal-metal bonding orbital on the cluster and the lone pair orbital on CO) that need to be orthogonal upon interaction. Counter-intuitively, there is a little change in the total orbital interaction energies between these two types. An energy decomposition analysis in conjunction with natural orbitals for chemical valence (EDA-NOCV) [106, 107] was employed to decompose the orbital interaction energies further into different interactions, such as σ-donation and π-backbonding. While there seemed to be an enhanced interaction in σ-donation of carbonyl ligands trans to bridged hydrides, this effect is cancelled by the reduced fourth interaction caused by the second-order polarisation of the cluster by the carbonyl ligand. There does not seem to be a significant trend in the change of π-backbonding interactions and there is no notable variation in the dispersion energies in all ligands.

3.5 Reactivity of carbonyl deficient clusters

With a deeper understanding of the structures and energetics of the CO-deficient clusters in hand, we now consider the implications for the reactions of these clusters with a representative alkene, 1-hexene. At its simplest level, the alkene can be

regarded as a 2-electron donor which can replace one or more CO, and so alkene binding represents a reversal of the CO loss. However, the evidence from experiment is that significant amounts of the 1-hexene adduct are observed only after three CO ligands have been lost, suggesting that the reaction is more than a simple ligand replacement. Before exploring the DFT-generated potential energy surfaces in detail, it is useful to apply electron-counting principles to anticipate some of the possible outcomes of these reactions (Figure 3.13). The four columns in this scheme represent distinct total valence electron counts: 54, 56, 58 and 60, the latter being the saturated count for $[H_3Ru_4(CO)_{12}]^-$. The three CO-deficient clusters, $[H_3Ru_4(CO)_n]^-$, discussed in the previous sections then lie along the lower left diagonal, with the most electron-deficient species, $[H_3Ru_4(CO)_9]^-$, at the far left. Simple addition of an alkene to each of these electron-deficient species generates an η^2-ene cluster ($[Ru_4(CO)_n(C_6H_{12})H_3]^-$), increasing the electron count by two. If the cluster is still unsaturated at this point, a further oxidative addition of a C-H bond can occur, increasing the electron count by a further two and generating an η^3-allyl species ($[Ru_4(CO)_n(C_6H_{11})H_4]^-$), along with an additional hydride ligand. Repeating this process leads to an η^4-diene and two additional hydrides ($[Ru_4(CO)_n(C_6H_{10})H_5]^-$). This represents the limit of bond activation available to the species considered here, although even more electron-deficient clusters such as $[H_3Ru_4(CO)_8]^-$, could, in principle, induce even more extensive bond activations. Thus, by successive C-H bond activations, the stable 60-electron count can be restored, albeit at the expense of increasing the average oxidation state at the Ru_4 core. In the following sections, we explore the thermodynamics of these oxidative addition steps, the ultimate aim being to provide molecular-level insight into the origins of the peaks identified in the mass spectrometry. In our survey, our starting point will be the lowest-energy isomer of each composition derived from the C_{3v} parent: C_{3v}-6, C_{3v}-(6,9) and C_{3v}-(6,9,12). These three isomers share the important feature that the vacant sites are all directed towards the same side of the cluster, providing a locus for the reaction chemistry. In contrast to the clusters with the same composition derived from the C_2 family, where the vectors linking the vacancies to the centre of the cluster are approximately orthogonal (Figure 3.10). Throughout this section, the energies reported correspond to LNO-CCSD(T) (with r^{2}SCAN-D4 geometries).

3.5.1 Addition of 1-hexene to $[H_3Ru_4(CO)_n]^-$, $n = 11, 10, 9$

A peak corresponding to a cluster of composition $[H_3Ru_4(CO)_{11}(C_6H_{12})]^-$, albeit one with very low intensity, can be identified at coordinate (m/z 800,V = 33) in the

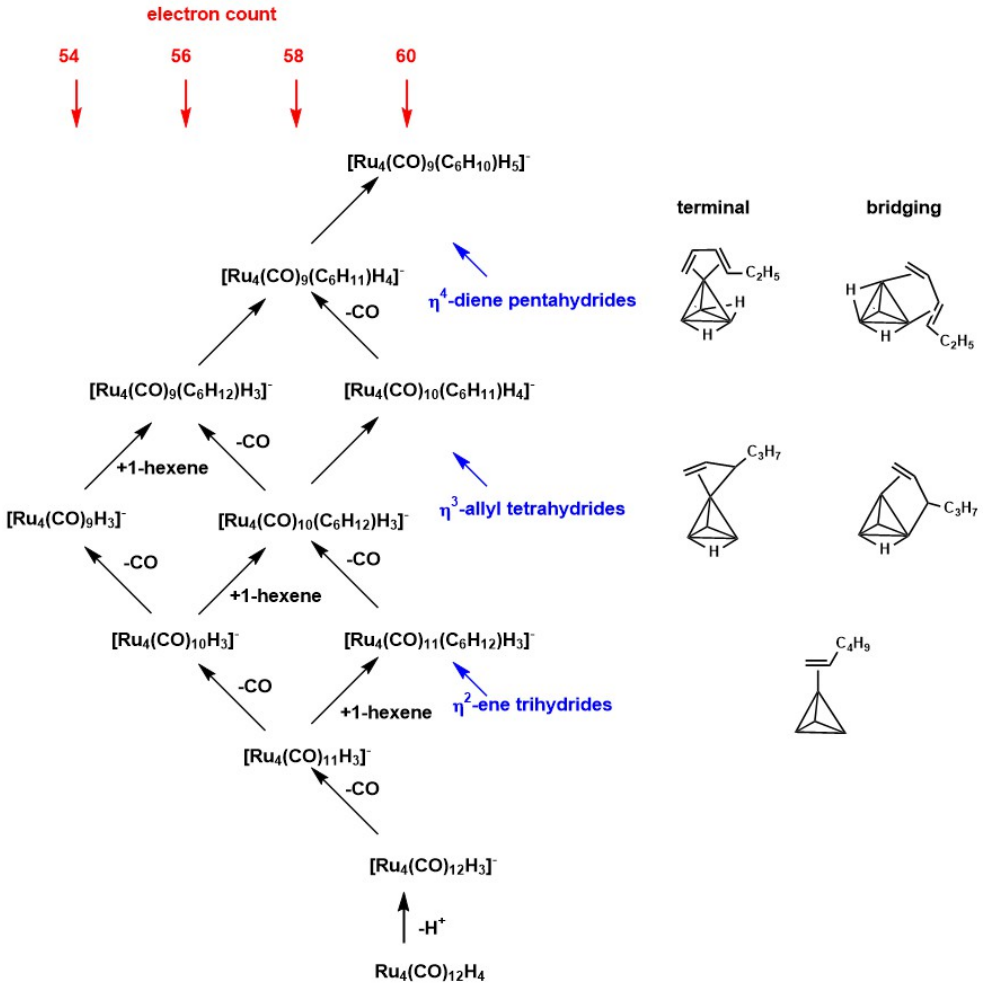
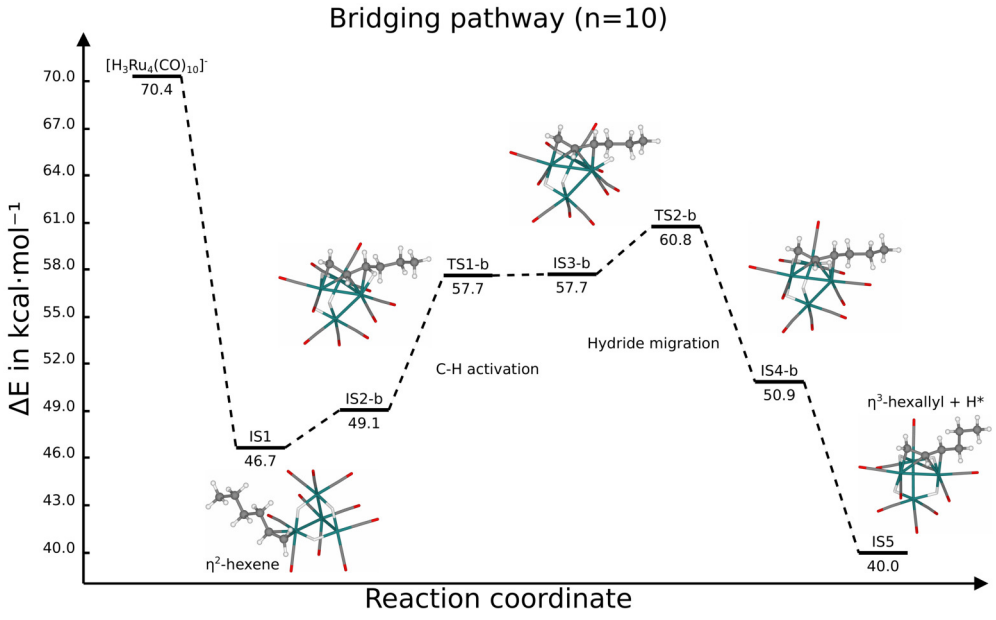


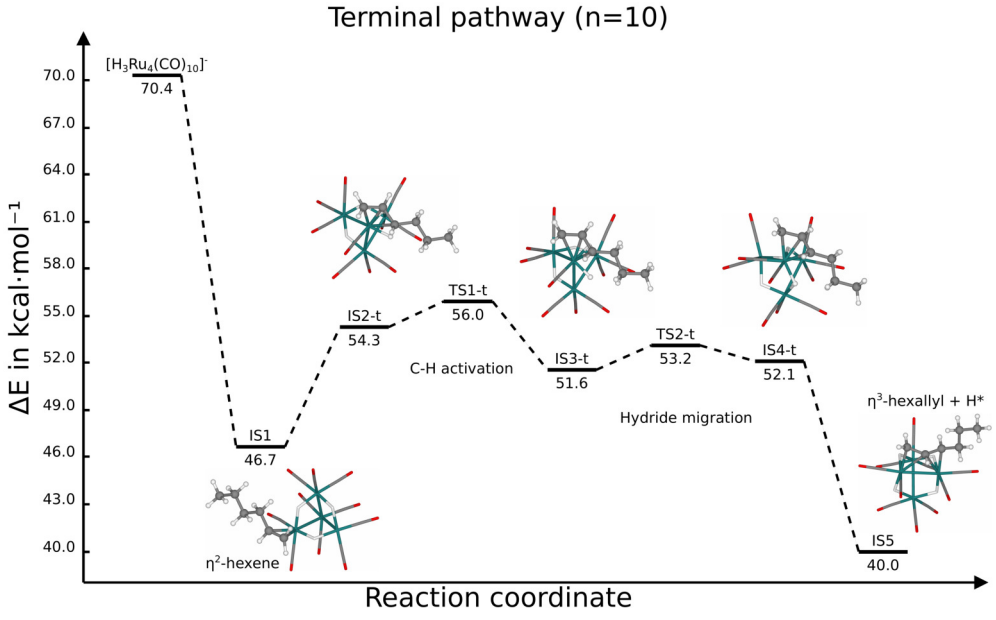
Figure 3.13: Possible formulations of the clusters identified in Figure 3.3. The number in red at the top of each column corresponds to the total valence electron count, with 60 electrons marking a saturated, electron-precise, structure. Each successive C-H bond activation step increases the formal electron count by two, and so compensates for the loss of one CO.

EDESI-MS spectrum shown in Figure 3.3. Starting from the lowest-energy isomer of 58-electron $[H_3Ru_4(CO)_{11}]^-$, C_{3v}-6, the ene group simply coordinates to the vacant site (site 6) in an η^2 mode, restoring the 60-electron count (IS1). The alkene ligand is coordinated to the $Ru_3(\mu-H_3)$ face and the C=C is elongated to 1.40 Å (compared to 1.35 Å for the free 1-hexene at the same level of theory) indicating substantial back-bonding. All six Ru-Ru bonds are restored to values close to those for the $[H_3Ru_4(CO)_{12}]^-$ parent (2.80 Å for the unbridged bonds, 2.93 Å for those with bridges), a clear marker of electronic saturation.

The reaction of 1-hexene with the 56-electron cluster $[H_3Ru_4(CO)_{10}]^-$ presents a greater challenge because the cluster remains coordinatively and electronically unsaturated even after binding of the alkene in an η^2 mode. The relevant peak in the EDESI-MS spectrum corresponding to a composition of $[H_3Ru_4(CO)_{10}(C_6H_{12})]^-$ is located around (m/z 770,$V = 41$): it is much more intense than the feature due to $[H_3Ru_4(CO)_{11}(C_6H_{12})]^-$ discussed above. Possible candidate structures include an η^2-ene (with or without an agostic interaction) and a 60-electron η^3-allyl/hydride. In Figure 3.14a and 3.14b, we identify two possible pathways linking these intermediates. The first step, common to both, is the binding of 1-hexene η^2 to the lowest-energy isomer of $[H_3Ru_4(CO)_{10}]^-$, C_{3v}-(6,9). This restores the local 18-electron count at one of the two previously unsaturated Ru centres, and the Ru_4 core shows only a single short Ru-Ru bond, just as in $[H_3Ru_4(CO)_{11}]^-$ (Figure 3.9). In Figure 3.14a, the alkyl chain can then rotate such that it lies above the vacant site on the adjacent Ru centre, forming first an intermediate (IS2-b) with an agostic bond, showing the elongation of the C-H bond that it characteristic of this class (1.09 Å to 1.18Å). From here, oxidative addition of the C-H to form the allyl-hydride proceeds over a very low transition state, generating the allyl-hydride isomer, IS4-b, where the allyl group bridges two adjacent Ru centres. There is abundant precedent for binding motifs of this type from the work of Cabeza and Hiroshi [108, 109]. The final step is then the migration of the hydride ligands to their thermodynamically most favoured arrangement, similar to that found in the D_{2d}-symmetric isomer of $[H_3Ru_4(CO)_{12}]^-$. The rate-limiting step in this process is the rotation of the alkyl chain, such that it lies over the vacant site on the adjacent Ru. A variant on this scheme, labelled 'terminal pathway', is shown in Figure 3.14b. The initial step of binding of 1-hexene is identical to the bridging pathway, but then a CO ligand on the Ru centre carrying the alkene ligand migrates to fill the vacant site, leaving the unsaturation localised on the Ru centre carrying the alkene ligand. The same pattern and oxidative addition of the C-H bond then leads to an η^3 allyl hydride (IS4-t) where now the allyl group is bound only to a single



(a)



(b)

Figure 3.14: Bridging and terminal pathways for alkene binding and C-H bond activation at $[H_3Ru_4(CO)_{10}]^-$

Ru centre. There appears to be very little to choose between these two pathways in a thermodynamic sense. Whichever pathway is followed, the η^3-allyl hydride, where the allyl ligand bridges two Ru centres, appears to be the most likely candidate for the low-intensity peak observed at (m/z 770,$V = 41$).

Finally, we consider the reaction pathways for 1-hexene with the most electron-deficient of the clusters considered here, the 54-electron species $[H_3Ru_4(CO)_9]^-$. The relevant peak in the EDESI-MS plot for $[H_3Ru_4(CO)_9(C_6H_{12})]^-$ (Figure 3.3) is centred at approximately (m/z 740,$V = 43$), and it is noticeably more intense than either of those for $[H_3Ru_4(CO)_{10}(C_6H_{12})]^-$ or $[H_3Ru_4(CO)_{11}(C_6H_{12})]^-$ discussed previously. The potential for reactivity is now greater again because even after the first C-H activation step leading to the allyl-hydride, the cluster remains at the 58-electron level and hence susceptible to further reactivity. The available pathway is summarised in Figure 3.15, where the initial steps *via* a 56-electron η^2-hexene intermediate (IS1) through to the η^3-allyl-hydride (IS5) are very similar to those reported in Figure 3.14a. From this point, a further moderately exothermic C-H bond activation leads to a 60-electron diene-hydride IS6 ($[Ru_4H_5(CO)_{10}(\mu\text{-}\eta^{2:2}\text{-}C_6H_{10})]^-$), where the diene bridges two of the three previously unsaturated Ru centres, and the H^- ligand is occupying the third vacant site.

Given that the second C-H oxidative addition is only marginally exothermic, it is not obvious that the formation of the diene is, in isolation, responsible for the much greater intensity of the (m/z 740,$V = 43$) peak in Figure 3.3. However, the formation of IS6 introduces a new structural motif, not previously observed in any of the other intermediates: the fifth hydride ligand occupies a terminal position. Loss of a molecule of H_2 then restores the electron deficiency (IS7 is a 58-electron species), which opens up a new, almost barrierless, pathway for further C-H bond activation leading, ultimately, to the η^5-hexadienyl bridged cluster, IS11. The H_2 loss step in Figure 3.15 is endothermic, although this will be offset to some extent by the increase in entropy. In the absence of any knowledge of the effective pressure of H_2 in the reaction, we do not attempt to make any correction for entropy here, but we reiterate that the broadness of the peaks in the EDESI-MS spectrum is fully consistent with the loss of H_2. The loss of H_2 from IS6 would render the reaction irreversible, offering a possible explanation for the much higher intensity of the (m/z 740,$V = 43$) peak relative to those arising from the less unsaturated clusters.

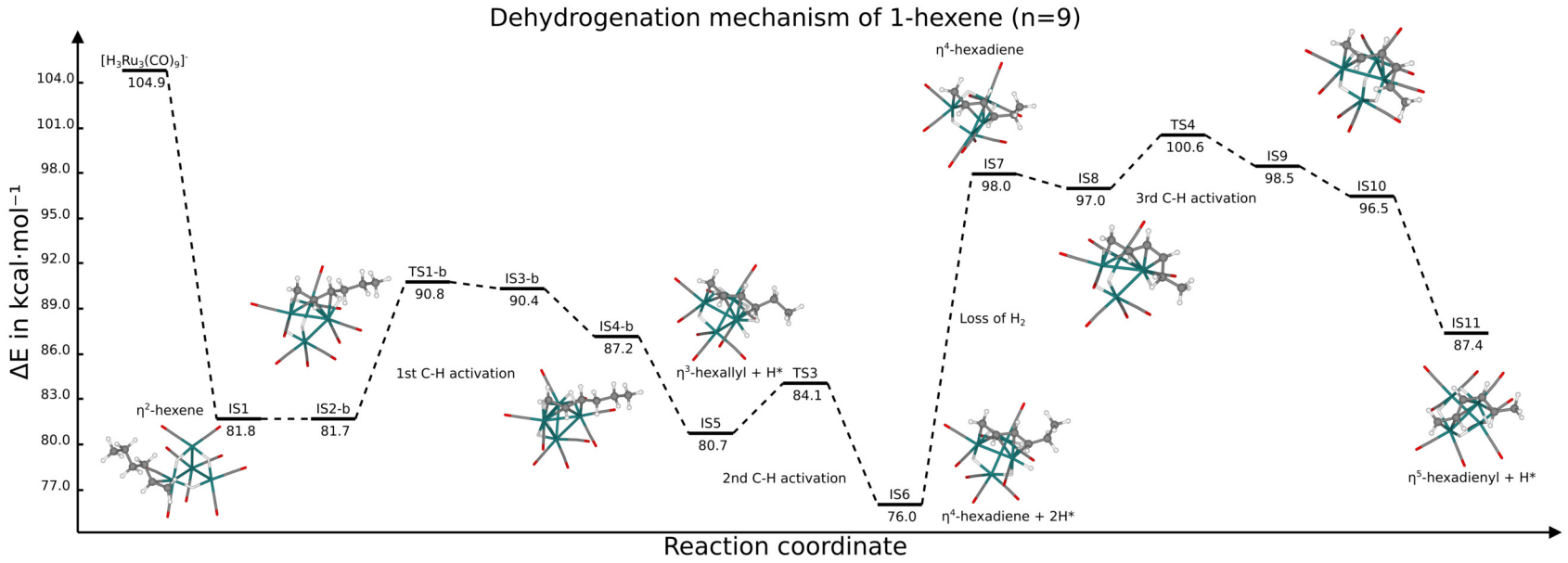


Figure 3.15: An example of possible reaction pathways for 1-hexene dehydrogenation by $[HeRu_4(CO)_9]^-$. Subsequent hydride migration steps after C-H activation are omitted for brevity.

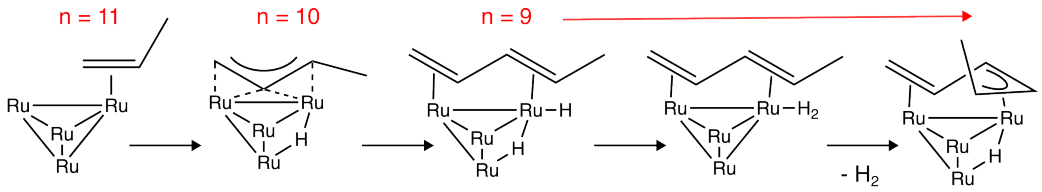


Figure 3.16: A schematic summarising the evolution of an alkane chain and hydride ions along the reaction pathways. For brevity, carbonyl ligands and original hydride ions on the cluster are omitted, and the length of the alkane chain is shortened. The red notes indicate the highest number of carbonyl ligands that can be bonded to the cluster while adopting that bonding mode of the alkane chain.

3.6 Conclusion

In this work we have explored the potential of coordinatively and electronically unsaturated Ru_4 carbonyl hydrides to bind an alkene, 1-hexene, and then activate one or more C-H bond. The calculations are motivated by observations from voltage-dependent mass spectrometry published, which indicate a rich reaction chemistry, particularly for highly unsaturated species. [78, 82, 105] The successive loss of one, two, and three CO ligands from $[H_3Ru_4(CO)_{12}]^-$ is compensated by the formation of enhanced bonds from neighbouring, saturated, Ru centres, that were previously weaken due to the trans influence of carbonyl ligands, leading to an increase in Ru-Ru bond order. These changes in Ru-Ru bonding are localised in the subset of bonds that do not have hydride bridges, and bond lengths contract from characteristic Ru-Ru single bond lengths of ∼2.80 Å to ∼ 2.65 Å, reflecting the emergence of enhanced bonding character. In the limit that three CO ligands are removed ($[H_3Ru_4(CO)_9]^-$), the cluster presents an essentially flat $Ru_3(\mu\text{-H})_3$ surface that supports the subsequent reactivity with alkenes (summarised in Figure 3.17).

In the initial stages of reactivity, the alkene, 1-hexene in this case, acts as a simple 2-electron donor, binding in an η^2 mode to saturate one metal centre. While the cluster remains unsaturated, a series of almost barrierless C-H bond activation steps can occur to generate η^3-allyl and $\eta^{2:2}$-diene a motifs, along with additional hydrides. When the hydride load builds beyond 4, loss of H_2 can initiate a further sequence of C-H bond activation steps.

The potential energy surfaces of all of the unsaturated intermediates discussed here are extremely flat, and feature multiple minima of very similar energies, differing in the location of the CO ligands. In such circumstances, models of reactivity based on a small sample of minimum energy structures must be treated with some caution.

However, by placing the emphasis clearly on the general features of structure and bonding that are common to all low-energy isomers (in this case the increase in Ru-Ru bond order), we can hope to draw out aspects of the reaction that are not dependent on the precise locations of the CO ligands.

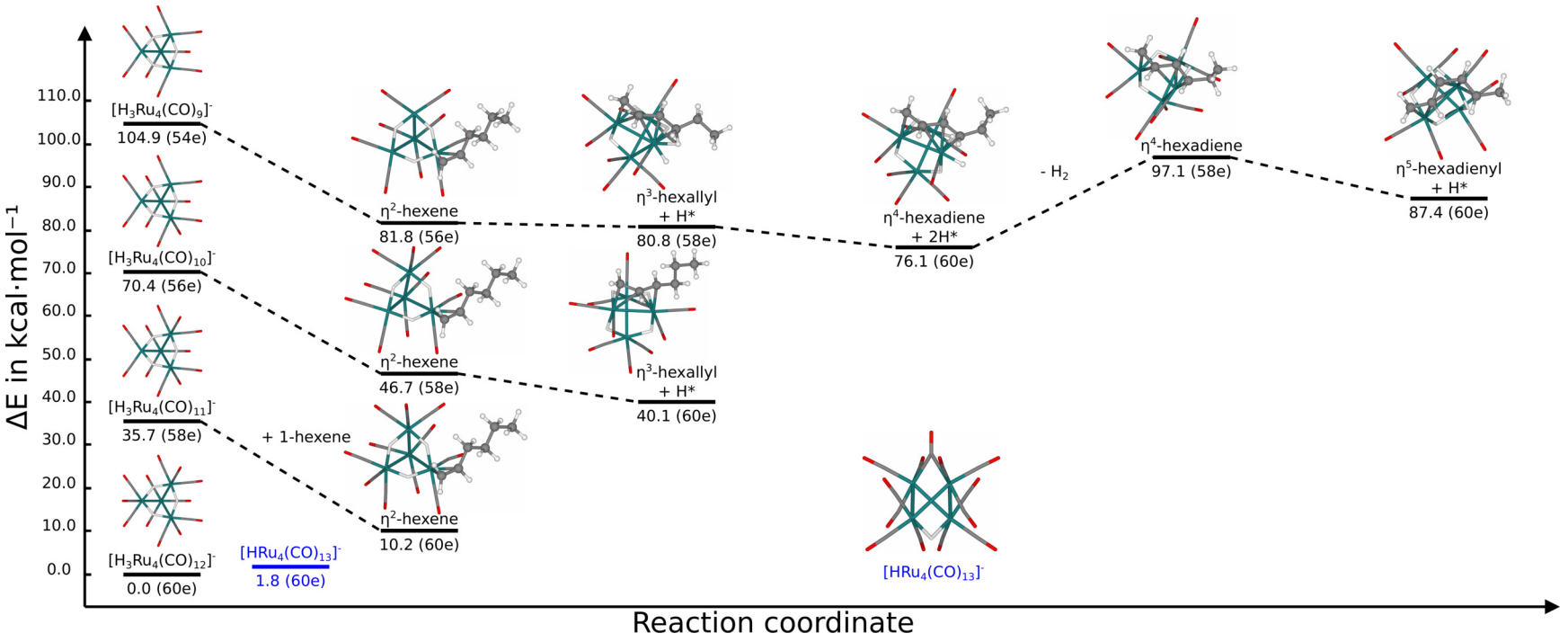


Figure 3.17: Summary of structures and energies of 1-hexene bonded clusters with different bonding modes.

Chapter 4

Adsorbate-substrate interactions for a single Ru_3 cluster embedded on N-graphene

4.1 Introduction

Heterogeneous catalysts play crucial roles in the majority of important industrial chemical manufacturing process, including the Haber-Bosch process, Fischer-Tropsch synthesis, water-gas shift reaction and fat hydrogenation. They also provide the foundations for clean and renewable technologies such as fuel cells [110] and carbon capture and storage [111]. Traditionally, heterogeneous catalysts based on precious metal groups (for example, platinum) have been utilised extensively in the aforementioned processes, and they remain amongst the best performing catalysts for many reactions. However, due to the high economic cost of their production, along with significant environmental concerns, the use of precious metals at current rates is unlikely to be sustainable in the long term. As a result, there is a strong impetus to identify alternatives. Amongst these, metal-free catalysis is an important and emerging field, but within the transition metal domain, a popular strategy is to reduce the number of precious metal atoms required per active site, by dispersing atoms on solid-state surface in various forms such as metal-semiconductor heterostructures [112, 113], high-entropy alloys [114, 115], metal-organic frameworks [116–118], single-atom catalysts [119, 120], single-atom alloys [121, 122] and single-cluster catalysts [6, 123]. Homogeneous catalysts can be seen as the logical limit of this approach, but challenges in recycling solution-phase components are well documented.

Graphene-based single atom catalysts (SACs) [124, 125] have emerged as an important class, and they have been the subject of numerous experimental and com-

(a) Periodic models of SACs (b) Adsorption free energy of OH versus the descriptor φ

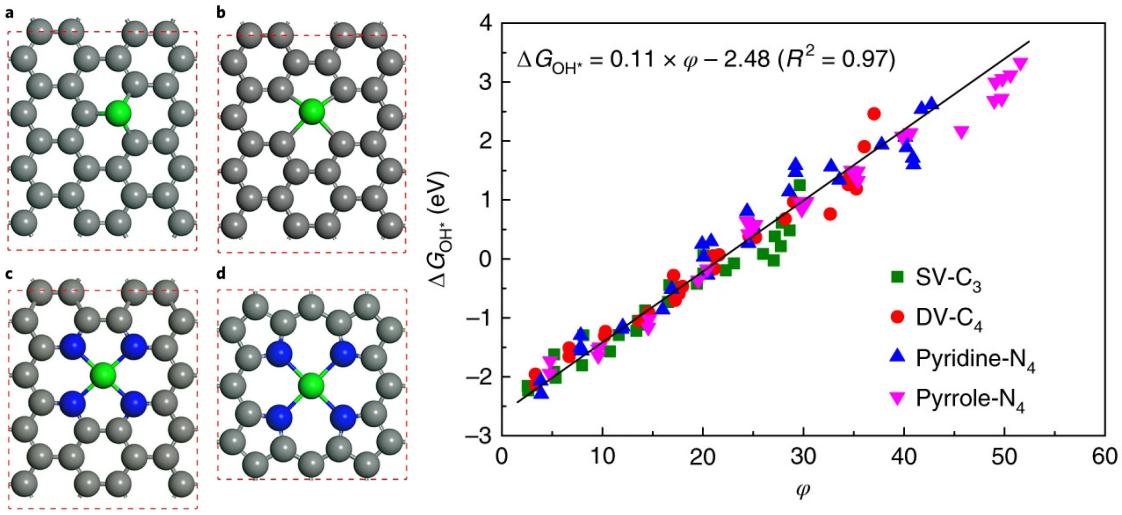


Figure 4.1: (a) Various periodic models of graphene-based SACs with different species supporting the metal centre and (b) the role of the universal descriptor φ in determining the adsorption energy of OH on the models. The models considered here are single metal centres supported on single-vacancy (SV-C_3), double-vacancy (DV-C_4), double-vacancy with 4 pyrridinic-N (pyridine-N_4) and porphyrine-like defect with 4 pyrrolic-N (pyrrole-N_4) on graphene. Figures adapted with permission from reference [128], Springer Nature.

putational frameworks in recent years. Graphene-based SACs, typically described as "M-N-C" where a transition metal atom (M) is supported on nitrogen (N) doped graphene (C) in a range of local coordination environments, have been shown to exhibit (electro)catalytic activities. They are, in many ways, analogous to the structures of metal-macrocyclic molecules, which are also catalytically active, and bridge the gap between homogeneous and heterogeneous catalysts. Many different forms of graphene-based catalysts are accessible through a wide range of experimental techniques [126], and their (electro)catalytic performance typically exceeds that of pristine graphene. However, trial-and-error approaches to optimise performance have been criticised by Wang *et al.* for not bringing significant insight – the authors went so far as to suggest that "if we spit on graphene it becomes a better electrocatalyst" [127]. This statement makes clear the need for careful experimental and computational studies of SACs, such that the nature of the active sites and an intimate understanding of reaction mechanism can emerge.

Computational studies are now an integral component of most studies of catalysis, and the SAC field is no exception. For example, detailed theoretical analysis of various

graphene-based SACs and metal–macrocycle complexes has been conducted by Xu and co-workers [128], who related their catalytic activity to composition. In their work, they proposed a universal descriptor for SACs:

$$\varphi = \theta_{\mathrm{d}} \times \frac{E_{\mathrm{M}} + \alpha \times (n_{\mathrm{N}} \times E_{\mathrm{N}} + n_{\mathrm{C}} \times E_{\mathrm{C}})}{E_{\mathrm{O/H}}} \tag{4.1.1}$$

where θ_{d} represents the *d*-electron count on the metal atom, E_{X} represents the electronegativity of element X, and n_{N} and n_{C} represent the number of nearest-neighbour N and C atoms respectively. $E_{\mathrm{O/H}}$ represents the enelectronegativity of either oxygen or hydrogen atoms, depending on the type of adsorbates. The corrective coefficient α is set to be 5/4 for pyrrolic-N_4 defects and 1 in other model supports (Figure 4.1). This universal descriptor of graphene-based SACs can provide crucial information that is required to predict catalytic performance. For example, the adsorption free energy of OH on a metal atom can be calculated from the empirically fitted equation:

$$\Delta G_{\mathrm{OH^*}}/eV = 0.11 \times \varphi - 2.48. \tag{4.1.2}$$

The presence of parameters derived from the local environment of graphene-based SACs in the universal descriptor, equation 4.1.1, such as coordination number and type of N defect, immediately flags the importance of defect structures in controlling catalytic performance. Using the adsorption energies calculated from this descriptor, the authors were able to predict the optimal catalytic centres, for examples, Fe-pyridine/pyrrole-N4 for the Oxygen reduction reaction (ORR), Co-pyrrole-N4 for Oxygen evolution (OER) and Mn-pyrrole-N4 for Hydrogen evolution (HER).

A new class of catalysts has emerged as a derivative of single-atom catalysts, the so-called single-cluster catalysts (SCCs). SCCs typically consist of atomic clusters that are 2 to 20 atoms in size and are supported on substrates, but are smaller than nanoparticle catalysts and have precise numbers of atoms that determine their properties. The transition from single to multiple atoms not only expands the chemical space available for exploration within active sites but also unlocks novel reaction pathways previously inaccessible with single atoms. As was the case with SACs, computational approaches are vital in identifying the active species and understanding the electronic structures of SCCs, in particular those with double or triple atoms. However, while data-driven techniques including machine learning are extensively used in the realm of SACs, the complexity of structure and bonding inherent in SCCs poses a challenge for building predictive models for designing a new catalysts. As of now, general theoretical frameworks concerning the structure and catalytic activities of active species in SCCs remain elusive.

The particular example of an SCC that we consider here was prepared by Ji *et al.* [13] by confining triruthenium dodecacarbonyl, $Ru_3(CO)_{12}$, inside the cage of a metal-organic framework (MOF). The material was then pyrolysed to replace carbonyl ligands with donors from the nitrogen-doped carbon composite. The authors have reported that the catalytic activity of this material in oxidation of 2-aminobenzyl alcohol to 2-aminobenzaldehyde is up to 23 times faster than a ruthenium single atom or nanoparticle supported on similar materials. (A mixture of 2-aminobenzyl alcohol, 0.001 equiv Ru species catalysts and 2.0 equiv. H_2O_2 in MeCN solution heated at 80 ^{o}C and vigorously stirred.) From their analysis based on X-ray absorption spectroscopy and scanning transmission electron microscopy, a model of the triruthenium cluster supported by four pyrrolic nitrogen atoms of porous nitrogen-doped graphene (Figure 4.6a) was proposed. However, it should be noted that, while the structures of small graphene defects have been widely studied, the structures of larger defects that can accommodate clusters are less understood and no single model has yet been universally accepted by the community. Therefore, one must carefully consider the structures of N-graphene models and their effects on the electronic structure of supported metal atom/clusters.

Several computational investigations of the structure of triple-atom clusters supported by carbon materials have been reported in the literature, but only a handful on SCCs supported on the specific N-graphene model proposed by Ji *et al.* Notable amongst these was a DFT study of the electrochemical nitogen reduction reaction (NRRs) on triatomic A_3 clusters by Zheng *et al.* [129] and the screening of triple-atom A_2B clusters for CO_2 electrochemical reduction with active learning in conjunction with DFT by Li *et al.* [130]. Other than these, computational studies on catalysis by triple-atom clusters have mostly been conducted using an N-graphene model that presents 6 pyrinidic-N species as donors[131–133], a graphdiyne model coordinating *via* π orbitals of the *sp*-hybridised carbon species [134–136], or reduced graphene oxides with C/O bonding modes[137]. In all cases, significantly fewer donor atoms are available to support the clusters compared to electron-precise molecular counterparts such as $Ru_3(CO)_{12}$. The importance of electronic and coordinative unsaturation makes a direct link to the work on Ru_4 clusters, also reported in Chapter 3 of this thesis, where the catalytic species are much closer to saturation. Direct comparisons of different local atomic environments around the cluster and their effects on the stability and catalytic activity have not yet been investigated in detail.

In this chapter, I discuss the structure and stability of a model SCC Ru_3 supported on various models of nitrogen-doped graphene-based materials. Specifically, three

different models of N-graphene will be considered:

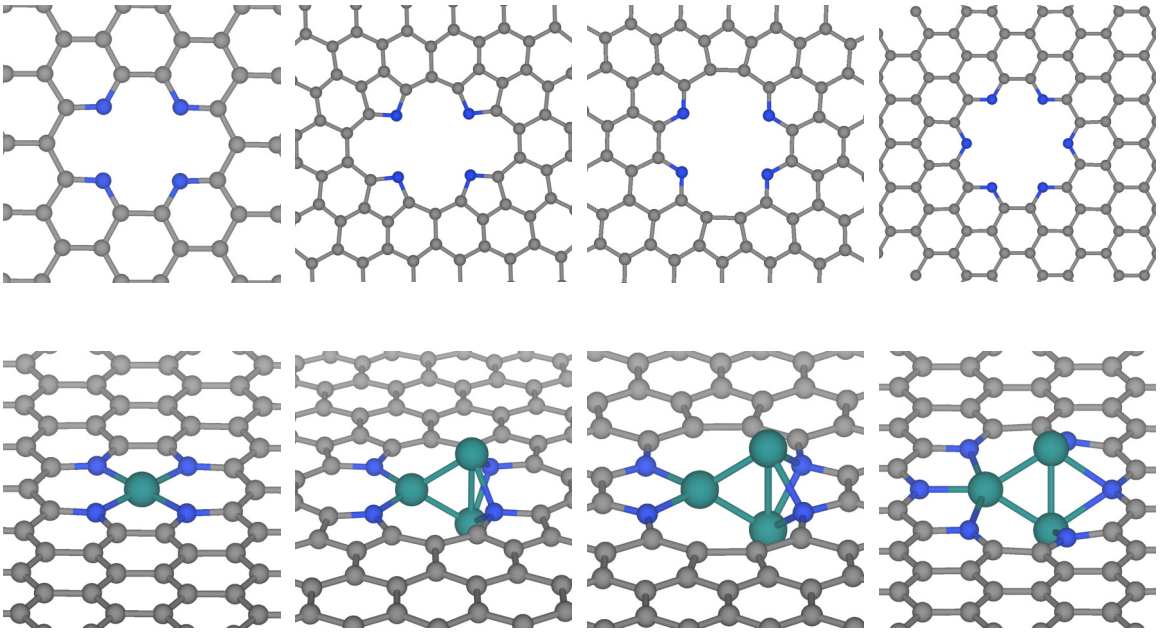
1. the pyrrolic-N_4 model proposed by Ji *et al.*
2. a pyridinic-N_4 model
3. a pyridinic-N_6 model which has been widely used in the literature.

Both periodic and cluster DFT calculations were run in order to analyse the electronic interactions between the cluster and the substrates, and to gain additional insight into the important orbital interactions. Adsorption energies of oxygen atoms are calculated on different models and the extent to which the local atomic environments influences catalytic performance will be determined. Finally, a reaction mechanism for ethanol to formaldehyde conversion is proposed based on the assumption that oxo species are the active sites, as H_2O_2 was used as an oxidiser in the experiment.

4.2 Computational Methods

The calculations of periodic models in this chapter were performed using plane-wave pseudopotential DFT with periodic boundary conditions (PBC) as implemented in the Vienna *ab initio* Software Package (VASP) [138–141]. The exhange-correlation energy was modelled using the PBE functional [142], as well as r^{2}SCAN [91]. The dispersion corrections are added by using the D4 corrections. The valence electron configurations are $2s^2\,2p^2$ for C, $2s^2\,2p^3$ for N, $2s^2\,2p^4$ for O and $4s^2\,4p^6\,5s\,4d^7$ for Ru, with the core electrons treated using PAW pseudopotentials. The plane-wave cut-off was set to 400 eV and the Brillouin zone was sampled on a $3 \times 3 \times 1$ Γ-centred grid for a 9×9 supercell of graphene, and $6 \times 6 \times 1$ Γ-centred grid for smaller unit cells shown in Figure 4.11. The structural relaxation was considered converged when the Hellmann-Feynman force fell below 10^{-3} eV Å^{-1}, while the SCF convergence criterion was 10^{-6} eV. The minimum energy path between the orthorhombic and tetragonal phases was determined using the climbing-image nudged elastic band (NEB) approach as implemented in the VTST code [143, 144] with 6 intermediate images between the two minima and a spring constant of 5 eV Å^{-2} between images. The total force was minimised to below 10^{-3} eV Å^{-1} in these calculations.

Energies of finite models were obtained by using DFT calculations as implemented in ADF 2021.104 [86]. All-electron Slater-type TZ2P orbital was used as a basis set [87] and scaler ZORA correction [88–90] was used. Meta-GGA r^{2}SCAN functional



(a) SAC/N_4V_2 (b) pyrrolic-N_4 (c) pyridinic-N_4 (d) C_2N-type

Figure 4.2: Periodic models of defects on N-graphene (top), and the embedded Ru atom or Ru_3 cluster on the defect (bottom). (a) Ru SAC embedded on divacancy with 4 pyridinic-N (SAC/N_4V_2), and Ru_3 clusters embedded on defects with (b) 4 pyrrolic N (pyrrolic-N_4), (c) 4 pyridinic N (pyridinic-N_4) and (d) 6 pyridinic N (C_2N-type). Carbon, nitrogen, and ruthenium atoms are coloured grey, blue and cyan respectively.

[91] implemented in LibXC [92] along with D4 dispersion corrections parametrised by Ehlert et al [93, 94]. was employed as density functional approximation. "Good" numerical quality was used throughout the calculation.

4.3 Electronic structure

4.3.1 Structure of N-graphene

Four different models for defect N-graphene are summarised in Figure 4.1. The simplest defect model, labelled the 4 pyridinic N defect (N_4V_2), is formed simply by removing two adjacent carbon atoms from the graphene surface and replacing the unsaturated carbons around the vacancy with pyridinic nitrogen atoms. This simplest vacancy is, however, not large enough to accommodate more than one Ru atom. The pyrrolic-N_4 defect, in contrast, has a much larger pore. This type of defect is created by removing a unit of naphthalene from graphene and bridging two adjacent unsaturated carbons with a nitrogen atom that forms a five-membered ring. A similar defect structure has been experimentally reported for silicon analogue by Lee

et al. [145]. Defects with pyridinic N atoms can be derived conceptually by excising a benzenoid unit and replacing the terminating carbon atoms with nitrogen atoms, analogous to pores found in graphitic-C_2N structures (Figure 4.2d) with a similar size that can acommodate the cluster. However, in these pores, the cluster would be interacting with 6 nitrogen species as opposed to 4 nitrogen species suggested by Ji *et al.* [13]. In this project, we also have constructed a model that is similar to the C_2N model but 2 N atoms are removed so that only 4 nitrogen species are there to support the cluster. However, this also results in the defect structure with two five-membered rings in the vicinity of the cluster. Note that all four defect models are charge-neutral and no charge carriers are generated in the substrate due to these defects.

The geometries of embedded Ru_3 clusters on these models are all optimised to the most stable configurations. In the pyrrolic-N_4 model, the structure of the cluster converges into a C_{2v}-symmetric configuration, with one Ru atom is in plane with the graphene sheet while two other Ru atoms are located out of the plane. The in-plane Ru atom is coordinated with two nitrogen atoms terminally, and the out-of-plane Ru atoms are bridged by the other two nitrogen atoms. The bond length between the in-plane Ru and terminal N is 1.95 Å, and the out-of-plane Ru and bridging N is 2.04 Å. The geometry of the Ru_3 cluster itself is close to equitriangular, with bond lengths 2.34 Å between in and out-of-plane Ru atoms, and 2.58 Å between two out-of-plane Ru atoms. These observations are consistent with the extended X-ray absorption fine structure (EXAFS) measurements from Ji *et al.* [13], reporting mean bond lengths of 1.99 ± 0.02 Å and 2.53 ± 0.02 Å for Ru–N and Ru–Ru bonds respectively. The structure of the cluster on the pyridinic-N_4 model adopts the same configuration, with similar Ru-N bond lengths: 2.05 Å (terminal-N) and 2.12 Å (bridging-N) but slightly shorter average Ru-Ru bond lengths: 2.35 Å (in and out-of-plane) and 2.47 Å (two out-of-plane). On the other hand, the cluster embedded on the C_2N-type defect converges to the C_2-symmetric configuration, with one in-plane Ru coordinated by 3 N atoms terminally, and two out-of-plane Ru atoms are bridged by one N atom and each of them is terminally coordinated by two remaining N atoms. The bond lengths between in-plane Ru and terminal N atoms are 1.97 Å and 2.38 Å, and between out-of-plane Ru and bridging/terminal N atoms are 2.20/1.92 Å respectively. The structure of the cluster itself is again close to equitriangular with relatively shorter distances 2.32 Å and 2.45 Å compared to EXAFS.

In order to establish the relative stabilities of the N-graphene models, the formation energies of defects are calculated with plane-wave DFT and the following

	E_{vac}	E_{form}	E_{bind}	E_{agg}
SAC/N_4V_2	3.56	-4.26	-7.82	-0.34
pyrrolic-N_4	13.17	-5.00	-6.06	1.43
pyridinic-N_4	11.55	-3.28	-4.94	2.54
C_2N	4.42	-11.82	-5.41	2.07

Table 4.1: Formation energies of N-graphene defects and SAC/SCCs, and their binding and aggregation energies.

expression:

$$E_{\text{vac}} = E_{\text{N-graphene}} - E_{\text{graphene}} + n_{\text{C}}\mu_{\text{C}} - n_{\text{N}}\mu_{\text{N}} \tag{4.3.1}$$

where $E_{\text{N-graphene}}$ and E_{graphene} are the total energies of the N-graphene model and pristine graphene, n_{N} and n_{C} are the numbers of N atoms added and C atoms removed, and μ_{N} and μ_{C} are the chemical potentials of N and C respectively. The chemical potentials are obtained from the energies of a N atom in an N_2 molecule and a C atom in pristine graphene. The first two terms describe the 'pure' formation energy of an isolated defect, which is defined as the energy difference between large cells of graphene containing a defect and pristine graphene [146]. The formation energy of the defect model (by using a 9*9 supercell of pristine graphene) is found to be 13.17 eV for the pyrrolic-N_4 defect, 11.58 eV for the pyridinic-N_4 defect and 4.42 eV for the C_2N-type defect (one cavity in C_2N). The much higher formation energy of the pyrrolic-N_4 defect suggests that its concentration would be much lower than that of the pyridinic-N defect. However, in the original paper by Xu [128] the defect is stabilised by the Ru_3 cluster, and effect that is not taken into account in the above expression. We have, therefore, also calculated the formation energy of the triple atom catalysts using the following expression:

$$E_{\text{form}} = E_{\text{Ru}_n/\text{N-graphene}} - E_{\text{graphene}} + n_{\text{C}}\mu_{\text{C}} - n_{\text{N}}\mu_{\text{N}} - n\mu_{\text{Ru}} \tag{4.3.2}$$

where $E_{\text{Ru}_n/\text{N-graphene}}$ is the total energy of the catalyst, μ_{Ru} is the chemical potential of Ru calculated from the energy of the Ru atom in vacuum, and n is the number of Ru atoms in the catalyst. In this way, the overall stability of the SCCs with the effect of N-graphene support can be determined. To ensure that the supported atoms on the substrate are thermodynamically stable and do not aggregate to form nanoparticles, the aggregation energy is also calculated [147], according to:

$$E_{\text{agg}} = E_{\text{bind}} - E_{\text{coh}} \tag{4.3.3}$$

where

$$E_{\text{bind}} = (E_{\text{Ru}_n/\text{N-graphene}} - E_{\text{N-graphene}} - n\mu_{\text{Ru}})/n \tag{4.3.4}$$

is the binding energy of the catalyst per Ru atom, and

$$E_{\text{coh}} = E_{\text{Ru(bulk)}} - \mu_{\text{Ru}} \tag{4.3.5}$$

is the cohesive energy of a Ru atom in bulk. Negative values of this quantity indicate that the catalysts are stable and therefore not susceptible to aggregations.
The summary of the various formation energies can be found in Table 4.1, and it is found to be -4.26 eV for the pyrrolic-N_4, -5.00 eV for the pyridinic-N_4 and -3.28 eV for the C_2N-type models. While the formation energy of the pyrrolic-N defect is higher than that of the pyridinic-N defect, the formation energy of the SCC supported by the pyrrolic-N_4 defect is indeed lower than that of the SCC supported by the pyridinic-N_4 defect by 1.7 eV, and comparable to that of the SAC supported by a N_4V_2 defect with a difference of 0.7 eV. This is a clear indication that there is a much stronger interaction between the Ru_3 cluster and the pyrrolic-N_4 model. Amongst all models considered here, the SCC supported by the C_2N-type defect has the lowest formation energy (-11.82 eV). However, in terms of likelihood of aggregation, all of the SCC models considered have positive aggregation energies, with 1.43 eV for the pyrrolic-N_4, 2.54 eV for the pyridinic-N_4 and 2.07 eV for the C_2N-type models, implying that they might not be thermodynamically stable, in contrast to the SAC model which has a negative value. Nevertheless, amongst the three models, the pyrrolic-N model has the lowest aggregation energy and lower than the pyridinic-N_6 model by 0.5 eV, despite that it has two fewer N species. Therefore, the Ru_3 cluster on the pyrrolic-N model might be actually the least unstable model. For comparison, the aggregation energy of Ru_3 cluster on graphdiyne model is reported to be 1.74 eV by Liu *et al.* [148] and therefore they are also susceptible to aggregations.

To investigate the origin of difference in the formation and binding energies, we analysed the electronic structures of the defects on N-graphene on their own. The comparisons of band structures in pristine graphene and various N-graphene models can be found in figure 4.3. Note that a 9 by 9 supercell based on a graphene primitive cell is used in this band structure calculation. Due to band folding, the Dirac cones of pristine graphene (which are located at K and K' in the primitive cell) are now shifted to Γ in this superstructure. The band structures of our defect models show many of the features of other N-graphene models with smaller defects reported in previous studies [149]. These include a band gap opening, a shifting of the π band energies, flat bands originating from N dangling bonds and hybridisation of defect states with

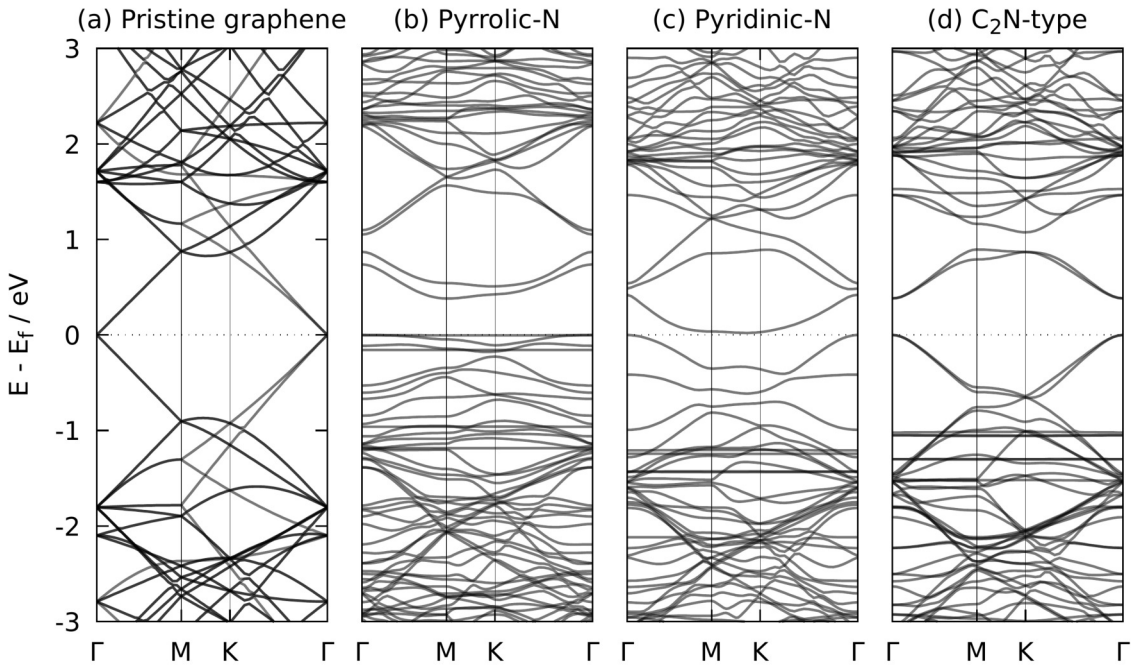


Figure 4.3: PBE band structures of a 9 by 9 pristine graphene cell and the cells with various types of defects. The Fermi energy is adjusted to the top of the conduction band.

the bulk π bands. In the pyrrolic-N_4 model, two defect states can be found between the Fermi level and the bottom of the conduction band. While these defect bands have similar character to the bulk π bands around the Γ point, the orbitals in other locations in the Brillouin zone show significant localisation. This is illustrated by the band-decomposed charge density of these defect bands shown in Figure 4.4, in which the density is strongly localised around the defect cavity for both the first and second unoccupied bands. The shapes of the charge densities are also consistent with the shapes of LUMO and LUMO+1 of the macrocyclic molecule, which we will use as a molecular model of the N-graphene in the next section. In the pyridinic-N_4 model, there is only one unoccupied defect band below the bottom of the bulk valence band, which is also hybridised with the bulk π bands and localised around the defect cavity. In the C_2N model, no defect π bands are generated in between bulk conduction and valence bands.

4.3.2 Cluster-defect orbital interactions

In order to investigate the electronic interactions between the cluster and various types of N-graphene models, we start by analysing the fragments. Consider the set

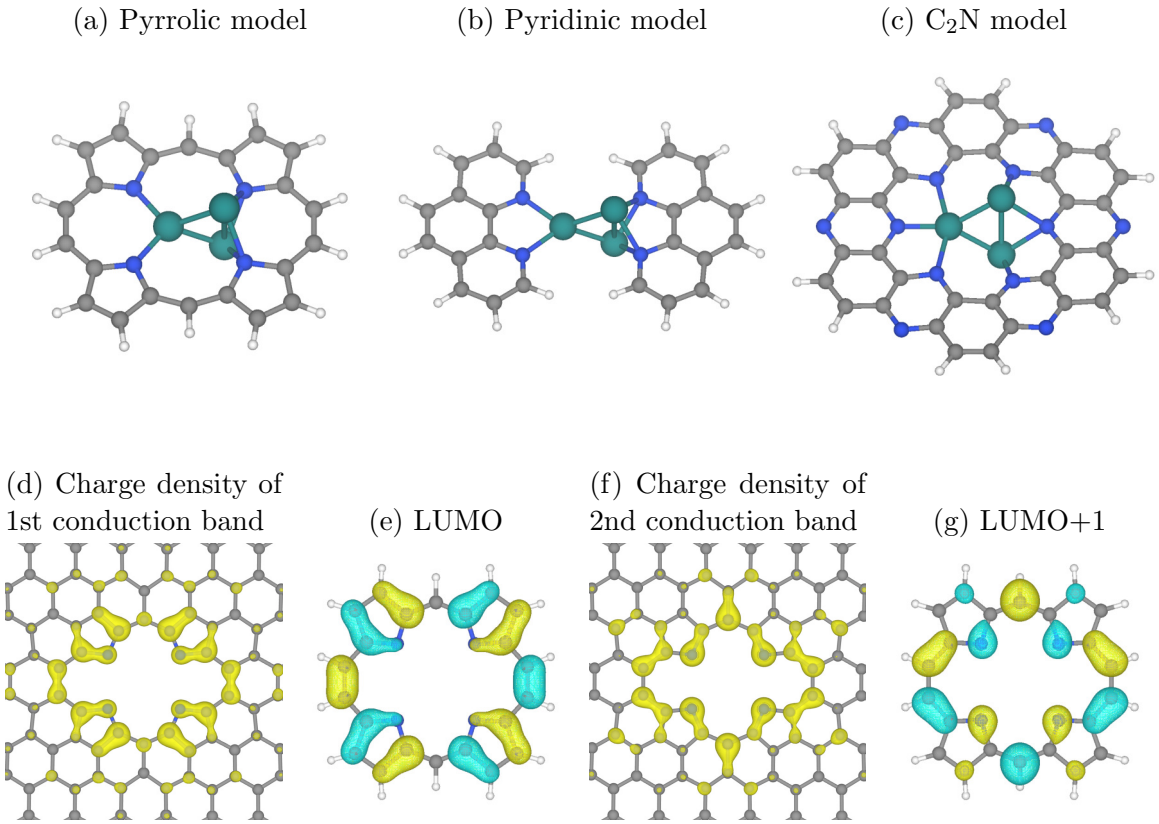


Figure 4.4: Macrocyclic molecular models for various N-graphene models.

of molecular orbitals of an isolated naked Ru_3 cluster with point group D_{3h} shown in Figure 4.5. The valence electrons of Ru fill orbitals formed by the linear combinations of Ru 4d and 5s. The result of DFT calculations at PBE-D4 and r2SCAN-D4 level shows that the ground-state of the cluster is a nonet and around the top of bands formed by the 4d atomic orbitals, molecular orbitals have a high antibonding character. While the energy level of 5s orbitals is significantly higher than that of 4d orbitals, the energy of the molecular orbital ($2a_1''$) formed by the totally symmetric combination of 5s atomic orbitals drops just below the most antibonding combination of 4d orbitals, $17e_1'$. Consequently, there are a total of 16 accessible molecular orbitals (15× Ru 4d + 1× Ru 5s) to accommodate 24 electrons. The result is 8 unpaired electrons, with the antibonding $5a_2'$ orbital being the SOMO. There are no signs of multiconfigurational character based on the value of $< S^2 >$ from the DFT calculations.

Turning now to the interaction between the Ru_3 fragment and the various defect models, we have identified simplified molecular fragments that replicate the key features of the periodic system: the Ru_3-containing defects and the models are shown

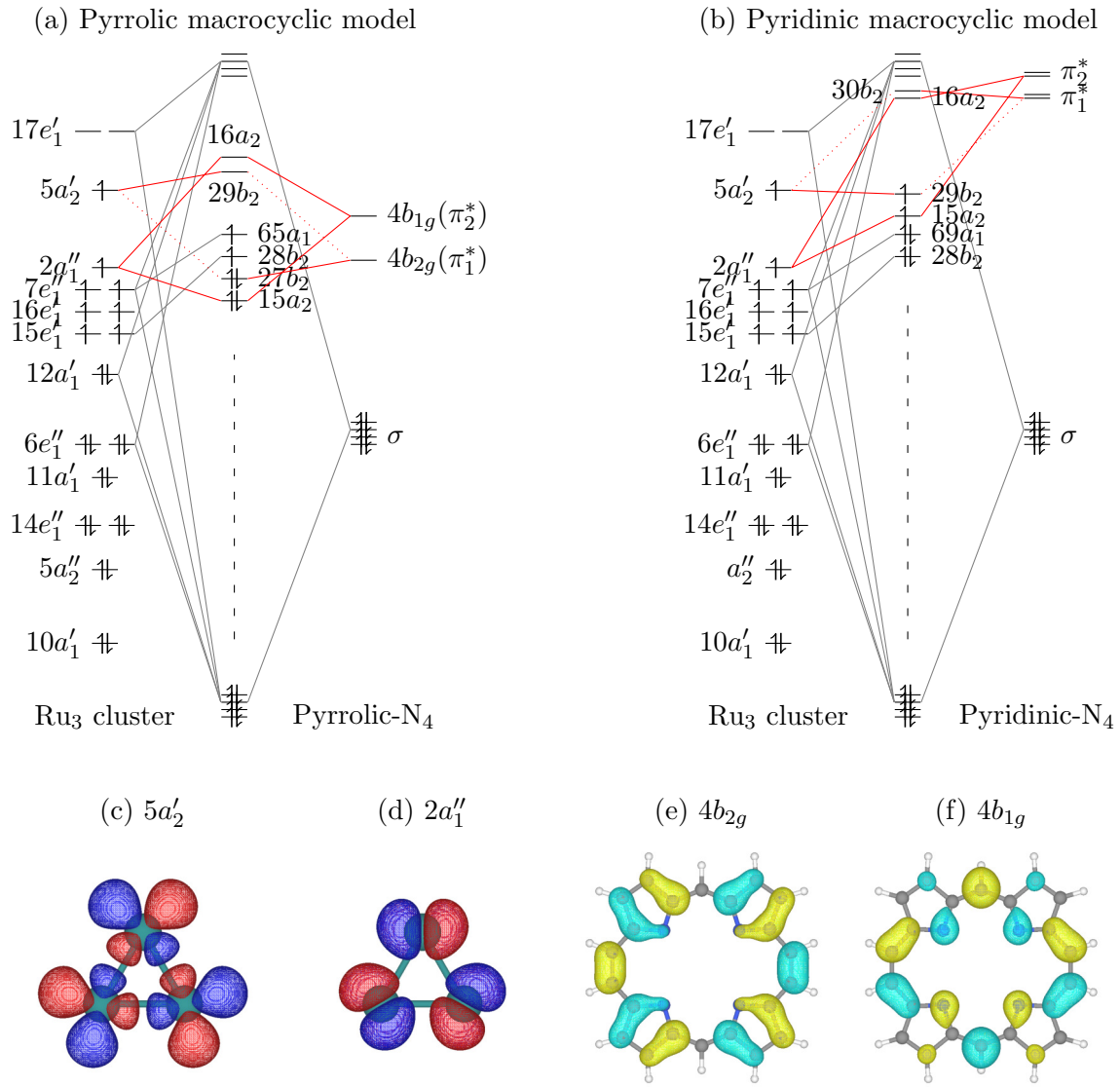


Figure 4.5: Molecular orbital diagram showing the orbital interaction between Ru_3 and N-graphene in (a) pyrrolic-N_4 and (b) pyridinic-N_4 models. Only the molecular orbitals significantly interacting *via* σ or π-type interactions with ligands are shown in the diagram orbitals and the rest are omitted as denoted by the dashed line. π-type interactions are coloured in red. In the pyrrolic-N_4 model, significant π-type interaction (indicated with red lines) can be found between (d) $2a_1''$ orbital of Ru_3 and (f) $4b_{1g}$ orbital of the macrocylic molecule. Two electrons are transferred from the cluster to (e) $4b_{2g}$ orbital of the macrocylic molecule and (c) $5a_2'$ orbital of Ru_3 cluster become completely unoccupied.

in Figure 4.2. The charge analysis based on both Mulliken and Bader protocols is summarised in Table 4.2, for both periodic (Bader only) and molecular (both). In all cases, the most significant interaction between the Ru_3 cluster and the N-graphene

Table 4.2: Summary of charge analysis on the Ru_3 cluster with different models calculated at PBE-D4 and r2SCAN-D4 level.

		PBE-D4		r^2SCAN-D4	
Model		$q_{Ru_3}^{Bader}$	$q_{Ru_3}^{Mulliken}$	$q_{Ru_3}^{Bader}$	$q_{Ru_3}^{Mulliken}$
Pyrrolic	Periodic	1.63		1.66	
	Finite	1.56	1.26	1.64	1.43
Pyridinic	Periodic	1.42		1.66	
	Finite	0.97	0.79	1.17	1.04
C_2N	Periodic	1.61		1.72	
	Finite	1.62	1.40	1.70	1.55

arises from the σ-type interaction from the dangling lone-pair electrons on the N atoms. These lone pairs interact with four of the Ru_3 molecular orbitals, three of which are occupied Ru_3 orbitals (two linear combinations of 4d and the symmetric combination of 5s orbitals noted above). The 4th linear combination of lone pairs on nitrogen interacts with one of E' linear combinations of Ru 5s orbitals that was already vacant in Ru_3. The end result is that three linear combinations of Ru-based orbitals are destabilised, leaving only 13 molecular orbitals to accommodate the 24 electrons, and hence a triplet ground state for both the pyridinic-N_4 and pyrrolic-N_4 models.

In addition to the σ interactions, interactions of π symmetry with the delocalised π system of the N-graphene are also significant. Typically, strong π-acceptors such as carbonyl and cyanide are required to stabilise the metal orbitals via π-backbonding. However, the ability of ligand to accept π density increases significantly if the ligand contains five-membered rings (indeed, this is the origin of the 5/4 enhancement factor in the universal descriptor for SACs in equation 4.1.1). The N-graphene model proposed by Xu Xu *et al.* [128] is reminiscent of porphyrins except that two extra carbon atoms are added to the inner ring component. The charge distribution for the pyrrolic-N_4 model shown in Table 4.2 indicates a very substantial charge transfer from the Ru_3 fragment to the LUMO of the N-graphene model, giving the cluster and the ligand formal charges of +2 and -2 respectively. The charge distributions are strikingly similar for both the periodic system and the molecular model, suggesting that the latter reproduces faithfully the frontier orbital array. The result is a total of 22 Ru electrons valence electrons to fill the available Ru_3 orbitals. The interaction of the π^* ligand on the surface effectively destabilises the $5a_2''$ orbital, leaving only 12 orbitals to accommodate 22 electrons, again yielding a triplet spin configuration.

In pyridinic-N_4 model, in contrast, the charge transfer from the cluster to ligand is less pronounced (Table 4.2) and the π-backbonding interaction is too weak destabilise the $5a_2''$ antibonding orbital, as was the case for the pyrrolic model (Figure 4.5(b)). Therefore, 24 valence electrons are available to fill 13 molecular orbitals, again yielding a triplet spin configuration. We note here that there is a significant discrepancy between the charge distributions for the periodic and molecular models in Table 4.2, which suggests that our finite model is a less faithful representation of the frontier region in this case.

By utilising the fragment analysis implemented in ADF, it is possible to partition the orbital interaction energy between the cluster and substrate into contributions from orbital interactions of different symmetry types. Since the pyrrolic-N_4 and pyridinic-N_4 models have the same C_{2v} symmetry, it is straightforward to separate the π and σ-type interactions ($\{a_2,\ b_2\}$ symmetry and $\{a_1,\ b_1\}$ symmetry, respectively. Table 4.3 shows the energy decomposition analysis of Ru3 and N-graphene fragments, calculated at both PBE-D4 and r^{2}SCAN-D4 levels. As expected based on the previous discussion of charge separation, the π-type contribution to the orbital interaction energy between the two fragments is higher with pyrrolic-N_4 models compared to the pyridinic-N_4 models, with a difference of roughly 5-7 eV. On top of the π-type orbital interaction, there are some minor differences in the steric interaction and the σ-type orbital interaction energies, presumably due to the effect of charge transfer from the cluster to N-graphene. There is a little difference in the corrective dispersion energy term in two models. The pyrrolic and C_2N models have very similar total bond energy decomposition but, due to the lack of symmetry in the latter model, it is not possible to decompose the orbital energy further into π and σ contribution. We would, however, anticipate a greater contribution from σ-type interactions in the C_2N models simply because there are six nitrogen species supporting the cluster as opposed to just four.

Table 4.3: Fragment analysis of molecular models listed in figure 4.2. The molecular models are decomposed into the Ru_3 cluster and macrocyclic molecular fragments. The values of energies are given in eV.

	PBE-D4			r^2SCAN-D4		
	pyrrolic-N_4	pyridinic-N_4	C_2N	pyrrolic-N_4	pyridinic-N_4	C_2N
Steric interaction:	19.80	16.07	20.19	20.22	17.06	19.98
Pauli repulsion	52.68	43.00	52.91	54.17	45.31	52.65
Electrostatic interaction	-32.87	-26.93	-32.72	-33.96	-28.25	-32.67
Orbital interaction:	-28.52	-20.35	-28.76	-29.29	-21.36	-28.60
σ contribution	-7.70	-6.08		-7.47	-5.23	
π contribution	-21.82	-14.27		-22.74	-17.58	
(MetaGGA correction)				0.92	1.45	1.46
Dispersion:	-1.10	-0.89	-1.18	-0.31	-0.28	-0.38
Frozen bond energy	-9.80	-5.17	-9.74	-9.38	-4.57	-9.00

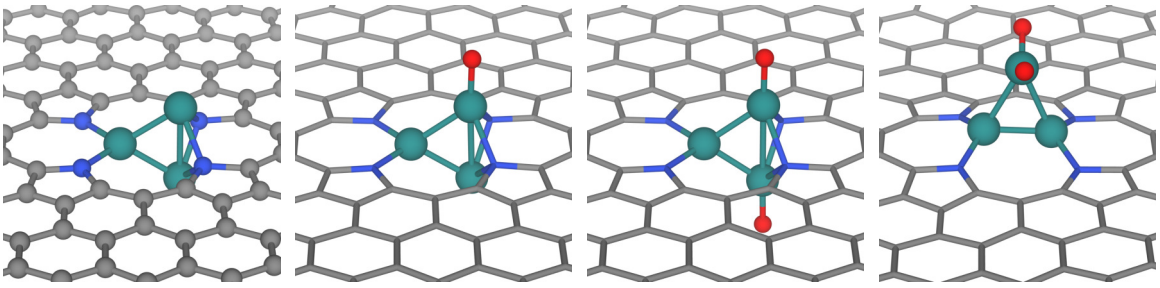
4.4 Formation of oxo active species

As H_2O_2 was used as the oxidising agent in the experiment, we first investigated the adsorption energy of H_2O_2 on Ru_3/N-graphene in the absence of solvents. However, we were unable to identify local minima corresponding to the adsorption of H_2O_2 by the Ru cluster; instead, the H_2O_2 molecule decomposes spontaneously into one water molecule and an oxygen atom which binds to the Ru cluster in the form of an oxo ligand (Figure 4.6b). Similarly, attempts to bind two molecules of H_2O_2 lead directly to dioxo species, either *trans*-dioxo-Ru_3/N-graphene (Figure 4.6c) or *cis*-dioxo-Ru_3/N-graphene (Figure 4.6d) depending on the direction of approach of the two H_2O_2 molecules. In the *cis*-dioxo structure, the Ru cluster is raised above the surface and is bound to it by only two of the three Ru atoms. The *trans*-dioxo species is reminiscent of the isolated $[Ru^{VI}(Por)(O)_2]$ complex, and is energetically more favourable than the *cis* isomer by 0.85 eV. It is not clear, however, whether both sides of the graphene sheet would be accessible to the solvents which is essential for the formation of the *trans* isomer, so it may be that the *cis* isomer is kinetically favoured under the prevailing reaction conditions. Placing H_2O_2 molecules on Ru_3/C_2N model yields similar results, the thermodynamic sink being *trans* or *cis*-dioxo-Ru_3 species along with surface-bound water molecules.

To understand the reactivity of the catalysts, we calculate the formation energies of oxo species, according to:

$$\Delta E_{n\text{oxo}} = E_{n\text{oxo}} + nE_{H_2O} - E_{Ru_3/N-\text{graphene}} - nE_{H_2O_2} \tag{4.4.1}$$

where $E_{n\text{oxo}}$, E_{H_2O}, and $E_{H_2O_2}$ are the energy of $n$oxo species, water and hydrogen peroxide respectively. The summary of formation energies of oxo species with various N-graphene models are given in Table 4.4. The formation energies of oxo species



(a) Ru_3/N-graphene (b) monooxo (c) trans-dioxo (d) cis-dioxo

Figure 4.6: Comparison of structures of Ru3/N-graphene and various oxo species. Oxygen atoms are coloured in red.

Table 4.4: The formation energies of oxo species with various N-graphene models. The values in bracket indicate the difference in the formation energies from its precursors. The structures of oxo species are shown in the Figure 4.7 and the energies are given in eV.

	SAC/N_4V_2		pyrrolic-N_4		pyridinic-N_4		C_2N	
$\Delta E_{\text{monooxo}}$	-2.97	(-2.97)	-3.99	(-3.99)	-4.10	(-4.10)	-4.20	(-4.20)
ΔE_{dioxo}	-4.91	(-1.93)	-7.96	(-3.98)	-7.75	(-3.65)	-7.98	(-3.78)
ΔE_{trioxo}			-11.56	(-3.60)	-12.35	(-4.60)	-10.95	(-2.97)
$\Delta E_{\text{tetraoxo}}$			-13.53	(-1.97)	-14.53	(-2.18)	-12.76	(-1.81)

with Ru_3 cluster models are generally higher than the formation energies of oxo species from the single atom catalyst analogue (SAC/N_4V_2). This is perhaps not surprising as the binding energies of the SCC themselves are also generally lower than that of SAC (Figure 4.1), hence they naturally tend to become reactive towards reactants. Comparing the formation energies among the cluster models, it can be seen that the first two formation energies ($\Delta E_{\text{monooxo}}$ and ΔE_{dioxo}) are broadly similar, but the models differ significantly in the formation of trioxo and tetraoxo species. In particular, the formation of the tri- and tetra-oxo forms of the pyrrolic-N_4 and C_2N models are significantly less favourable than the corresponding clusters on the pyridinic-N_4 model. This difference likely reflects the higher degree of charge transfer between the cluster and the surface in the pyrrolic-N_4 and C_2N models.

4.4.1 Density of states analysis

The DOS plots of oxo species on periodic N-graphene models (Figure 4.7) shows how the occupancy of the defect π states changes with the degree of oxidation of the Ru_3 cluster. The subplot 4.7a and 4.7b show the DOS plots of pyrrolic-N_4 defects and the supported Ru_3 cluster respectively. The shaded grey area shows the projected DOS of π orbitals of the N-graphene substrate, and the black line shows the sp^2-hybridised orbitals, including 4 dangling bonds of N atoms showing the localised character from 1.5 to 0.5 eV. The unoccupied broad π defect states lie just above the Fermi level, and are available to be coupled with the Ru_3 cluster orbitals, resulting in two bands of hybridised orbitals above and below the Fermi level in subplot 4.6a. These are reminiscent of the bonding and antibonding orbitals in the macrocyclic molecular model. It can also be seen that, while other Ru_3 orbitals have sharp peaks and localised characters, those hybridised with the π-defect states are remarkably

broadened to resonance states, a phenomenon that is described in the Anderson-Newns–Grimley model of chemisorption. The brief explanation of this model can be found in Appendix A. This is an indication that there is an interaction between the Ru_3 and N-graphene different from what is found in finite models, due to the delocalised character of π bands in the periodic model.

These renormalised bonding states remain completely occupied upon the oxidation of the Ru_3 cluster to monooxo and dioxo species, as seen in the subplot 4.7c and 4.7d. There is no sign of resonance states for the oxygen adsorbate in this case as the PDOS of oxygen p orbitals is characterised by very sharp peaks. However, when they are oxidised into trioxo and tetraoxo species (4.7e and 4.7f), the hybridised bonding bands on the Ru_3 cluster become partially unoccupied, indicating the charge that was transferred to the π defect states have been returned to the cluster. Indeed, the Bader charge analysis shows that the total charge of the Ru_3O_x cluster changes, becoming more positive upon the addition of oxo ligands, from +1.63 (naked) to +1.42 (monooxo), +1.41 (dioxo), +1.23 (trioxo), and +1.06 (tetraoxo). As a consequence, the stabilisation of electrons gained from the bonding-like interaction from broad π bands is being lost and the reaction energy of hydrogen peroxide decomposition or the formation energy of oxo species ($H_2O_2 + [Ru] \rightarrow H_2O + [Ru]O$) notably increases from the trioxo species. It is also worth noting that the broadening of oxygen PDOS starts to appear from this stage, indicating a coupling of oxygen with the π-bands. In a SAC study conducted by Liu *et al.* [150], it was shown that adsorption strength weakens as the centre of bands, that the adsorbate is coupling to, shift upwards relative to the Fermi level. The shift of renormalised π-bands here are also weakening the bonds between oxo ligands and the Ru_3 cluster.

Comparing to other periodic models of N-graphene, similar types of interactions between the Ru_3 cluster and the substrate can be observed (Figure 4.13, 4.14), hence the enhanced interactions in the periodic pyridinic-N_4 model. However, a significant distortion in the Ru_3 cluster is observed in the trioxo and tetraoxo species on the pyridinic-N_4 model. This difference between the pyrrolic and pyridinic-N_4 models could be related to the formal charges on the Ru_3 cluster.

4.5 Reaction Mechanism

We propose the reaction pathway for ethanol oxidation, starting from *trans*-dioxo species generated by the reaction with H_2O_2 (INT1), on the basis that there has been a previous study reporting catalytic performance of Ru_3O_2/rGO [137]. A complete

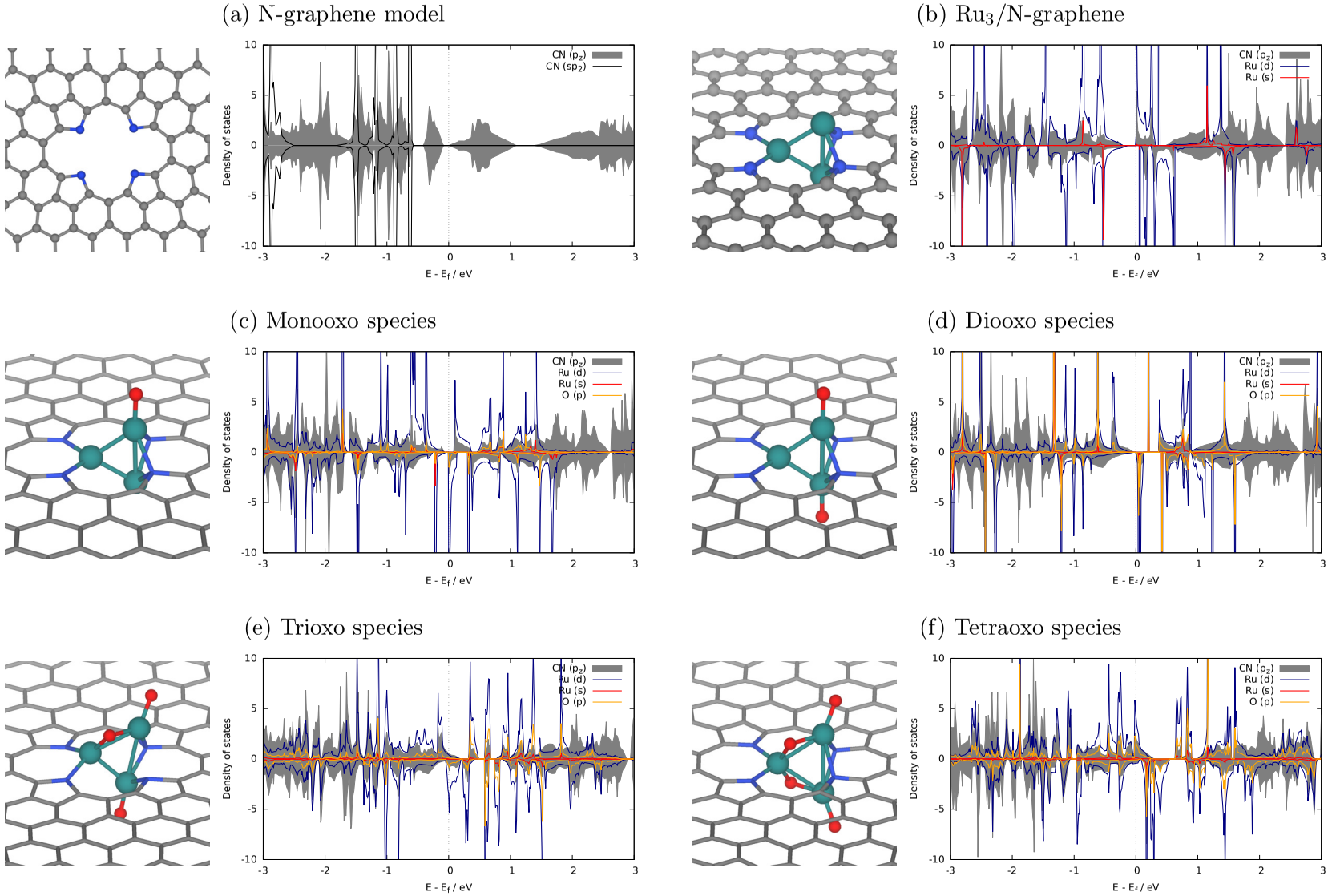


Figure 4.7: The geometries and projected DOS plots of oxo species embedded on pyrrolic-N_4. The energy is adjusted to the Fermi-energy and the positive and negative values of y-axis indicate DOS of up- and down-spin channels respectively.

reaction pathway is shown in Figure 4.8. In INT2 in the first half of the cycle (a), the alcohol group of adsorbate binds to the out-of-plane Ru atom. The oxo ligand attached to the same Ru atom moves its position to accommodate the alcohol group to the cis position, forming a five coordination environment for the Ru atom. From this point the hydrogen atom is transferred from the alcohol group to the neighbouring oxo ligand, creating hydroxyl and alkoxyl groups (INT3) attached in the same coordination. This process is a concerted proton-electron transfer (CPET) with a small activation barrier of 0.2 eV, and does not involve any radical intermediates. The resulting intermediate is the most energetically stable intermediate on this catalytic cycle, with energy difference of -0.54 eV from the previous one. The step to INT4, β-hydride elimination to form an η^2-aldehyde and a Ru-bound hydride, is the rate-determining step on the computed potential energy surface with an activation barrier of 0.82 eV, and the Ru atom (INT5) is now coordinated by 5 ligands. The reactant changes its hapticity from 2 to 1, preventing the hydride from returning to the β-C atom so that it remains terminally bonded to the same Ru atom (INT5). Release of the weakly-bound aldehyde leaves the cluster with hydroxide and the terminal hydride ligands that are now forming square-planar configuration on the Ru atom again (INT6). At this point, it is tempting to speculate that the cluster undergoes a reductive elimination step of water to generate a monooxo species, but the computed intermediate states are $\sim$0.5 eV higher than the starting intermediate (IS1). (The activation barrier for this transition state is found to be 0.85 eV, which is similar to the β-hydride elimination step.) Given that the temperature in the experiment is $\sim$350 K, we believe that this is not a realistic step in the catalytic cycle. Alternatively, one can speculate that these ligands remain intact and the cluster proceeds to catalyse the next alcohol dehydrogenation. In order for this to happen, the hydride can migrate from a terminal to a bridging position between the out-of-plane and in-plane Ru atoms freeing up a coordination site for binding of a second molecule of alcohol, with an activation barrier of 0.52 eV (INT7). The binding of a second reactant is significantly stronger than that in the first cycle, with an energy difference close to 1.0 eV (INT8) from the previous intermediate. This can then undergo a similar series of CPET (this time to a Ru-OH ligand, creating a water molecule) (INT9) and β-hydride elimination to generate the aldehyde, bound to a Ru_3 di-hydride (INT11) through an agostic intermediate (INT10). Again, the β-hydride elimination step is rate-determining but the activation barrier is 0.52 eV, lower than the previous cycle. Finally, after the release of the second aldehyde (INT13), reductive elimination of the two hydride ligands and release of H_2 can regenerate the monooxo species, which can

then be converted to the starting dioxo species (INT4). The two distinct pathways correspond to:

$$H_2O_2 + C_2H_5OH \rightarrow H_2O + CH_3CHO$$

$$H_2O_2 + 2C_2H_5OH \rightarrow H_2O + 2CH_3CHO + H_2$$

It is worth noting that if the reaction proceeds to the reductive elimination pathway, one equivalent of H_2O_2 can dehydrogenate only one equivalent of alcohol, while if the reaction continues with hydroxide, one equivalent of H_2O_2 can dehydrogenate two equivalents of alcohol.

This mechanism is similar to the dehydrogenation mechanism of double-atom catalysts proposed by Liu and co-workers Liu *et al.* [136], which also involves CPET to hydroxide ligand, β-hydride elimination and hydrogen evolution. One noticeable difference between our reaction mechanism and a widely-recognised mechanism of methane-to-methanol conversion by SACs is how a hydrogen atom is being transferred to the oxo species from the reactants. In the SAC, the reaction pathways are often radical based and require a hydrogen-atom transfer (HAT) steps in various pathways [151]. For example, the key step of methane activation is an abstraction of H atom on methane by an oxo ligand with homolytic C–H bond cleavage, that results in the formation of a free methyl radical and a hydroxyl ligand. In such reaction pathways with SACs, the HAT step is often the rate-determining step with the activation barriers varying from $\sim$0.6 to $\sim$1.6 eV . On the other hand, in our mechanism, the process of transferring hydrogen from alcohol group to the oxo or hydroxide ligand does not involve formations of radical intermediates. Throughout the reaction cycle until the reductive elimination of hydrogen molecule, the oxidation state of the cluster is unchanged. It is also worth noting that, because of the unsaturated structure inherent to the supported Ru_3 cluster, two ligands can be bound on the same atom which might facilitate the reactions. In the CPET step in Figure 4.8, the energy required to transfer the hydrogen atom is drastically reduced to $\sim$0.2 eV and only the β-hydride elimination is the rate-determining.

4.6 Conclusion and summary

In this Chapter, we have used DFT calculations to investigate the electronic structures and reactivity of Ru_3 clusters embedded on various models of N-graphene with different numbers and arrangements of nitrogen donors. Whilst the binding energies of Ru_3 clusters in all of our models are somewhat lower than that of SAC analogue,

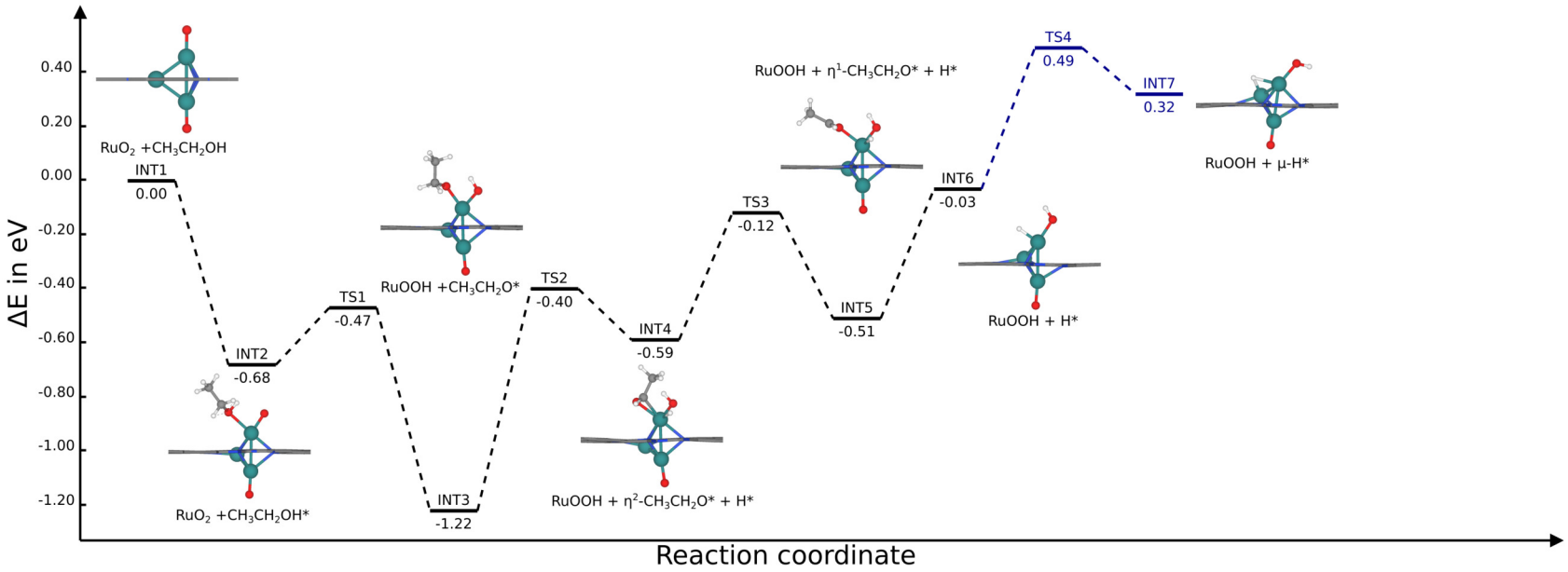


Figure 4.8: Dehydrogenation steps of ethanol with dioxo species

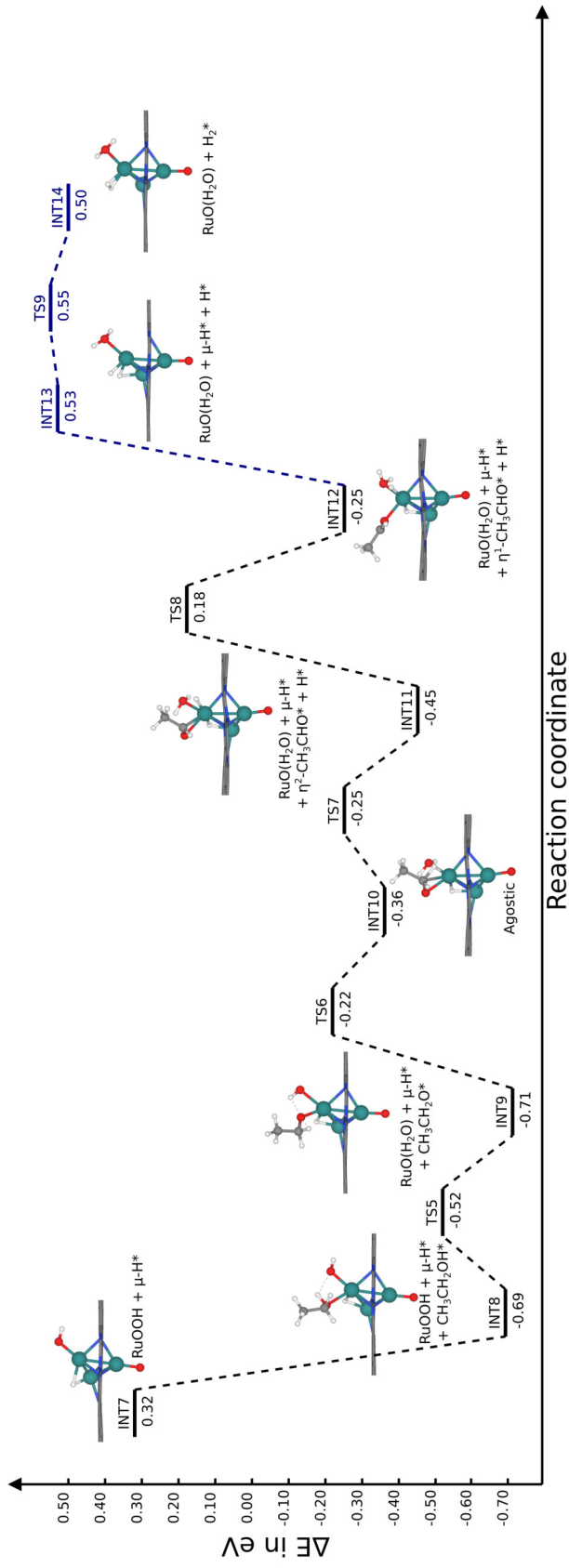


Figure 4.9: Dehydrogenation steps of of ethanol with dioxo species (continued)

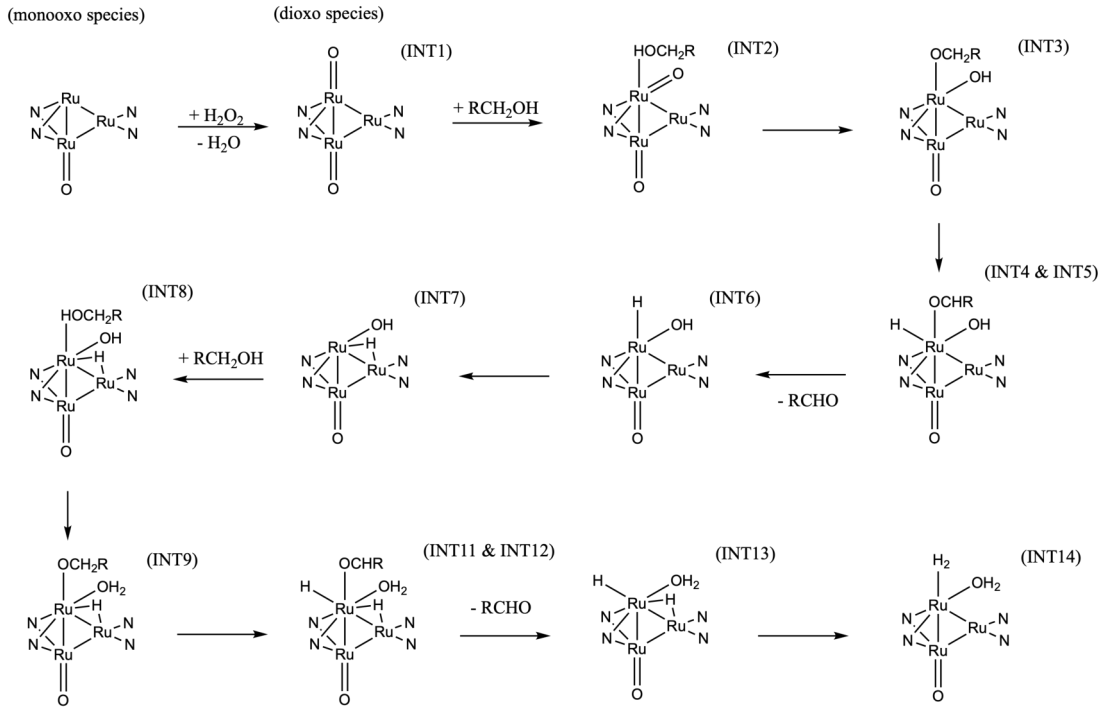
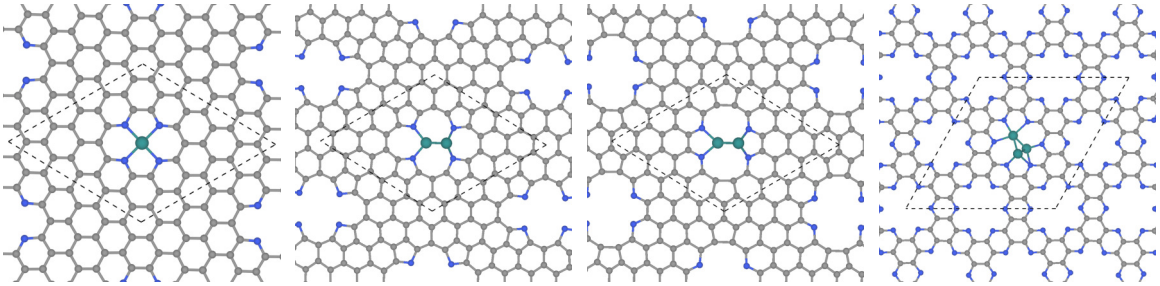


Figure 4.10: A schematic diagram summarising the proposed ethanol dehydrogenation by dioxo species as described in figure 4.8 and 4.9.

suggesting that they could be prone to aggregation, they are enhanced by the presence of pyrrolic-N_4 defects that stabilise the cluster. The orbital interaction analysis based on the finite models of N-graphene reveals that a significant charge transfer from the Ru_3 cluster to the macrocylic molecule occurs, giving the cluster a formal charge of 2+. This is also supported by the DOS analysis on the periodic models. The embedded clusters are reactive towards to formation of various oxo species in the presence of H_2O_2. We proposed a feasible reaction mechanism of ethanol dehydrogenation into formaldehyde with trans-dioxo species, in which β-hydride elimination is the rate-determining step.

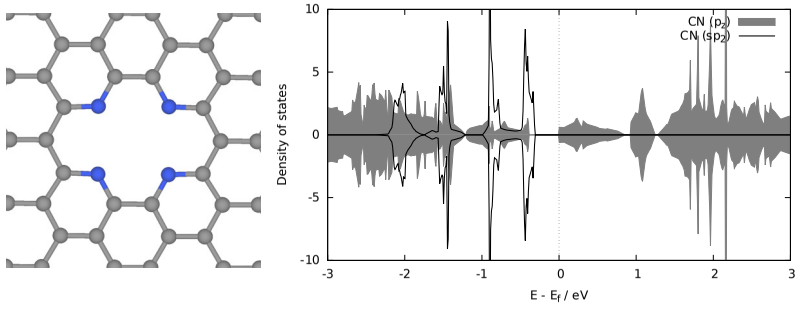
4.7 Supplementary materials

In this section, additional data and figures are presented to support the findings and assertions in the main body of this chapter. Figure 4.11 shows the smaller periodic models used to calculate DOS and potential energy surface. Figure 4.12, 4.13 and 4.14 show the the structures of oxo species derived from the models, and the projected DOSs for each species.

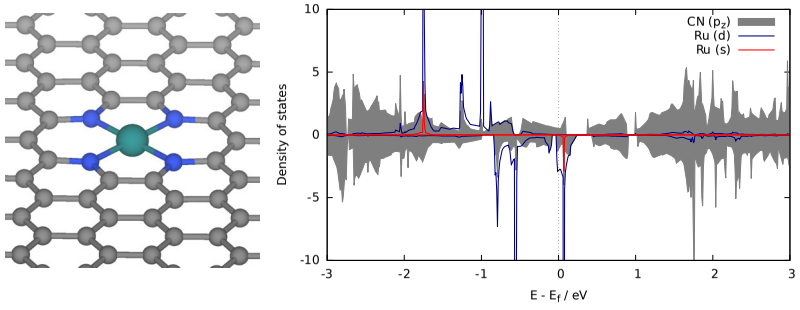


(a) SAC/N_4V_2 (b) pyrrolic-N_4 (c) pyridinic-N_4 (d) C_2N

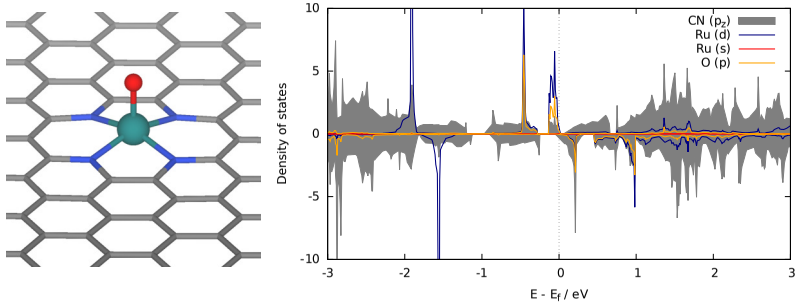
Figure 4.11: Top views of smaller periodic models used for DOS analysis and potential energy surface explorations. The dashed line indicate the boundary of the unit cell.



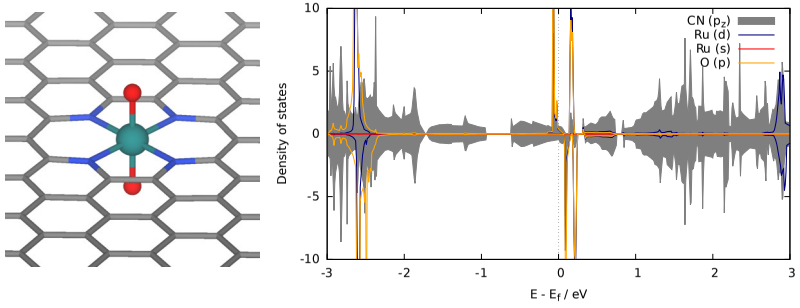
(a) N-graphene model



(b) Ru3/N-graphene

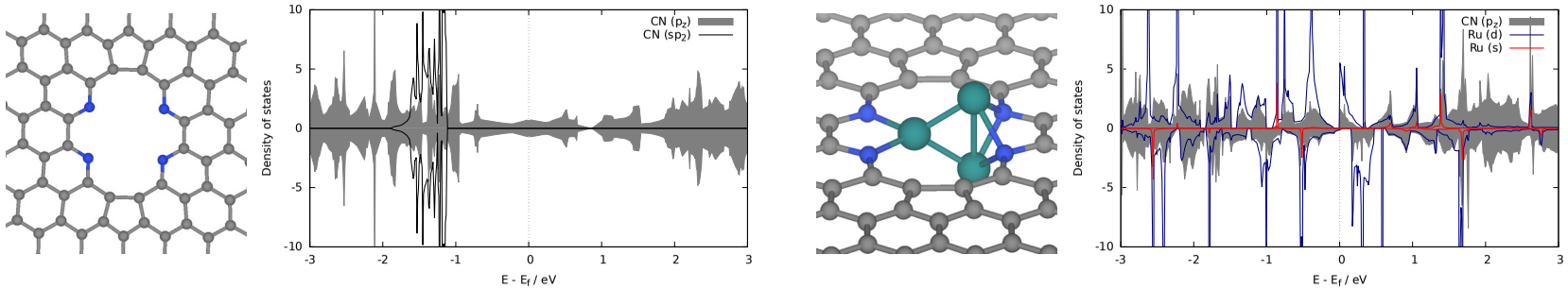


(c) Monooxo species



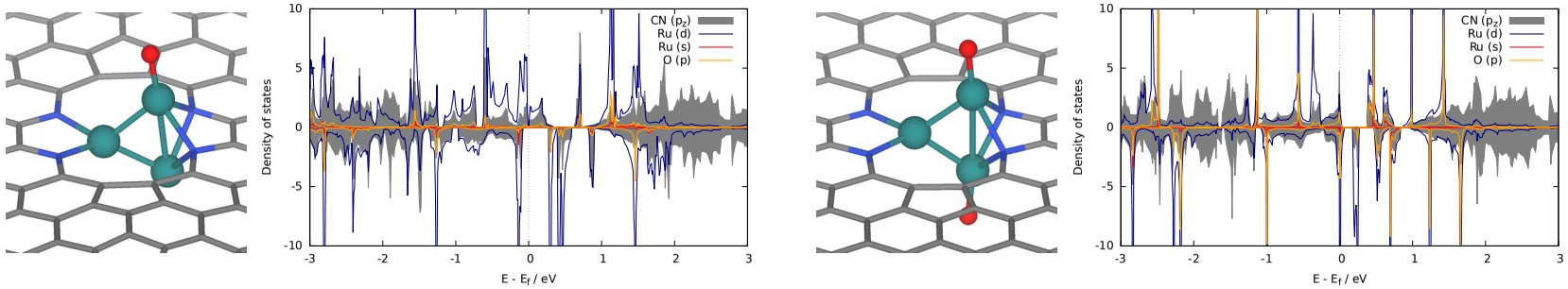
(d) Diooxo species

Figure 4.12: The geometries and projected DOS plots of SAC oxo species embedded on N_4V_2. The energy is adjusted to the Fermi-energy and the positive and negative values of y-axis indicate DOS of up- and down-spin channels respectively.



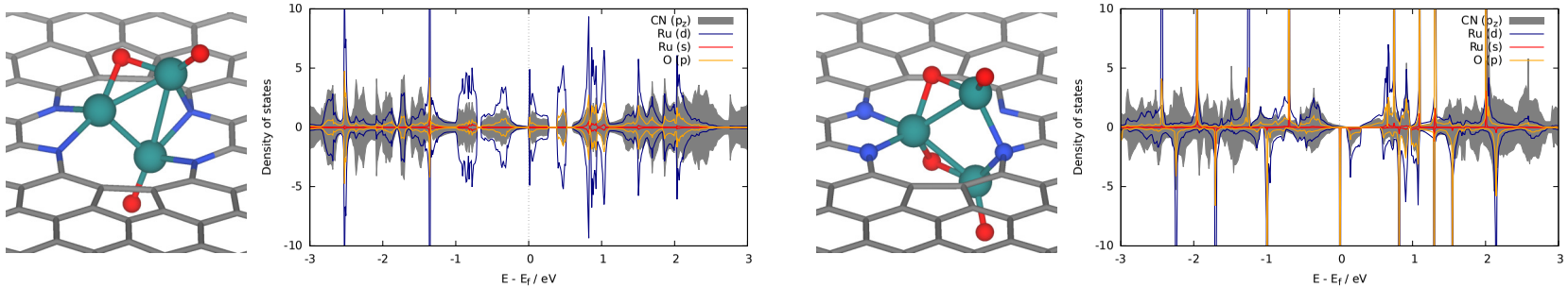
(a) N-graphene model

(b) Ru3/N-graphene



(c) Monooxo species

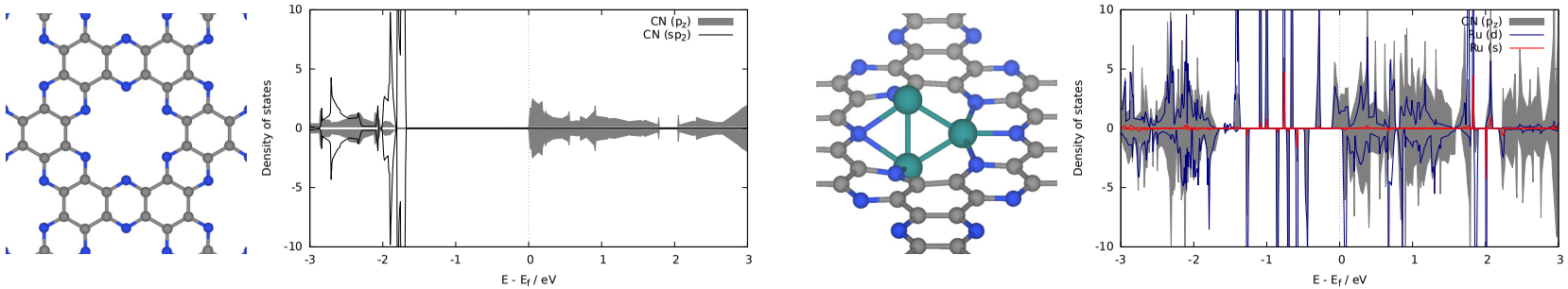
(d) Diooxo species



(e) Trioxo species

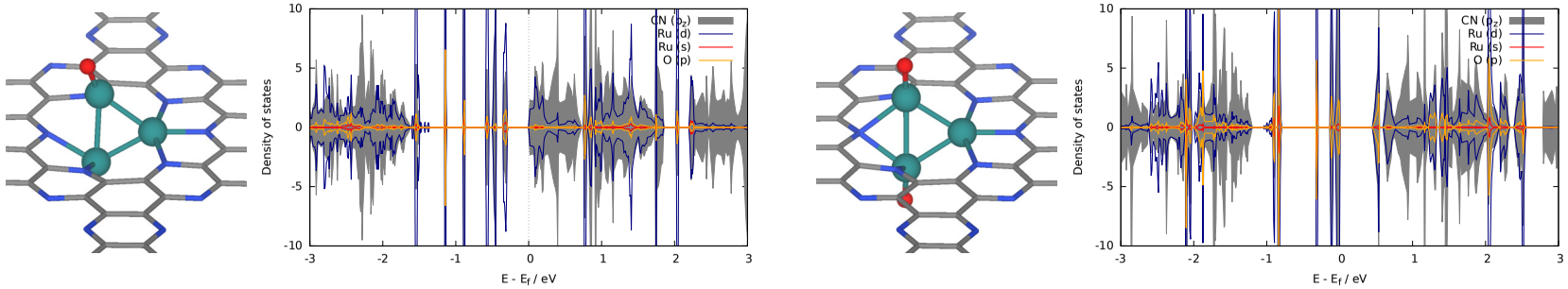
(f) Tetraoxo species

Figure 4.13: The geometries and projected DOS plots of oxo species embedded on pyridininc-N_4. The energy is adjusted to the Fermi-energy and the positive and negative values of y-axis indicate DOS of up- and down-spin channels respectively.



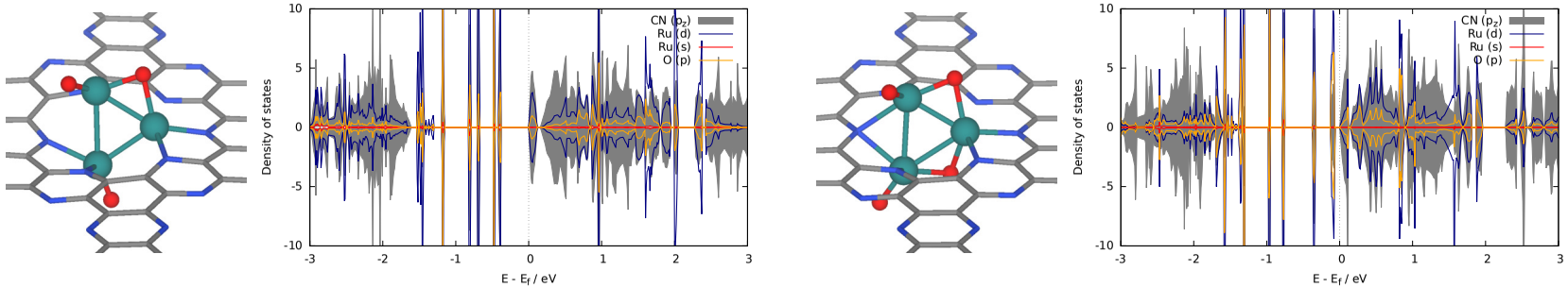
(a) N-graphene model

(b) Ru3/N-graphene



(c) Monooxo species

(d) Diooxo species



(e) Trioxo species

(f) Tetraoxo species

Figure 4.14: The geometries and projected DOS plots of oxo species embedded on graphitic-C_2N. The energy is adjusted to the Fermi-energy and the positive and negative values of y-axis indicate DOS of up- and down-spin channels respectively.

Chapter 5

Bond-stretch isomerism in Zintl phase Ba_3Ge_4 and Ba_3Si_4

5.1 Introduction

The chemistry of Zintl ions continue to present fascinating challenges to electronic structure theory [152–157]. Much of work in the field of computational cluster chemistry has focused on isolated Zintl ions, either in the gas phase or where the ions are separated in the solid state by encapsulated metal cations such as [M[2.2.2]crypt]$^+$. In the so-called Zintl phases, in contrast, the cations are not coordinated by additional ligands, leading to much smaller separations between the cluster anions [158–160]. Simple examples of this class would be the binary silicides and germanides such as NaSi,[161] which contains tetrahedral E_4^{4-} units, isoelectronic with P_4. The relatively close proximity of the cluster anions in these phases admits the possibility of forming additional inter-cluster bonds, and indeed in LiSi and LiGe, the very small Li^+ cation allows the tetrel atoms to form an extended network rather than a lattice of discrete E_4^{4-} tetrahedra [162]. Changes in temperature and/or pressure can also induce structural modifications, as for example in the report by Quesada-Cabrera *et al.* of the pressure-induce amorphisation of NaSi above 15 GPa, which may involve oxidative formation of Si-Si bonds between Si_4 clusters with concomitant reduction of Na^+ to Na [161]. The pressure dependence of phase equilibria for $BaGe_2$ have been studied extensively: at ambient pressure, $BaGe_2$ adopts the $BaSi_2$-type orthorhombic structure with discrete Si_4^{4-} tetrahedra,[163] but at high temperatures and pressures, the $ThSi_2$-type tetragonal lattice is preferred, where the Ge atoms are connected in a 3-dimensional network [164]. Similar behaviour emerges in the high-pressure (>10 GPa) metastable phases of $EuGe_3$ and $SrGe_6$ which have 2-, 3- and 4-connected networks, in marked contrast to the discrete tetrahedra in the MGe_2 decomposition

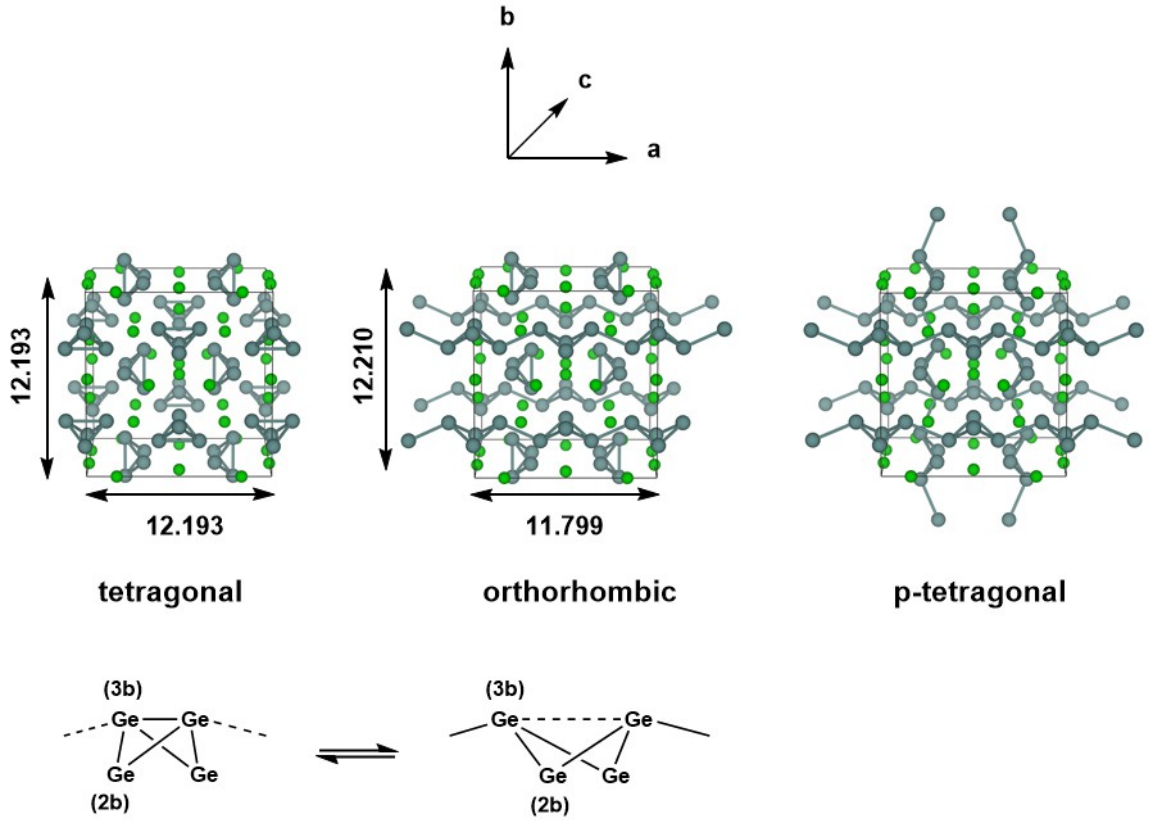


Figure 5.1: Optimised structures of the tetragonal, orthorhombic and (as yet unknown) p-tetragonal phases of Ba_3Ge_4 (all distances in Å.) The chains of Ge_4 units are polymerised along the a axis in the orthorhombic phase and along both a and b axes in the p-tetragonal phase. Ge and Ba atoms are coloured in grey and green respectively.

products found at ambient pressures [165, 166]. Wang *et al.* have explored the phase transition in $BaGe_2$ in some detail using density functional theory, establishing a pathway connecting the orthorhombic and tetragonal phases *via* flattening of the E_4 tetrahedra and inter-cluster bond formation, with a barrier of 0.35 eV/atom.[167]

In this chapter we focus on a more Ba-rich region of the Ba-Si/Ge phase diagrams,[168–170] and specifically on Ba_3Si_4 and Ba_3Ge_4, where E_4 units are again present but now in the more reduced 6- state (Figure 5.1). The silicon compound, first characterised by Eisenmann *et al.* in 1969,[171] has a tetragonal unit cell ($P4_2/mnm$) and contains isolated Si_4^{6-} units aligned in two orthogonal chains that run parallel to the crystallographic a and b axes. The Si_4 units adopt a butterfly-type structure where a single Si-Si bond of the tetrahedron has been cleaved, as might be anticipated for a Si_4^{6-} unit based on the Zintl-Klemm concept. There are two distinct atom types in each cluster: the atoms on the wing-tips of the butterfly that are bonded to two other

atoms and carry a formal 2- charge (denoted henceforth as '(2b)E' in Figure 1) and the atoms that constitute the body of the butterfly that are bonded to three others and carry a formal charge of 1- (denoted '(3b)E'). The (3b)Si-(3b)Si and (3b)Si-(2b)Si bond lengths were reported to be 2.29 Å and 2.34 Å, respectively, in Eisenmann's original paper, but a subsequent re-evaluation by Grin and co-workers in 2008 revised these to 2.4183(6)Å and 2.4254(3)Å, respectively.[172] The Ba_3Si_4 phase is weakly conducting (a 'bad metal'), consistent with band structure calculations that indicate substantial mixing of the Si 3*p* and the Ba 5*d* orbitals and a non-zero density of states at the Fermi level. Above 630 K, the Ge analogue Ba_3Ge_4 adopts a tetragonal phase that is structurally very similar to Ba_3Si_4, with isolated $Ge_4{}^{6-}$ anions and (3b)Ge-(3b)Ge and (3b)Ge-(2b)Ge bond lengths of 2.78 Å and 2.59 Å, respectively (the tetragonal phase).[173] Below this temperature, however, a first-order transition to a different, orthorhombic, phase where precisely half of the Ge_4 units polymerise along the crystallographic *a* axis (designated $(Ba^{2+})_6[Ge_4]^{6-}\ {}^1_\infty[Ge_4]^{6-}$), with concomitant cleavage of the intra-cluster (3b)Ge-(3b)Ge bond. The polymerisation causes a contraction along the *a* axis (11.799 Å *vs* 12.193 Å) and a ∼2% reduction in cell volume. The Zintl-Klemm concept is equally applicable to both tetragonal and orthorhombic phases: the (3b)Ge centres remain bonded to three others in both cases. The second chain of $Ge_4{}^{6-}$ units aligned along the crystallographic *b* axis is largely unaffected by the phase transition, as is the lattice parameter *b*. There has, as yet, been no evidence reported for a putative third phase (which we label p-tetragonal) where the Ge_4 units polymerise along both the *a* and *b* axes ($(Ba^{2+})_6\ {}^1_\infty[Ge_4]^{6-}\ {}^1_\infty[Ge_4]^{6-}$), a transition that would restore tetragonal symmetry. In a subsequent study, Pani and Palenzona reproduced the low-temperature orthorhombic phase of Ba_3Ge_4 but all attempts to isolate the high-temperature tetragonal phase by quenching of melts were unsuccessful.[168] A weak feature in the differential thermal analysis at ∼610 K was, however, consistent with the transition temperature identified by Zürcher and Nesper.[173]

A number of non-stoichiometric members of the $Ba_3Si_xGe_{4-x}$ series that interpolate between the limiting forms Ba_3Si_4 and Ba_3Ge_4 have also been characterised,[174] and the polymerised orthorhombic phase is found only for the most Ge-rich of these ($x = 0.3$ or lower). It is likely that the size of the cation plays an important role in determining the different behaviour of Si-rich and Ge-rich phases, just as it does in the LiSi/NaSi comparison made above: Nesper and co-workers have argued that whilst the Ba^{2+} cation is large enough to separate the $Si_4{}^{6-}$ anions to the point where inter-cluster bond formation is not possible, the more diffuse nature of the valence

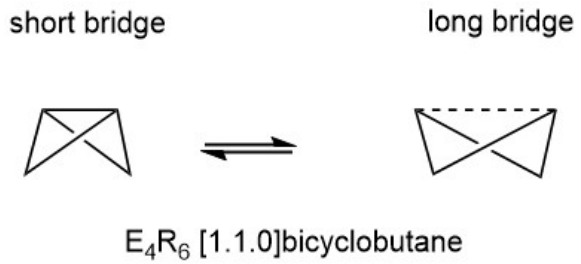


Figure 5.2: Bond-stretch isomerism in E_4R_6 clusters.

4*p* orbitals of Ge makes the polymerised orthorhombic phase accessible. [174] In that sense, the temperature-dependent behaviour of Ba_3Ge_4 mirrors many of the important features of the pressure-dependent behaviour of $BaGe_2$.

The E_4^{6-} unit is valence iso-electronic with the E_4R_6 bicyclo[1.1.0]butanes and heavier analogues which have been central to the discussion of bond-stretch isomerism, the phenomenon where two or more isomers differ primarily in the length of one or more bond.[175–183] In the case of the bicyclo[1.1.0]butanes, the two isomers in question differ in length of the E-E bond along the body of the butterfly, Figure 5.2 [184]. Schleyer and co-workers identified bond-stretch isomers of bicyclo[1.1.0]tetra-silane, Si_4H_6, that differ by almost 0.5 Å at the multi-configurational SCF level.[185] Calculations by Koch and co-workers using density functional theory (B3LYP functional) have explored both the impact of different substituents, R, and the switch from Si to Ge, on the potential energy surface [186, 187]. The heavier Ge_4R_6 analogues appear to favour the isomer with a long (3b)Ge-(3b)Ge bond, to the extent that the short isomer does not even constitute a stable minimum on the potential energy surface. The transition between the tetragonal and orthorhombic phases of Ba_3Ge_4 is clearly more complex in the sense that the cleavage of the intra-cluster Ge-Ge bond is intimately coupled to the polymerisation of the chain, but nevertheless the greater tendency to favour the 'long' isomer for Ge *vs* Si in a molecular context correlates with the apparently greater stability of the orthorhombic phase in Ba_3Ge_4 *vs* Ba_3Si_4. The relative stabilities of the two phases are, of course, determined by the free energy rather than the internal energy or enthalpy, and so the change in entropy is an important consideration. In a molecular context, we might anticipate that polymerisation results in a loss of entropy, but in a solid-state context there is no translational entropy to lose, in which case entropy changes are a more complex function of the phonon modes. With this in mind, we present here an analysis of the thermodynamics of the orthorhombic, tetragonal and p-tetragonal phases of Ba_3Si_4 and Ba_3Ge_4,

computed using density functional theory in conjunction with the quasi-harmonic approximation.

5.2 Computational methods

5.2.1 DFT details

The calculations reported in this project were performed using plane-wave pseudopotential density functional theory (DFT) with periodic boundary conditions (PBC) as implemented in the Vienna *ab initio* Software Package (VASP) [138–141, 188]. In the majority of cases, the exhange-correlation energy was modelled using the optB86b-vdW functional [57, 189], but we have also explored the impact on lattice parameters of alternative formulations including PBE [142], SCAN [190], and a combination of r^{2}SCAN [91] and the revised Vydrov–van Voorhis (rVV10) functional [191], r^{2}SCAN+rVV10 [192]. The valence electron configurations are $3s^2\,3p^2$ for Si, $3d^{10}\,4s^2\,4p^2$ for Ge and $5s^2\,5p^6\,6s^2$ for Ba, with the core electrons treated using PAW pseudopotentials. The plane-wave cut-off was set to 400 eV and the Brillouin zone was sampled on a $6 \times 6 \times 4$ Γ-centred grid. The structural relaxation was considered converged when the Hellmann-Feynman force fell below 10^{-4} eV Å^{-1}, while the SCF convergence criterion was 10^{-8} eV. Wannierisation was performed with a set of occupied valence bands using the Wannier90 software package [193]. The minimum energy path between the orthorhombic and tetragonal phases was determined using the solid-state climbing-image nudged elastic band (NEB) approach as implemented in the VTST code [143, 144, 194] with 8 intermediate images between the two phases and a spring constant of 5 eV Å^{-2} between images. The cell vectors of the primitive unit cells are allowed to be relaxed under the symmetry constraints of the orthorhombic phase. The total force was minimised to below 10^{-3} eV Å^{-1} in these calculations. The convergence of energies with respect to the plane-wave cut-off energy has been tested to ensure that Pulay stress is minimal and the calculated energy-volume curves are smooth.

5.2.2 Phonon properties

Harmonic phonon dispersion curves are calculated using the finite difference method implemented in Phonopy [61, 195]. Force constants were computed with atomic displacements of 0.01 Å in a supercell described by the following transformation of

the primitive vectors (Table 5.1 and 5.2):

$$S = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

k-point sampling was performed on a Γ-centered $4 \times 4 \times 4$ grid. Phonon DOS curves and thermodynamic properties were then calculated by interpolation onto a $12 \times 12 \times 8$ q-point grid applied to the primitive cell. Quasi-harmonic calculations were carried out by performing phonon calculations on optimised structures with up to 6% compression and 8% expansion of the cell volume relative to the equilibrium structures, or until imaginary phonons emerge, at which point the structure becomes dynamically unstable. The Helmholtz free energy, $F(T, V)$, the sum of the electronic energy, $U_0(V)$ and the phonon free energy, $F_{\mathrm{ph}}(T, V)$, for a given temperature is then fitted to a third-order Birch-Murnaghan equation of state:[196]

$$\begin{aligned} F(T,V) &= U_0(V) + F_{\mathrm{ph}}(T,V) \\ &= F_0 + \frac{9V_0B_0}{16}\left\{\left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^3 B_0' + \left[\left(\frac{V_0}{V}\right)^{\frac{2}{3}} - 1\right]^2 \left[6 - 4\left(\frac{V_0}{V}\right)^{\frac{2}{3}}\right]\right\} \end{aligned} \tag{5.2.1}$$

where V_0 and B_0 are the reference volume and bulk modulus, respectively, and B_0' is the derivative of the bulk modulus with respect to pressure. The Gibbs free energy of a phase for a given temperature and pressure is then obtained by finding the volume that minimises the sum of the lattice internal energy $U_0(V)$, the harmonic phonon Helmholtz free energy $F_{\mathrm{ph}}(T, V)$ and the PV term, as given in equation 2.7.23. The ranges of temperatures and pressures considered were 0–1200 K and 0–1.5 GPa for orthorhombic and tetragonal phases and 0–300 K and 0–1.5 GPa for the p-tetragonal phase: the lower limit for the latter reflects the emergence of dynamic instability at lower temperatures.

5.3 Results and discussions

5.3.1 Potential energy surfaces for Ba_3Si_4 and Ba_3Ge_4

Optimised structural parameters at 0 K for the three phases of interest in Ba_3Ge_4, the orthorhombic and tetragonal phases and the third, as-yet unknown, p-tetragonal phase where all of the Ge_4 units are polymerised in both orthogonal directions, are summarised in Table 5.1, where the available experimental data are also shown for

comparison. With the optBP86b-vdW functional, the optimised lattice parameters are broadly consistent with the crystallographic data, most strikingly in the orthorhombic phase of Ba_3Ge_4, where the a, b and c lattice parameters are within 0.01 Å of experiment. The experimentally-observed ∼0.2 Å contraction of a in the orthorhombic phase is reproduced in all cases, as is the ∼ 2% decrease in unit cell volume. This trend continues to the p-tetragonal phase, where a further 1.1% reduction in volume is associated with the polymerisation of the second chain of Ge_4 units along the b axis. The relative energies are also consistent with experiment in so much as the orthorhombic phase is predicted to be most stable at 0 K, lying 0.027 eV/f.u. (formula unit) below the tetragonal alternative which is, as noted in the introduction, stable only above 630 K. The energy of the p-tetragonal phase, which has not been observed under any conditions, is intermediate between orthorhombic and tetragonal, lying 0.014 eV/f.u. above the former. These relative energies are, however, somewhat sensitive to choice of functional, and all of PBE, SCAN and r^{2}SCAN+rVV10 predict a relative stabilisation of the tetragonal phase relative to orthorhombic, to the extent that the two phases are almost iso-energetic with PBE and tetragonal phase is predicted to be more stable for both SCAN and r^{2}SCAN+rVV10. The p-tetragonal phase is not predicted to be the most stable for any of the chosen functionals. PBE predicts lattice parameters that are systematically ∼0.1 Å longer than optBP86b-vdW while the SCAN-type functionals also overestimate lattice parameters, but to a lesser degree. Very similar patterns are observed in the data for Ba_3Si_4, although in this case the tetragonal phase is systematically stabilised by ∼0.06 eV/f.u. relative to orthorhombic, as a result of which the tetragonal phase is the most stable for all functionals tested in this work. The structural differences between the tetragonal and orthorhombic phases also follow the experimental trend, with polymerisation of the Si_4 chains along a causing a 2.5% decrease in cell volume. All attempts to locate a local energy minimum for the p-tetragonal phase of Ba_3Si_4 relaxed instead on the experimentally characterised tetragonal alternative.

The potential energy surfaces connecting the tetragonal and orthorhombic phases of Ba_3Ge_4 and Ba_3Si_4, computed using the nudged elastic band method and the optBP86b-vdW functional, are shown in Figure 5.3. For Ba_3Si_4, a transition state separates the tetragonal phase from its less stable orthorhombic counterpart by a barrier of ∼0.05 eV/f.u. The transition structure is confirmed to be a first-order saddle point through the presence of a single imaginary phonon mode at the Γ point, and has a geometry intermediate between the two phases. There is significant elongation of the intra-cluster bonds (2.88 Å *vs* 2.44 Å in orthorhombic phase) and concomitant

Table 5.1: Summary of optimised energies and structural parameters for the orthorhombic, tetragonal and p-tetragonal phases of Ba_3Ge_4. d1 is the intra-cluster (b3)Ge-(b3)Ge distance while d2 is the inter-cluster (b3)Ge-(b3)Ge distance . Distances are given in Å, volumes in Å^3/formula unit (f.u.) and energies in eV/f.u. Note that the a (=b) parameters for the tetragonal cells (the tetragonal and p-tetragonal phases) are $\sqrt{2} \times \sqrt{2}$ expansions of the primitive unit cell. We were unable to converge the geometry of orthorhombic phase with SCAN as the optimiser failed to reach the convergence criterion even with an extended number of steps.

		Ba_3Ge_4						
		a	b	c	V/f.u.	$d1$	$d2$	ΔE/f.u.
X-ray	tetragonal	12.193		12.032	223.6	2.71	3.63	
	orthorhombic	11.799	12.210	12.097	217.8	3.27	2.87	
	p-tetragonal			no data				
optB86b-vdW	tetragonal	12.093		12.064	220.5	2.65	3.63	0
	orthorhombic	11.799	12.201	12.096	216.8	3.29	2.86	-0.027
	p-tetragonal	11.892		12.129	214.4	3.16	3.03	-0.013
PBE	tetragonal	12.250		12.215	229.1	2.64	3.72	0
	orthorhombic	11.909	12.321	12.252	224.7	3.30	2.91	-0.003
	p-tetragonal	12.019		12.295	222.0	3.14	3.10	0.027
SCAN	tetragonal	12.161		12.123	224.1	2.61	3.69	0
	orthorhombic	11.822	12.261	12.144	220.0	3.24	2.90	0.013
	p-tetragonal			not converged				
r^{2}SCAN+rVV10	tetragonal	12.175		12.071	223.6	2.57	3.75	0
	orthorhombic	11.799	12.250	12.113	218.8	3.27	2.87	0.007
	p-tetragonal	11.911		12.185	216.1	3.13	3.05	0.060

Table 5.2: Summary of optimised energies and structural parameters for the orthorhombic, tetragonal and p-tetragonal phases of Ba_3Si_4. We were unable to converge on p-tetragonal-like structures for Ba_3Si_4 – all attempts reverted instead to the more stable tetragonal phase.

		Ba_3Si_4						
		a	b	c	V/f.u.	$d1$	$d2$	ΔE/f.u.
X-ray	tetragonal	12.053		11.832	214.9	2.42	3.81	
	orthorhombic			no data				
	p-tetragonal			no data				
optB86b-vdW	tetragonal	12.020		11.857	214.2	2.44	3.77	0
	orthorhombic	11.507	12.236	11.845	208.5	3.37	2.63	0.022
PBE	tetragonal	12.135		11.959	220.2	2.43	3.83	0
	orthorhombic	11.562	12.393	11.937	213.8	3.43	2.61	0.053
SCAN	tetragonal	12.157		11.901	219.8	2.39	3.88	0
	orthorhombic	11.553	12.317	11.977	213.0	3.36	2.54	0.091
r^2SCAN+rVV10	tetragonal	12.122		11.889	218.4	2.40	3.85	0
	orthorhombic	11.529	12.360	11.883	211.7	3.42	2.60	0.090

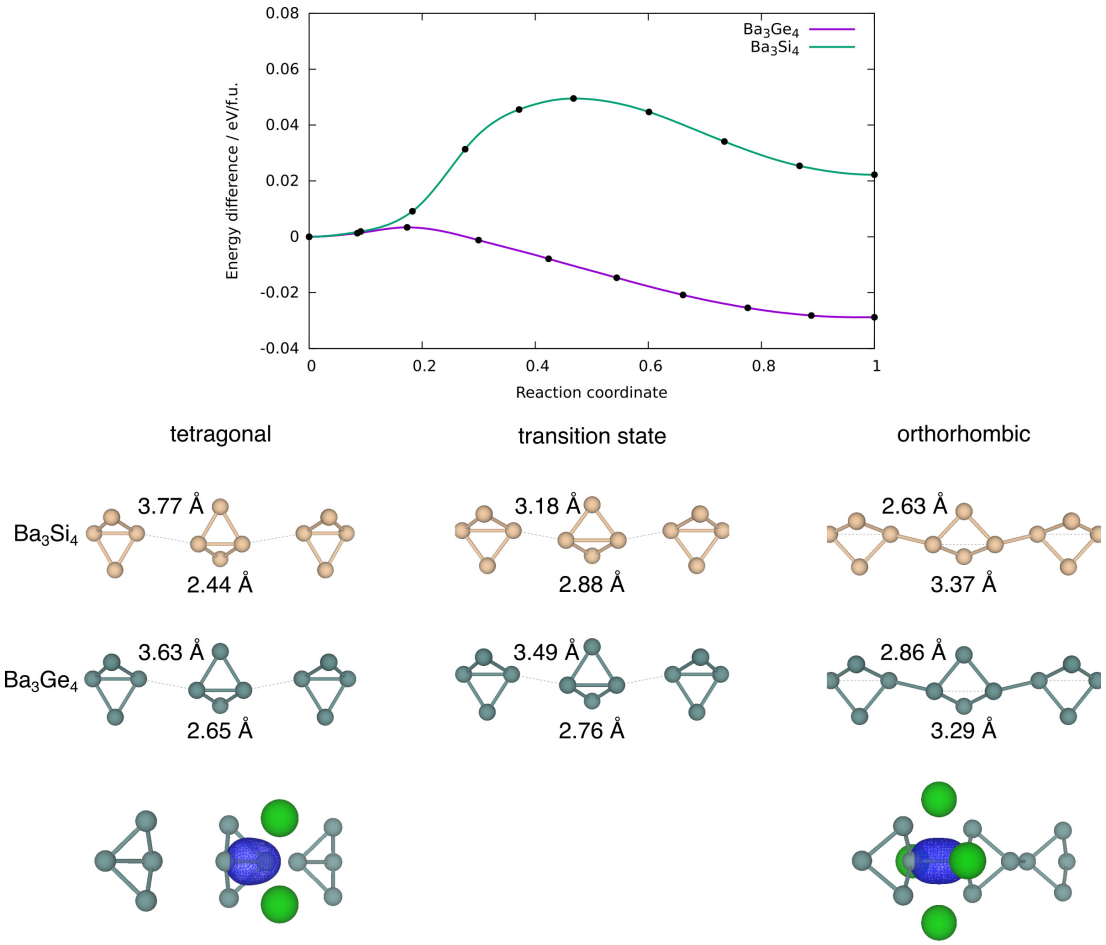


Figure 5.3: Potential energy curve (optB86b-vdW) for the concerted phase transition from the tetragonal to the orthorhombic phase in Ba_3Ge_4 and Ba_3Si_4. Maximally localised Wannier functions corresponding to the intra- and inter-molecular bonds of the tetragonal and orthorhombic phases, respectively, are also shown. Si atoms are coloured in beige.

contraction of the inter-cluster bonds (3.18 Å *vs* 3.77 Å in orthorhombic phase). The correspondence between the intra-cluster Si-Si bond length of 2.88 Å at the transition structure and the value of 2.86 Å reported by Koch *et al.* in the 'long' bond-stretch isomers in molecular Si_4R_6 is quite striking [186]. It appears, then, that the Si_4^{6-} units are indeed effectively isolated in Ba_3Si_4, and cleavage of the intra-molecular Si-Si is almost complete at the transition state. For the Ba_3Ge_4 system (purple line in Figure 5.3), the reaction is almost barrier-less, with the transition structure only 0.003 eV/f.u. above the tetragonal phase. Consistent with the exothermicity of the forward reaction, the transition structure shown in Figure 5.3(b) is 'early', with a marginally elongated Ge-Ge bond along the body of the butterfly (2.76 Å *vs* 2.65 Å in the tetragonal phase), and very marginally shortened intra-cluster distance of 3.49 Å *vs* 3.63 Å in the tetragonal phase. Even for the Ba_3Si_4 case, the barrier to rearrangement is an order of magnitude smaller than those reported by Wang *et al.* for the orthorhombic - tetragonal phase transition in $BaGe_2$,[167] the difference probably reflecting the relative weakness of the bonds in the Si_4^{6-} and Ge_4^{6-} tetrahedra, and also the fact that the phase transitions in Ba_3E_4 involve only the cleavage/formation of E-E bonds along a single axis rather than the complete collapse of the tetrahedral units into a 3-dimensional network. The evolution of the electronic structure across the potential energy surface has been discussed extensively by Nesper and co-workers [173, 174], and our analysis of the density of states is fully consistent with those reported previously. The making and breaking of Ge-Ge bonds can be tracked through the maximally-localised Wannier functions (MLWFs) shown at the bottom of Figure 5.3, which show the intra-cluster σ-bond in the tetragonal phase and the inter-cluster σ-bond in the orthorhombic phase. These MLWFs show delocalisation tails on the Ba^{2+} ions surrounding the cluster, consistent with Nesper and co-workers' observation that the Ba 5*d* levels are substantially populated in Ba_3Si_4 [173]. An analysis of the Bader charges, shown in supporting information, Table S1, indicates that polymerisation induces no significant changes in the charges of the ions.

5.3.2 Phonon modes and thermodynamic properties at finite temperature

Given the rather delicate energetic balance between the orthorhombic and tetragonal phases of Ba_3Ge_4 in Table 5.1, we now consider the relative Gibbs free energies which requires the computation of the vibrational entropy. To do this, we use the quasi-harmonic approximation (QHA) [197] as implemented in phonopy to allow for volume relaxation at finite temperature. The relative free energies (computed with the

Table 5.3: Bader charge analysis of Ba_3Ge_4 and Ba_3Si_4 (optB86b-vdW).

	Ba_3Ge_4						Ba_3Si_4			
atom	tetragonal		orthorhombic		p-tetragonal		tetragonal		orthorhombic	
Ba	4e	1.09	4k	1.08	4e	1.10	4e	1.10	4k	1.07
			4l	1.11					4l	1.11
	4d	1.14	8m	1.11	4d	1.10	4d	1.14	8m	1.10
	4g	1.15	4g	1.11	4g	1.11	4g	1.16	4g	1.11
			4j	1.13					4j	1.12
Ge,Si	8i	-0.67	8q	-0.67	8i	-0.67	8i	-0.69	8q	-0.72
			8p	-0.67					8p	-0.70
	8j	-1.02	8n	-0.96	8j	-0.98	8j	-1.01	8n	-0.86
			8o	-1.03					8o	-1.03

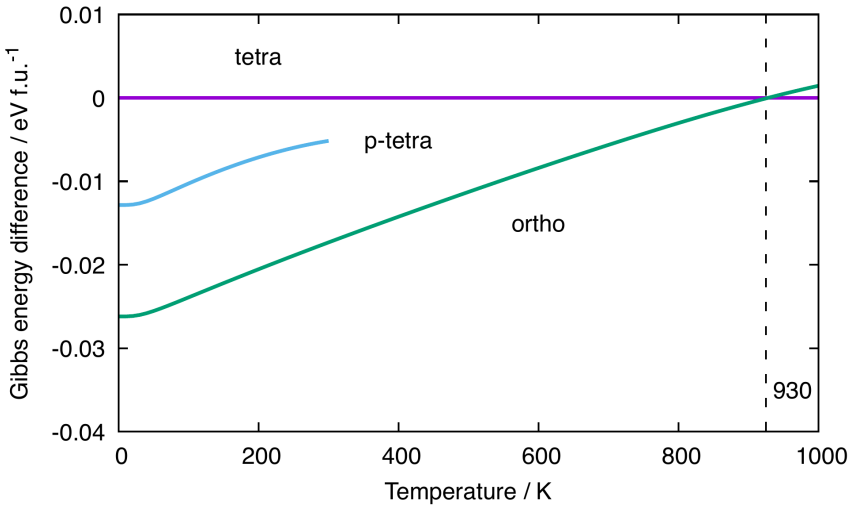


Figure 5.4: Relative Gibbs free energies (optB86b-vdW) of the different phases of Ba_3Ge_4. The reference value is set to the value of the tetragonal phase.

optB86b-vdW functional) of the orthorhombic and tetragonal phases as a function of temperature are shown in green and magenta in Figure 5.4(a), while the p-tetragonal phase is shown in cyan. The tetragonal phase is chosen as the zero of energy. Note that the values at 0 K and correspond to $U_0(V)$, the data points in Table 5.1, with the stability of the phases decreasing in the order orthorhombic < p-tetragonal < tetragonal. The free energy of the p-tetragonal phase is truncated at 300 K because beyond this temperature the minimum energy point on the Helmholtz free energy surface corresponds to a structure with very low or imaginary harmonic frequencies that render it dynamically unstable. As the temperature is raised, the orthorhombic phase is destabilised relative to tetragonal, leading, ultimately, to a cross-over at ∼930 K,

beyond which the tetragonal phase becomes thermodynamically stable. This value is 300 K higher than the experimental value of ∼630 K, implying an over-stabilisation of the low-temperature (α) phase. We can identify a number of possible reasons for the discrepancy between the experimental and computed transition temperatures, the most obvious being the strong functional dependence of the electronic energies, $U_0(V)$, shown in Table 5.1. The values of $U_0(V)$ correspond to the intercepts in Figure 5.4(a), and so a relative stabilisation of the β phase at 0 K, as observed in all functionals other than optB86b-vdW, would necessarily raise the intercepts and therefore depress the computed transition temperature. Deringer *et al.* have identified the density functional as a potential source of error in the calculated transition temperature for the orthorhombic-cubic phase transition in GeSe.[198] Another potential source of error is the approximate treatment of anharmonicity offered by the QHA approach. Several previous studies using the QHA have noted the overestimation of thermal expansion coefficients and lattice constants at elevated temperatures [199, 200]. In such circumstances the harmonic phonon free energy, $F_{\mathrm{ph}}(T, V)$, often becomes markedly non-linear as a function of T due to the presence of soft phonons at expanded volumes, causing a divergence in the thermal expansion coefficients (see Figure 5.5) and inaccuracies in the calculated thermodynamic parameters. The anharmonicity can, in principle, be explicitly included using computationally intensive methods such as self-consistent phonon theory (SCPH), which has been shown to improve the wurtzite to rocksalt phase boundary in GaN [202–204]. In another recent study of phase transitions in SnS and SnSe, Pallikara and Skelton noted that a ∼350 K over-estimation of transition temperatures between *Pnma* and *Cmcm* phases of that could be corrected, albeit by only ∼20 K, by renormalising selected imaginary modes by numerical solution of the Schrödinger equation on the corresponding 1-dimensional potential energy surface [205]. The dynamic instability of the p-tetragonal phase above 300 K prevents an exact determination of the crossover temperature with the tetragonal phase, but by extrapolation of the low-volume region [198] we can be confident that there is no point in the ambient-pressure phase diagram where it is the global minimum. The root cause of the temperature dependence in Figure 5.4 is the much lower intrinsic entropy for the orthorhombic and p-tetragonal phases of Ba_3Ge_4 compared to tetragonal, summarised in Table 5.4, where the tetragonal phase is again taken as the reference value. The total change in harmonic phonon entropy, $\Delta S_{\mathrm{total}}$, for the transition from orthorhombic to tetragonal phases is -0.042 meV K^{-1} $\mathrm{f.u.}^{-1}$ (∼0.5 k_B, within the usual range of 0-5 k_B for phase transitions) [204]. For the orthorhombic to p-tetragonal transition it is even lower at -0.054 meV K^{-1} $\mathrm{f.u.}^{-1}$.

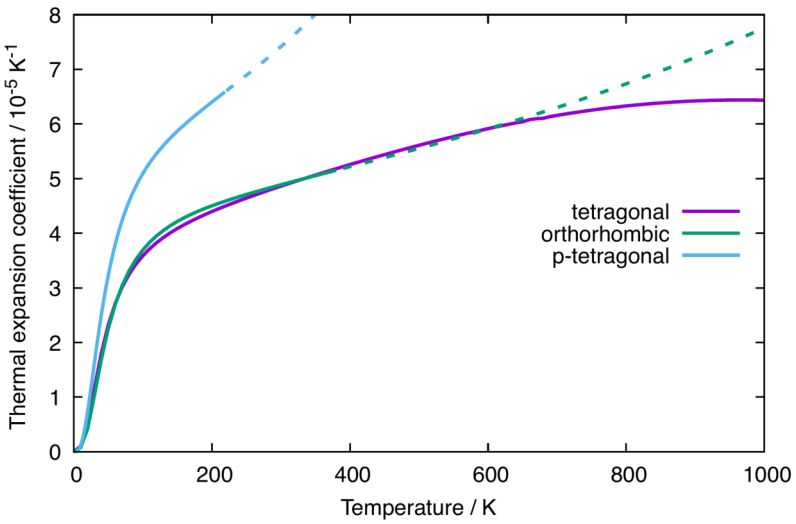


Figure 5.5: Comparison of temperature dependencies of thermal expansion coefficients of Ba_3Ge_4 at ambient pressure. The transition from full to dashed lines indicates the point where when the $[d^2\alpha/dT^2]_p \geq 0$. Beyond this point, the validity of the QHA is uncertain according to criteria set out by Wentzcovitch *et al.* [200, 201]. The thermal expansion coefficients for the p-tetragonal phase diverge markedly at ∼400 K and the Gibbs free energy is therefore truncated beyond 300 K in this study. The orthorhombic phase shows only minor signs of divergence above ∼1000 K and $[d^2\alpha/dT^2]_p < 0$ for the the tetragonal phase at all temperatures below the melting point.

The trends in ΔS_{total} summarised in Table 5.4 are at least qualitatively consistent with the idea that an increasing degree of polymerisation of the Ge_4 chains should lead to a loss in entropy. However, as we noted in the introduction, there is no loss of translational entropy associated with the linking of the clusters in the solid state, as there would be in a molecular version of polymerisation, and the formation of inter-cluster Ge-Ge bonds in the orthorhombic phase is in any case offset, at least partially, by the cleavage of the intra-molecular bonds. The vibrational entropy, and therefore the free energy, is in fact a complex function of all the phonon modes of the lattice. Figure 5.6 shows the phonon dispersion curves and phonon density of states for the orthorhombic, tetragonal and p-tetragonal phases of Ba_3Ge_4. The density of states plots are projected onto the Ge atoms (red) and the Ba^{2+} ions (grey). Each $Ge_4{}^{6-}$ unit has $3N - 6 = 6$ vibrational modes, of which 5 correspond to Ge-Ge stretches and the sixth to a wagging motion of the wingtips of the Ge_4 'butterfly'.

These six vibrational modes are illustrated in the top panel of Figure 5.6 for the isolated $Ge_4{}^{6-}$ cluster: they range in frequency from 7.63 THz (a stretching mode of the (3b)Ge-(3b)Ge bond) to 2.59 THz (the wagging motion of the wing-tips noted above). The corresponding frequencies for the $Si_4{}^{6-}$ cluster, shown in Figure 5.7, are higher, reflecting the stronger Si-Si bonds and the lower mass of Si. The phonon modes of the Zintl phases can be separated into those involving relative motion of the cations and anions in the lattice, found at low frequencies, and those related to the six internal vibrational modes of each $Ge_4{}^{6-}$ unit. There are four Ge_4 units per unit cell, and so each of the fundamental modes generates four linear combinations in the dispersion curve, giving 24 modes in total. The high-frequency regions of the phonon dispersion curves for both phases are dominated by the five linear combinations of the Ge-Ge stretches, $3a_1$, $1b_2$, $1b_1$, $1a_2$ and $2a_1$, shown in red. In the tetragonal phase, these internal modes appear around 7.0 THz, 6.3 THz, 5.2 THz, 5.0 THz and 4-5 THz, respectively, and the marked dispersion between 4 and 5 THz reflects the coupling between the $2a_1$ modes where the largest amplitude of motion is aligned along the directions of the Ge_4 chains. The wagging mode lies in the lower frequency region, where it is strongly coupled to motions of the Ba^{2+} ions.

The impact of polymerisation to form the orthorhombic phase is found primarily in the 4-5 THz region where dispersion of the $2a_1$ fundamental mode is now seen only in two of the four linear combinations, corresponding to the vibrations of the unpolymerised $Ge_4{}^{6-}$ clusters along the b axis. The $2a_1$ fundamentals of the polymerised chain along a instead form a narrow band around 4.2 THz. Similarly, the two linear combinations of the $3a_1$ modes of the unpolymerised chain remain at 7.0 THz, but

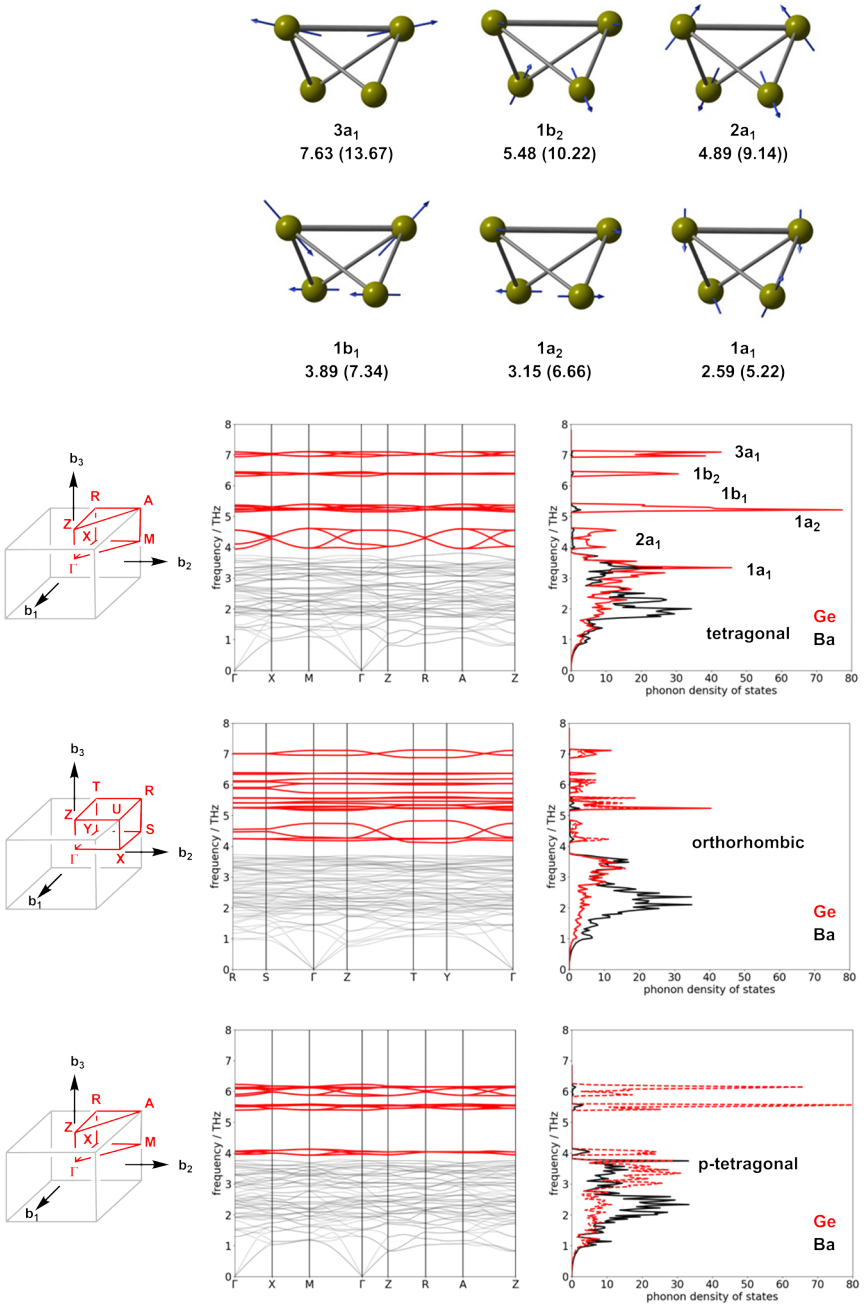


Figure 5.6: Phonon dispersion curves and densities of states for the tetragonal, orthorhombic and p-tetragonal phases of Ba_3Ge_4. The density of phonon states is projected onto the unpolymerised (full red line) and polymerised (dashed red line) chains of Ge_4 units. Vibrational modes of the isolated C_{2v}-symmetric Ge_4^{6-} anion, computed with density functional theory with localised Slater-type basis (ADF package, PBE functional, high-dielectric COSMO solvent model ($\epsilon = \infty$) used to approximate the confining effect of the cation lattice), are shown for comparison. Note that this isolated model is a crude approximation and the order of vibrational modes may be different. Frequencies are given in THz, values for Si_4^{6-} are shown in brackets.

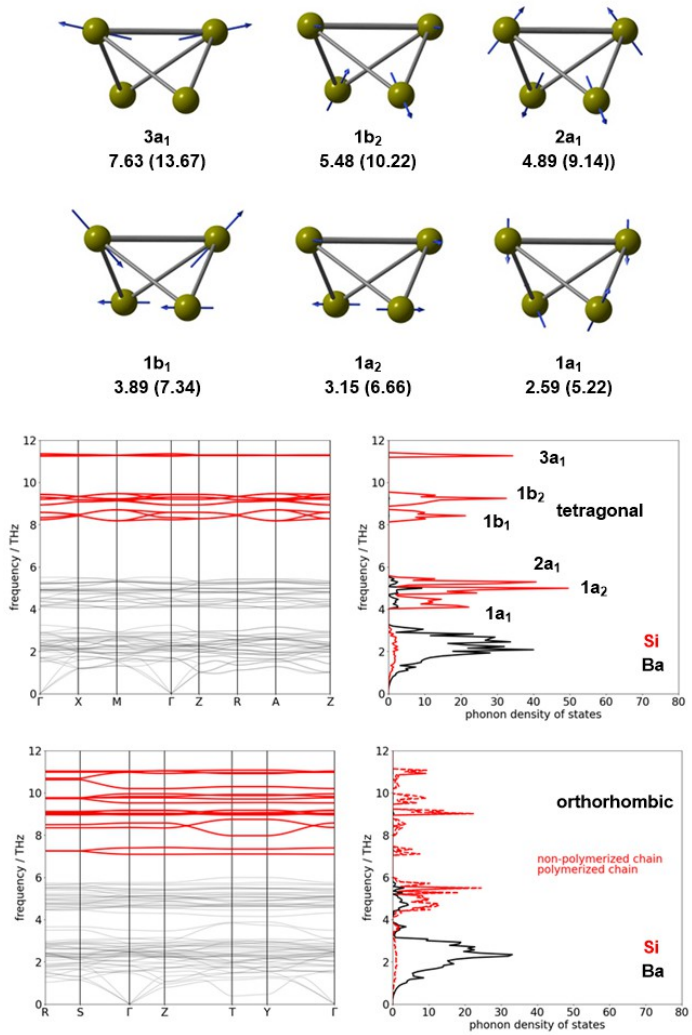


Figure 5.7: Phonon dispersion curves and densities of states for the tetragonal, orthorhombic and p-tetragonal phases of Ba_3Si_4.

Table 5.4: Harmonic phonon contributions to entropy changes, ΔS between the tetragonal β phase of Ba_3Ge_4 and the alternative orthorhombic (α) and p-tetragonal phases, calculated at 300 K. The total entropy difference is decomposed as $\Delta S_{\text{total}} = \Delta S_{\text{stretch}} + \Delta S_{\text{ion-ion}}$ where $\Delta S_{\text{stretch}}$ and $\Delta S_{\text{ion-ion}}$ is the change in entropy due to the phonon modes in the high- (>3.9 THz) and low-frequency (<3.9 THz) regions respectively. The values are given in meV K^{-1} f.u.$^{-1}$.

	Ba_3Ge_4			Ba_3Si_4	
	tetragonal	orthorhombic	p-tetragonal	tetragonal	orthorhombic
ΔS_{total}	0	-0.042	-0.054	0	-0.007
$\Delta S_{\text{stretch}}$	0	0.005	0.013	0	0.018
$\Delta S_{\text{ion-ion}}$	0	-0.046	-0.067	0	-0.025

the corresponding two modes for the polymerised chain shift to around 6.0 THz. In the p-tetragonal phase, the dispersion collapses almost entirely as the second chain polymerises along the *b* direction, leading to very sharp peaks in the phonon density of states curve at 4.0 THz.

The clean separation between the internal stretching modes of the Ge_4 unit and the vibrations of the ionic lattice at lower frequency allows us to assess the relative importance of these different modes to the overall entropy change. By integrating the partition function over the frequency range 0 –3.9 THz we can compute the contribution of the low-frequency region containing the ionic motions, and the wagging mode of the Ge_4 butterfly (which we anticipate will be broadly unchanged by the polymerisation) to the entropy (identified as $\Delta S_{\text{ion-ion}}$ in Table 5.4). The difference between this and the total ΔS obtained from integrating over the entire energy range can then be associated with the contribution of the Ge-Ge stretches, $\Delta S_{\text{stretch}}$. The data in Table 5.4 show very clearly that the subtle changes in the high-frequency modes linked directly to the polymerisation make negligible contributions to the total entropy difference which is, instead, dominated by changes in the densely packed low-frequency region associated with the relative motions of the Ba^{2+} and ${Ge_4}^{6-}$ ions. The most significant factor in determining the relative entropies appears, therefore, not to be the polymerisation of the Ge_4 units *per se*, but rather the accompanying ∼2% compression of the unit cell, which causes the modes to harden due to the enforced close approach of the ion cores [27, 206]. To validate this assertion, we have recomputed the entropy and its components for the orthorhombic phase with the cell volume constrained at 220.5 Å^3 f.u.$^{-1}$ (the optimised value for the tetragonal phase), and indeed we find that the difference in entropy almost vanishes (0.005 meV K^{-1} f.u.$^{-1}$).

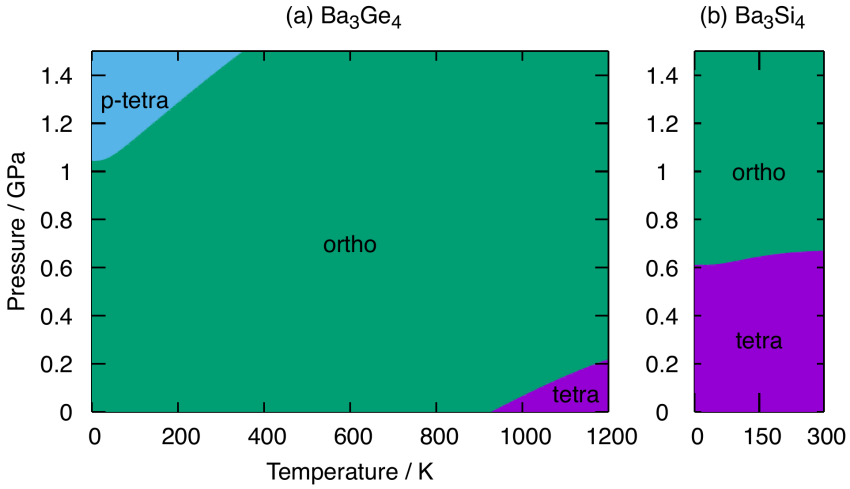


Figure 5.8: The pressure-temperature phase diagrams of Ba_3Ge_4 and Ba_3Si_4 constructed from the Gibbs free energy calculated within the QHA. Green, magenta and cyan indicate the regions of stability for the orthorhombic, tetragonal and p-tetragonal phases, respectively. The temperature range in Ba_3Si_4 is truncated at 300 K, beyond which the orthorhombic phase becomes dynamically unstable.

5.3.3 Pressure-temperature phase diagram

Figure 5.8 shows the phase diagram for Ba_3Ge_4 and Ba_3Si_4 as a function of pressure and temperature. At ambient pressure, the orthorhombic phase of Ba_3Ge_4 is stable up to 930 K, above which the tetragonal phase emerges, precisely as shown in one dimension in Figure 5.4. At elevated pressures, the contraction of the lattice parameters associated with polymerisation increasingly stabilises the orthorhombic and particularly the p-tetragonal phases as a result of the PV term in the Gibbs free energy. At 0 K, a transition from the orthorhombic to the p-tetragonal phases is predicted just above 1.0 GPa, rising to ∼1.5 GPa at 375 K. The critical pressure of ∼1 GPa is accessible, and indeed modest compared to the computed values of ∼4 GPa for the orthorhombic to tetragonal phase transition in $BaGe_2$ [167], again reflecting the more substantial structural changes involved in that case as well as the intrinsically weaker Ge-Ge bonds in ${Ge_4}^{6-}$ *vs* ${Ge_4}^{4-}$. For Ba_3Si_4, in contrast, the tetragonal phase remains stable until ∼0.6 GPa, above which the orthorhombic phase forms. Our analysis of Ba_3Si_4 is restricted to temperatures up to 300 K, above which dynamic instability of the orthorhombic phase limits the application of the QHA. In the accessible regime, the gradient of the orthorhombic-tetragonal phase boundary is close to zero, a result of the smaller differences in ΔS_{total} in Table 5.4 for the Si cluster.

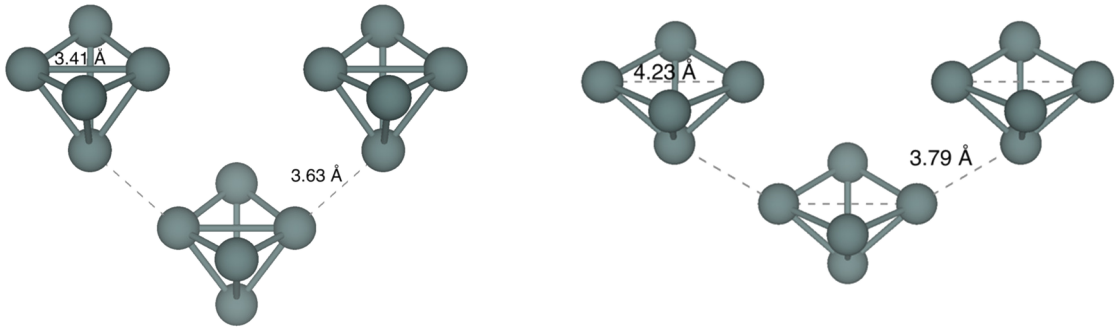


Figure 5.9: Reported structures of Sn_5 clusters in high and low temperature phase of $Sr_{2.04}Ca_{0.96}Sn_5$ from reference [2].

5.4 Elusive bond-stretching isomerism in Ca-doped Sr_3Sn_5

At its inception, our goal in this project was to explore vertical periodic trends down group 14, specifically in the context of the ability of clusters to undergo transitions that make and break bonds between main-group metals. It was, therefore, with great interest that we noted a 2005 report of 'incipient bond-stretch isomerism' [2] in a closely related system, the mixed Ca/Sr phase $Sr_{2.04}Ca_{0.96}Sn_5$ cluster shown in Figure 5.9. The all-Sr analogue, Sr_3Sn_5 adopts the Pu_3Pd_5 structure (*Cmcm*, Z=4) with a network of Sn_5 square pyramids,[207] and indeed $Sr_{2.04}Ca_{0.96}Sn_5$ adopts the same structure at room temperature, with Ca and Sr ions disordered over the cation sites. Within the Wade-Mingos framework, the *arachno* square pyramidal geometry is consistent with a Sn_5^{4-} state rather than the 6- encountered here. Fassler, Nesper and co-workers have analysed the electronic structure of Sr_3Sn_5 and its Ba analogue in some detail and noted that the additional two electrons occupy antibonding states near the Fermi level, causing an elongation of the bonds between the apical and basal Sn atoms (3.15 Å compared to 3.01 Å for the basal-basal distances). The Sn_5 units in $Sr_{2.04}Ca_{0.96}Sn_5$ are linked *via* Sn-Sn contacts between basal sites (3.343 Å) and between apical and basal sites (3.785 Å) which are only marginally longer than the average intra-cluster Sn-Sn bond lengths of 3.08 Å. The corresponding inter-cluster distances in Sr_3Sn_5 are approximately 0.08 Å longer, consistent with the larger radius of Sr^{2+} vs Ca^{2+}. At 100 K, however, the authors reported a structural distortion where the base-base distance increases to 4.05 Å while the base-apex contact contracts marginally to 3.603 Å. At the same time, the intra-cluster base-base bonds contract from 3.016 Å to 2.744 Å and the square base is distorted to a rhombus with one short

diagonal Sn-Sn distance of 3.411 Å. Prompted by the striking similarities between the structural patterns summarised in Figure 5.9 and those in the Ba_3Ge_4 case, particularly the apparent interplay between inter- and intra-cluster bonding, we used density functional theory to probe the potential energy surface for an idealised model with composition close to the experimental one, Sr_2CaSn_5 with the Ca^{2+} and Sr^{2+} ions distributed over the cation sites. Optimisations starting from the high-temperature structure converged without difficulty on a geometry that closely approximates experiment: the average intra-cluster Sn-Sn bond lengths are 3.10 Å, while the base-base and base-apex inter-cluster contacts are 3.3 Å and 3.8 Å, respectively. However, all optimisations started from the 100 K structure reported in reference [2] reverted back to the high-temperature form. Even with the lattice parameters fixed at the 100 K values reported in the paper, the atomic arrangement corresponding to the high-temperature structure is substantially more stable. In short, our analysis of the electronic structure does not offer any plausible explanation for the existence of the low-temperature structure. One possible resolution to this puzzle emerges from a close analysis of the data presented in the supporting information in reference [2]. The space groups of the 273 K and 100 K structures are identical ($Cmnm$), as are the volumes of the unit cell (other than the anticipated thermal expansion at higher temperatures). The individual lattice parameters are also rather similar, other than the fact that the a and b parameters have exchanged places: in the 273 K structure, a = 10.5178(9) Å and b = 8.4789(8) Å while for the 100 K experiment, a = 8.4256(13) Å and b = 10.4659(16) Å. If the a and b axes are swapped in the 100 K structure, all three lattice parameters contract by approximately 0.6%, and the structure of the Sn_5 cluster is very similar to that in the 273 K structure, save for a uniform contraction in all Sn-Sn bond lengths. In the absence of support from theory, it seems likely, therefore, that this particular report of bond-stretch isomerism is a crystallographic artefact.

5.5 Summary and conclusions

This chapter explored the electronic origins of bistability in the Zintl phase Ba3Ge4, where experiments show a transition at 630 K from a tetragonal phase to an orthorhombic one, where half of the Ge_4 tetrahedra are polymerised in a 1-dimensional chain. The energetic balance between the two phases is delicate, with the competing effects of making and breaking Ge–Ge bonds, along with changes in cation/anion interactions, combined to make the overall internal energy change less than 0.11 eV per

f.u. The challenges in accurately capturing these competing effects are illustrated by the functional dependence of the relative energies of the orthorhombic and tetragonal phases. The optBP86b-vdW functional correctly predicts the orthorhombic form to be the most stable at 0 K, but others predict that either the phases are effectively iso-energetic or even the tetragonal phase is more stable. In contrast, analogous calculations for the corresponding silicon compound, Ba_3Si_4, show a strong preference for the (unpolymerised) tetragonal phase, which is indeed the only phase that has been observed experimentally. An analysis of the phonon modes for Ba_3Ge_4 confirms that the tetragonal phase, where the Ge_4 units are not linked, is entropically favoured over the other isomers, leading to a predicted transition temperature of 930 K compared to 630 K measured experimentally. The phonon spectrum can be separated into five Ge–Ge stretching modes at a relatively high frequency (above 3.9 THz) and modes involving the relative motion of the cations and anions at lower frequencies. The making and breaking of Ge–Ge bonds appear to have only a minor impact on the differences in entropy, the major contribution coming instead from changes in the frequencies of the anion/cation vibrations caused by the contraction of the lattice upon polymerisation.

Chapter 6

Summary

This thesis explores the electronic structures and reactivity of clusters under various conditions, including in the gas phase (Chapter 3), on the surface of a substrate (Chapter 4) and embedded in an ionic solid (Chapter 5). Chapter 3 is concerned with the reactivity of coordinatively and electronically unsaturated Ru_4 carbonyl hydrides towards an alkene, 1-hexene in the gaseous phase. Both DFT and LNO-CCSD(T) electronic structure methods are used to identify the global minimum structures of the partially-ligated clusters and to calculate accurate potential energy surfaces for their reactions with 1-hexene. The optimal structures can be explained using the frontier orbital theory, and the *trans* influence of Ru-Ru bonds plays an important role in determining the structures. Moreover, the increased reactivity of $[H_3Ru_4(CO)_9]^-$ is proposed to originate in the combination of hydride ligands to release H_2, preventing the micro reversibility of the 1-ne dehydrogenation step. This chapter highlights the importance of understanding the precise number and location of binding sites in the partially deligated clusters. Chapter 4 moves on to consider the interactions of Ru_3 SCCs with various models for an N-graphene support: pyrrolic-N_4, pyridinic-N_4 and pyridinic-N_6. DFT is used to determine the stability and therefore concentration of each model defect, as well as to analyse the underlying electronic structures. The pyrrolic-N_4 model proves to be the most effective N-graphene support, with significant interactions of π symmetry. The change in the stability of the cluster when it is oxidised is related to the change in the occupancy of π bonding bands, and a reaction mechanism of alcohol oxidation with oxo species is proposed. This chapter highlighted the importance of substrate structures in modelling the support interactions, and consequent reactivity of the cluster. In Chapter 5, phonon frequencies of Zintl phase Ba_3Ge_4 were calculated with DFT and the origin of structural phase transitions was explored. Several DFT functionals were tested for their ability to predict the most stable phase and geometries of different phases, and the vdW functional

optB86b-vdW proved to be optimal for this task. Based on the analysis of phonon dispersion curves and the decomposition of entropy, it was found that it is the volume that drives the phase transition between polymerised and unpolymerised phases. The possibility of the third phase, p-tragonnal phase, was also speculated. This chapter highlights the crucial role of thermal effects on the structure and stability of clusters. Such an understanding of the relationship between local atomic environments and the reactivity of clusters is important for gaining a deeper insight into catalysis and for building predictive models to facilitate a faster and more efficient discovery of catalysts. These include high-throughput screening of compounds with given properties, as well as data-driven approaches, such as machine learning, to inversely design catalysts for a targeted chemical reaction.

Appendix A

Anderson-Newns-Grimley model

A.1 A brief overview

The Anderson-Newns-Grimley (ANG) model can be used as an alternative way to illustrate the interactions between an adsorbate and the surface of metals [208, 209]. In this model, a model Hamiltonian for an adsorbate-substrate system is considered, given as

$$\hat{H} = \epsilon_a \left|a\right\rangle\!\left\langle a\right| + \sum_k \epsilon_k \left|k\right\rangle\!\left\langle k\right| + \sum_k V_{ak} \left|a\right\rangle\!\left\langle k\right| + \text{h.c.} \tag{A.1.1}$$

where $|a\rangle$ and $|k\rangle$ are adsorbate and substrate states respectively, and h.c. denotes hermitian conjugate. In this Hamiltonian, the first and second terms can be thought of as unperturbed Hamiltonians for the adsorbate and substrate states respectively, and the third and fourth terms can be thought of as coupling terms between two sub-systems. The Schrödinger equation can be solved within the Green's function approach and the density of states projected onto the adsorbate orbital is expressed as

$$\begin{aligned} \rho_{\mathrm{a}}(\varepsilon) &= \frac{1}{\pi}\mathrm{Im}\left\langle a|G(\varepsilon)|a\right\rangle \\ &= \frac{1}{\pi}\frac{\Delta(\varepsilon)}{\left[\epsilon - (\varepsilon_{\mathrm{a}} + \Lambda(\varepsilon))\right]^2 + \Delta(\epsilon)^2} \end{aligned} \tag{A.1.2}$$

where

$$\Delta(\varepsilon) = \pi \sum_k \left|V_{ak}\right|^2 \delta(\varepsilon - \varepsilon_k) \tag{A.1.3}$$

and

$$\Lambda(\varepsilon) = \frac{1}{\pi}\mathrm{p.v.}\int_{-\infty}^{\infty} \mathrm{d}\varepsilon' \frac{\Delta(\varepsilon')}{\varepsilon - \varepsilon'} \tag{A.1.4}$$

are called the chemisorption functions, with $\Lambda(\varepsilon)$ being the Hilbert transform of $\Delta(\varepsilon)$ and p.v. denoting the Cauchy principal value of the integral. If the coupling constants between adsorbate and substrate states V_{ak} in equation A.1.3 are assumed to

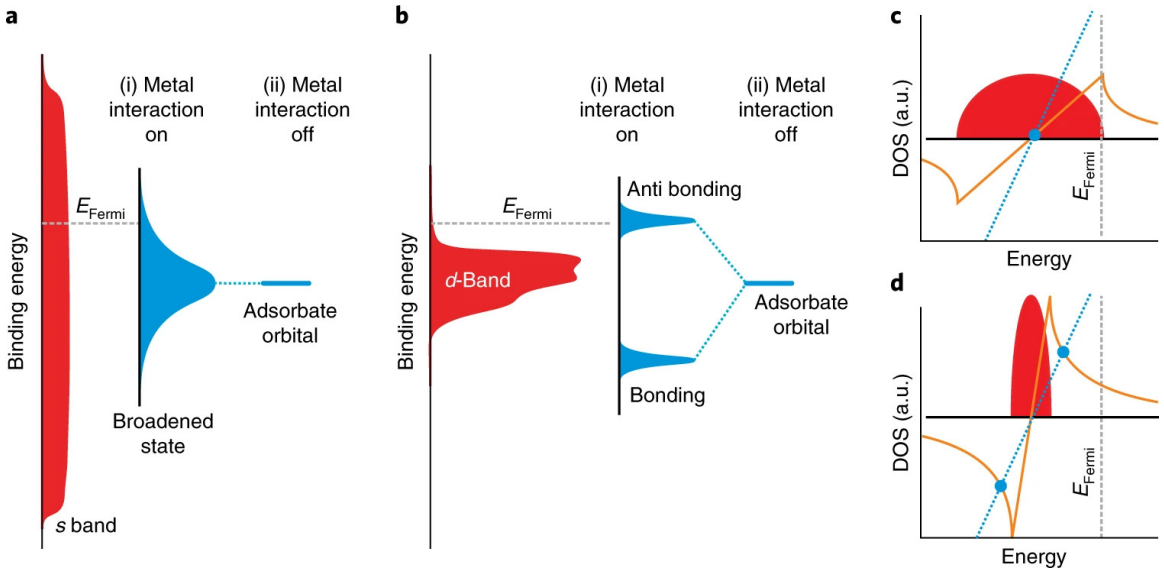


Figure A.1: Illustration of interactions between an adsorbate and substrates and graphical solutions to the ANG model. The DOS of (a) s-band and (b) d-band and their influence on the PDOS of the adsorbate after the chemisorption. Graphical solutions to the model in case of (c) wide metal bands and narrow metal (d) bands (indicated in red). The yellow lines indicate the chemisorption function $\Lambda(\varepsilon)$ and dotted blue lines show the linear function $\varepsilon - \varepsilon_a$. An intersection between these two functions causes the PDOS to peak at that point. Figures adapted with permission from reference [210], Springer Nature.

be constant for all k, then the expression for $\Delta(\varepsilon)$ would be just the density of states on substrate scaled by a constant. By inspecting the expression of PDOS in equation A.1.2, we expect two major different outcomes depending on the shapes of density of states on metal, and the graphical illustrations are depicted in Figure A.1. The first outcome is a development of resonance by the adsorbate with a wide metal band composed of mostly s orbitals in Figure A.1a. This is illustrating that delocalised electrons in metal significantly resonate with the adsorbate orbital, resulting in the broadening of states in PDOS. The second outcome is a split of these renormalised states into two as a result of strong coupling with a narrow d-band on metal, as illustrated in Figure A.1b. This scenario is reminiscent of formations of bonding and antibonding states in the frontier orbital theory that was discussed earlier, but the resulting orbitals are broadened due to resonance. It is also worth noting that, while we made a wild approximation to the chemisorption function for this discussion, the roles of orbital overlaps can be incorporated in V_{ak}. Hence, the adsorption on the metal surface can be qualitatively inspected in terms of orbital interactions, with occupations of (anti)bonding bands (dis)favouring the adsorption. Nevertheless, in both

cases, the ANG model provides an insight into adsorbate-substrate interactions that are unique to solids. Since it was introduced, there have been further developments of several models based on the ANG model to predict the reactivity of metal surfaces, namely d-band reactivity theory and its variations [211–214]. However, these models are inadequate for complex multi-component systems, and also neglect the electric effects of adsorbate induced on its local environment on the substrate [215]. Nevertheless, they have been widely applied to qualitatively evaluate the reactivity of various novel catalysts, such as single-atom catalysts [150] and single-atom alloys [210].

References

[1] Y. Zhao and J. E. McGrady, *Physical Chemistry Chemical Physics*, 2024, **26**, 7318–7328.

[2] S. Gupta and A. K. Ganguli, *Inorganic Chemistry*, 2005, **44**, 7443–7448.

[3] M. W. G. d. Bolster, *Pure and Applied Chemistry*, 1997, **69**, 1251–1304.

[4] B. M. Hoffman, D. Lukoyanov, Z.-Y. Yang, D. R. Dean and L. C. Seefeldt, *Chemical Reviews*, 2014, **114**, 4041–4062.

[5] B. Han, H. Meng, F. Li and J. Zhao, *Catalysts*, 2020, **10**, 974.

[6] J.-C. Liu, X.-L. Ma, Y. Li, Y.-G. Wang, H. Xiao and J. Li, *Nature Communications*, 2018, **9**, 1610.

[7] W. Kang, C. C. Lee, A. J. Jasniewski, M. W. Ribbe and Y. Hu, *Science*, 2020, **368**, 1381–1385.

[8] S. M. Lang, T. M. Bernhardt, R. N. Barnett, B. Yoon and U. Landman, *Journal of the American Chemical Society*, 2009, **131**, 8939–8951.

[9] X. Li, S. Mitchell, Y. Fang, J. Li, J. Perez-Ramirez and J. Lu, *Nature Reviews Chemistry*, 2023, **7**, 754–767.

[10] E. Jimenez-Izal and A. N. Alexandrova, *Annual Review of Physical Chemistry*, 2018, **69**, 377–400.

[11] J. Dadras, L. Shen and A. Alexandrova, *The Journal of Physical Chemistry C*, 2015, **119**, 6047–6055.

[12] S. Mitchell and J. Pérez-Ramírez, *Nature Reviews Materials*, 2021, **6**, 969–985.

[13] S. Ji, Y. Chen, Q. Fu, Y. Chen, J. Dong, W. Chen, Z. Li, Y. Wang, L. Gu, W. He, C. Chen, Q. Peng, Y. Huang, X. Duan, D. Wang, C. Draxl and Y. Li, *Journal of the American Chemical Society*, 2017, **139**, 9795–9798.

[14] K. Fukui, T. Yonezawa and H. Shingu, *The Journal of Chemical Physics*, 1952, **20**, 722–725.

[15] R. B. Woodward and R. Hoffmann, *Angewandte Chemie International Edition*, 1969, **8**, 781–853.

[16] R. Hoffmann, *Angewandte Chemie International Edition*, 1982, **21**, 711–724.

[17] K. Wade, *Journal of the Chemical Society D: Chemical Communications*, 1971, 792–793.

[18] D. M. P. Mingos, *Nature Physical Science*, 1972, **236**, 99–102.

[19] E. D. Jemmis, M. M. Balakrishnarajan and P. D. Pancharatna, *Journal of the American Chemical Society*, 2001, **123**, 4313–4323.

[20] G. Blyholder, *The Journal of Physical Chemistry*, 1964, **68**, 2772–2777.

[21] A. Nilsson, M. Weinelt, T. Wiell, P. Bennich, O. Karis, N. Wassdahl, J. Stöhr and M. G. Samant, *Physical Review Letters*, 1997, **78**, 2847–2850.

[22] M. K. Sabbe, M.-F. Reyniers and K. Reuter, *Catalysis Science & Technology*, 2012, **2**, 2010–2024.

[23] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Courier Corporation, 2012.

[24] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods*, Cambridge University Press, Cambridge, 2nd edn., 2020.

[25] R. G. Parr and Y. Weitao, *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, 1995.

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

[27] J. M. Ziman, *Electrons and Phonons: The Theory of Transport Phenomena in Solids*, Oxford University Press, 2001.

[28] D. R. Hartree, *The calculation of atomic structures.*, J. Wiley, New York, 1957.

[29] J. C. Slater, *Physical Review*, 1930, **35**, 210–211.

[30] J. C. Slater, *Physical Review*, 1929, **34**, 1293–1322.

[31] E. U. Condon, *Physical Review*, 1930, **36**, 1121–1133.

[32] Z. Rolik and M. Kállay, *The Journal of Chemical Physics*, 2011, **135**, 104111.

[33] Z. Rolik, L. Szegedy, I. Ladjánszki, B. Ladóczki and M. Kállay, *The Journal of Chemical Physics*, 2013, **139**, 094105.

[34] P. R. Nagy and M. Kállay, *The Journal of Chemical Physics*, 2017, **146**, 214106.

[35] P. R. Nagy, G. Samu and M. Kállay, *Journal of Chemical Theory and Computation*, 2018, **14**, 4193–4215.

[36] C. Riplinger, B. Sandhoefer, A. Hansen and F. Neese, *The Journal of Chemical Physics*, 2013, **139**, 134101.

[37] Y. Guo, C. Riplinger, U. Becker, D. G. Liakos, Y. Minenkov, L. Cavallo and F. Neese, *The Journal of Chemical Physics*, 2018, **148**, 011101.

[38] F. Neese, *WIREs Computational Molecular Science*, 2012, **2**, 73–78.

[39] Q. Ma and H.-J. Werner, *Journal of Chemical Theory and Computation*, 2018, **14**, 198–215.

[40] Q. Ma and H.-J. Werner, *WIREs Computational Molecular Science*, 2018, **8**, e1371.

[41] S. G. Balasubramani, G. P. Chen, S. Coriani, M. Diedenhofen, M. S. Frank, Y. J. Franzke, F. Furche, R. Grotjahn, M. E. Harding, C. Hättig, A. Hellweg, B. Helmich-Paris, C. Holzer, U. Huniar, M. Kaupp, A. Marefat Khah, S. Karbalaei Khani, T. Müller, F. Mack, B. D. Nguyen, S. M. Parker, E. Perlt, D. Rappoport, K. Reiter, S. Roy, M. Rückert, G. Schmitz, M. Sierka, E. Tapavicza, D. P. Tew, C. van Wüllen, V. K. Voora, F. Weigend, A. Wodyński and J. M. Yu, *The Journal of Chemical Physics*, 2020, **152**, 184107.

[42] D. P. Tew, *Journal of Chemical Theory and Computation*, 2019, **15**, 6597–6606.

[43] H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, T. F. Miller, III, A. Mitrushchenkov, K. A. Peterson, I. Polyak, G. Rauhut and M. Sibaev, *The Journal of Chemical Physics*, 2020, **152**, 144107.

[44] I. Efremenko and J. M. L. Martin, *The Journal of Physical Chemistry A*, 2021, **125**, 8987–8999.

[45] E. Semidalas and J. M. L. Martin, *Journal of Chemical Theory and Computation*, 2022, **18**, 883–898.

[46] N. Sylvetsky, A. Banerjee, M. Alonso and J. M. L. Martin, *Journal of Chemical Theory and Computation*, 2020, **16**, 3641–3653.

[47] G. Santra, E. Semidalas, N. Mehta, A. Karton and J. M. L. Martin, *Physical Chemistry Chemical Physics*, 2022, **24**, 25555–25570.

[48] P. R. Nagy, L. Gyevi-Nagy, B. D. Lőrincz and M. Kállay, *Molecular Physics*, 2023, **121**, e2109526.

[49] G. Santra and J. M. Martin, *The Journal of Physical Chemistry A*, 2022, **126**, 9375–9391.

[50] P. Hohenberg and W. Kohn, *Physical Review*, 1964, **136**, B864–B871.

[51] W. Kohn and L. J. Sham, *Physical Review*, 1965, **140**, A1133–A1138.

[52] J. Harris, *Physical Review A*, 1984, **29**, 1648–1659.

[53] A. D. Becke, *The Journal of Chemical Physics*, 1993, **98**, 1372–1377.

[54] J. P. Perdew, M. Ernzerhof and K. Burke, *The Journal of Chemical Physics*, 1996, **105**, 9982–9985.

[55] P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *The Journal of Physical Chemistry*, 1994, **98**, 11623–11627.

[56] E. Brémond and C. Adamo, *The Journal of Chemical Physics*, 2011, **135**, 024106.

[57] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth and B. I. Lundqvist, *Physical Review Letters*, 2004, **92**, 246401.

[58] T. Kato, *Communications on Pure and Applied Mathematics*, 1957, **10**, 151–177.

[59] F. Weigend and R. Ahlrichs, *Physical Chemistry Chemical Physics*, 2005, **7**, 3297–3305.

[60] T. H. Dunning, Jr., *The Journal of Chemical Physics*, 1989, **90**, 1007–1023.

[61] A. Togo, *Journal of the Physical Society of Japan*, 2023, **92**, 012001.

[62] P. Buchwalter, J. Rosé and P. Braunstein, *Chemical Reviews*, 2015, **115**, 28–126.

[63] R. D. Adams, B. Captain and L. Zhu, *Journal of Organometallic Chemistry*, 2006, **691**, 3122–3128.

[64] A. F. Abdel-Magied, Y. Theibich, A. K. Singh, A. Rahaman, I. Doverbratt, A. K. Raha, M. Haukka, M. G. Richmond and E. Nordlander, *Dalton Transactions*, 2020, **49**, 4244–4256.

[65] T. Gutmann, B. Walaszek, X. Yeping, M. Wächtler, I. del Rosal, A. Grünberg, R. Poteau, R. Axet, G. Lavigne, B. Chaudret, H.-H. Limbach and G. Buntkowsky, *Journal of the American Chemical Society*, 2010, **132**, 11759–11767.

[66] E. L. Muetterties, T. N. Rhodin, E. Band, C. F. Brucker and W. R. Pretzer, *Chemical Reviews*, 1979, **79**, 91–137.

[67] J. Lewis and B. F. G. Johnson, *Pure and Applied Chemistry*, 1975, **44**, 43–79.

[68] G. A. Ozin, *Catalysis Reviews*, 1977, **16**, 191–289.

[69] D. B. Brown, M. Cripps, B. F. G. Johnson, C. M. Martin, D. Braga and F. Grepioni, *Chemical Communications*, 1996, 1425–1426.

[70] G. Süss-Fink, M. Faure and T. R. Ward, *Angewandte Chemie International Edition*, 2002, **41**, 99–101.

[71] G. Süss-Fink, B. Therrien, L. Vieille-Petit, M. Tschan, V. B. Romakh, T. R. Ward, M. Dadras and G. Laurenczy, *Journal of Organometallic Chemistry*, 2004, **689**, 1362–1369.

[72] C. M. Hagen, L. Vieille-Petit, G. Laurenczy, G. Süss-Fink and R. G. Finke, *Organometallics*, 2005, **24**, 1819–1831.

[73] G. Süss-Fink, *Journal of Organometallic Chemistry*, 2014, **751**, 2–19.

[74] J. Zhang and W. An, *Catalysis Science & Technology*, 2022, **12**, 2604–2617.

[75] J. Shan, C. Ye, Y. Jiang, M. Jaroniec, Y. Zheng and S.-Z. Qiao, *Science Advances*, 2022, **8**, eabo0762.

[76] I. G. Powers and C. Uyeda, *ACS Catalysis*, 2017, **7**, 936–958.

[77] J. G. Vitillo, C. J. Cramer and L. Gagliardi, *Israel Journal of Chemistry*, 2022, **62**, e202100136.

[78] E. Crawford, P. J. Dyson, O. Forest, S. Kwok and J. Scott McIndoe, *Journal of Cluster Science*, 2006, **17**, 47–63.

[79] C. P. G. Butcher, A. Dinca, P. J. Dyson, B. F. G. Johnson, P. R. R. Langridge-Smith and J. S. McIndoe, *Angewandte Chemie International Edition*, 2003, **42**, 5752–5755.

[80] P. J. Dyson, A. K. Hearley, B. F. G. Johnson, J. S. McIndoe, P. R. R. Langridge-Smith and C. Whyte, *Rapid Communications in Mass Spectrometry*, 2001, **15**, 895–897.

[81] C. P. G. Butcher, P. J. Dyson, B. F. G. Johnson, P. R. R. Langridge-Smith, J. S. McIndoe and C. Whyte, *Rapid Communications in Mass Spectrometry*, 2002, **16**, 1595–1598.

[82] M. A. Henderson, S. Kwok and J. S. McIndoe, *Journal of the American Society for Mass Spectrometry*, 2009, **20**, 658–666.

[83] K. L. Vikse and J. S. McIndoe, *Pure and Applied Chemistry*, 2015, **87**, 361–377.

[84] S. M. Lang, S. U. Förtig, T. M. Bernhardt, M. Krstić and V. Bonačić-Koutecký, *The Journal of Physical Chemistry A*, 2014, **118**, 8356–8359.

[85] P. J. Dyson, A. K. Hearley, B. F. G. Johnson, P. R. R. Langridge-Smith and J. S. McIndoe, *Inorganic Chemistry*, 2004, **43**, 4962–4973.

[86] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders and T. Ziegler, *J. Comput. Chem.*, 2001, **22**, 931–967.

[87] E. Van Lenthe and E. J. Baerends, *Journal of Computational Chemistry*, 2003, **24**, 1142–1156.

[88] E. van Lenthe, E. J. Baerends and J. G. Snijders, *The Journal of Chemical Physics*, 1994, **101**, 9783–9792.

[89] E. v. Lenthe, E. J. Baerends and J. G. Snijders, *The Journal of Chemical Physics*, 1993, **99**, 4597–4610.

[90] E. van Lenthe, A. Ehlers and E.-J. Baerends, *The Journal of Chemical Physics*, 1999, **110**, 8943–8953.

[91] J. W. Furness, A. D. Kaplan, J. Ning, J. P. Perdew and J. Sun, *The Journal of Physical Chemistry Letters*, 2020, **11**, 8208–8215.

[92] S. Lehtola, C. Steigemann, M. J. T. Oliveira and M. A. L. Marques, *SoftwareX*, 2018, **7**, 1–5.

[93] E. Caldeweyher, S. Ehlert, A. Hansen, H. Neugebauer, S. Spicher, C. Bannwarth and S. Grimme, *The Journal of Chemical Physics*, 2019, **150**, 154122.

[94] S. Ehlert, U. Huniar, J. Ning, J. W. Furness, J. Sun, A. D. Kaplan, J. P. Perdew and J. G. Brandenburg, *The Journal of Chemical Physics*, 2021, **154**, 061101.

[95] P. R. Nagy and M. Kállay, *Journal of Chemical Theory and Computation*, 2019, **15**, 5275–5298.

[96] M. Kállay, P. R. Nagy, D. Mester, Z. Rolik, G. Samu, J. Csontos, J. Csóka, P. B. Szabó, L. Gyevi-Nagy, B. Hégely, I. Ladjánszki, L. Szegedy, B. Ladóczki, K. Petrov, M. Farkas, P. D. Mezei and A. Ganyecz, *The Journal of Chemical Physics*, 2020, **152**, 074107.

[97] D. Rappoport and F. Furche, *The Journal of Chemical Physics*, 2010, **133**, 134105.

[98] A. Hellweg and D. Rappoport, *Physical Chemistry Chemical Physics*, 2015, **17**, 1010–1017.

[99] R. Hoffmann, B. E. R. Schilling, R. Bau, H. D. Kaesz and D. M. P. Mingos, *Journal of the American Chemical Society*, 1978, **100**, 6088–6093.

[100] H. Ding, Y. Lu, Y. Xie, H. Liu and H. F. I. Schaefer, *Journal of Chemical Theory and Computation*, 2015, **11**, 940–949.

[101] H. Wang, Y. Xie, R. B. King and H. F. Schaefer, *Journal of Organometallic Chemistry*, 2008, **693**, 3201–3212.

[102] Y. Xie, H. F. Schaefer and R. B. King, *Journal of the American Chemical Society*, 2000, **122**, 8746–8761.

[103] B. Peng, Q.-S. Li, Y. Xie, R. B. King and H. F. S. Iii, *Dalton Transactions*, 2008, 6977–6986.

[104] H. Wang, Y. Xie, R. B. King and H. F. Schaefer, *Journal of the American Chemical Society*, 2005, **127**, 11646–11651.

[105] P. J. Dyson, J. E. McGrady, M. Reinhold, B. F. G. Johnson, J. S. McIndoe and P. R. R. Langridge-Smith, *Journal of Cluster Science*, 2000, **11**, 391–401.

[106] F. M. Bickelhaupt and E. J. Baerends, in *Reviews in Computational Chemistry*, John Wiley & Sons, Ltd, 2000, pp. 1–86.

[107] M. Mitoraj and A. Michalak, *Organometallics*, 2007, **26**, 6576–6580.

[108] J. A. Cabeza, I. del Río, J. M. Fernández-Colinas, E. Pérez-Carreño and D. Vázquez-García, *Organometallics*, 2010, **29**, 4818–4828.

[109] T. Chihara and H. Yamazaki, *Journal of the Chemical Society, Dalton Transactions*, 1995, 1369–1377.

[110] J. T. S. Irvine, D. Neagu, M. C. Verbraeken, C. Chatzichristodoulou, C. Graves and M. B. Mogensen, *Nature Energy*, 2016, **1**, 1–13.

[111] H.-C. Fu, F. You, H.-R. Li and L.-N. He, *Frontiers in Chemistry*, 2019, **7**, 525.

[112] Y. Lv, S. Duan and R. Wang, *Progress in Natural Science: Materials International*, 2020, **30**, 1–12.

[113] S. K. Dutta, S. K. Mehetor and N. Pradhan, *The Journal of Physical Chemistry Letters*, 2015, **6**, 936–944.

[114] Y. Xin, S. Li, Y. Qian, W. Zhu, H. Yuan, P. Jiang, R. Guo and L. Wang, *ACS Catalysis*, 2020, **10**, 11280–11306.

[115] Y. Sun and S. Dai, *Science Advances*, 2021, **7**, eabg1600.

[116] D. Yang and B. C. Gates, *ACS Catalysis*, 2019, **9**, 1779–1798.

[117] M. S. Alhumaimess, *Journal of Saudi Chemical Society*, 2020, **24**, 461–473.

[118] A. Bavykina, N. Kolobov, I. S. Khan, J. A. Bau, A. Ramirez and J. Gascon, *Chemical Reviews*, 2020, **120**, 8468–8535.

[119] B. Qiao, A. Wang, X. Yang, L. F. Allard, Z. Jiang, Y. Cui, J. Liu, J. Li and T. Zhang, *Nature Chemistry*, 2011, **3**, 634–641.

[120] J. Liu, *ACS Catalysis*, 2017, **7**, 34–59.

[121] G. Kyriakou, M. B. Boucher, A. D. Jewell, E. A. Lewis, T. J. Lawton, A. E. Baber, H. L. Tierney, M. Flytzani-Stephanopoulos and E. C. H. Sykes, *Science*, 2012, **335**, 1209–1212.

[122] G. Giannakakis, M. Flytzani-Stephanopoulos and E. C. H. Sykes, *Accounts of Chemical Research*, 2019, **52**, 237–247.

[123] X.-L. Ma, J.-C. Liu, H. Xiao and J. Li, *Journal of the American Chemical Society*, 2018, **140**, 46–49.

[124] Y.-H. Lu, M. Zhou, C. Zhang and Y.-P. Feng, *The Journal of Physical Chemistry C*, 2009, **113**, 20156–20160.

[125] S. Sun, G. Zhang, N. Gauquelin, N. Chen, J. Zhou, S. Yang, W. Chen, X. Meng, D. Geng, M. N. Banis, R. Li, S. Ye, S. Knights, G. A. Botton, T.-K. Sham and X. Sun, *Scientific Reports*, 2013, **3**, 1775.

[126] A. Wang, J. Li and T. Zhang, *Nature Reviews Chemistry*, 2018, **2**, 65–81.

[127] L. Wang, Z. Sofer and M. Pumera, *ACS Nano*, 2020, **14**, 21–25.

[128] H. Xu, D. Cheng, D. Cao and X. C. Zeng, *Nature Catalysis*, 2018, **1**, 339–348.

[129] G. Zheng, L. Li, Z. Tian, X. Zhang and L. Chen, *Journal of Energy Chemistry*, 2021, **54**, 612–619.

[130] X. Li, H. Li, Z. Zhang, J. Q. Shi, Y. Jiao and S.-Z. Qiao, *Nano Energy*, 2023, **115**, 108695.

[131] W. Pei, S. Zhou, J. Zhao, X. Xu, Y. Du and S. X. Dou, *Nano Energy*, 2020, **76**, 105049.

[132] W. Pei, W. Zhang, X. Yu, L. Hou, W. Xia, Z. Wang, Y. Liu, S. Zhou, Y. Tu and J. Zhao, *Journal of Materials Informatics*, 2023, **3**, 26.

[133] X. Shi, Y. Li, S. Zhang, R. Hu, S. Gao, P. Jin, J. Shang and J. Shui, *Nano Research*, 2023, **16**, 8042–8050.

[134] D. Ma, Z. Zeng, L. Liu, X. Huang and Y. Jia, *The Journal of Physical Chemistry C*, 2019, **123**, 19066–19076.

[135] Z. Yin, Z. Wang, Y. Gao, Z. Wang, Z. Wei and H. Wang, *Molecular Catalysis*, 2023, **550**, 113556.

[136] C. Liu, T. Li, X. Dai, J. Zhao, D. He, G. Li, B. Wang and X. Cui, *Journal of the American Chemical Society*, 2022, **144**, 4913–4924.

[137] X. Li, N. Guo, Z. Chen, X. Zhou, X. Zhao, Y. Du, L. Ma, Y. Fang, H. Xu, H. Yang, W. Yu, S. Lu, M. Tian, Q. He, K. P. Loh, S. Xi, C. Zhang and J. Lu, *Advanced Functional Materials*, 2022, **32**, 2200933.

[138] G. Kresse and J. Hafner, *Physical Review B*, 1993, **47**, 558–561.

[139] G. Kresse and J. Furthmüller, *Computational Materials Science*, 1996, **6**, 15–50.

[140] G. Kresse and J. Furthmüller, *Physical Review B*, 1996, **54**, 11169–11186.

[141] G. Kresse and D. Joubert, *Physical Review B*, 1999, **59**, 1758–1775.

[142] J. P. Perdew, K. Burke and M. Ernzerhof, *Physical Review Letters*, 1996, **77**, 3865–3868.

[143] G. Henkelman and H. Jónsson, *The Journal of Chemical Physics*, 2000, **113**, 9978–9985.

[144] G. Henkelman, B. P. Uberuaga and H. Jónsson, *The Journal of Chemical Physics*, 2000, **113**, 9901–9904.

[145] J. Lee, Z. Yang, W. Zhou, S. J. Pennycook, S. T. Pantelides and M. F. Chisholm, *Proceedings of the National Academy of Sciences*, 2014, **111**, 7522–7526.

[146] D. M. Thomas, Y. Asiri and N. D. Drummond, *Physical Review B*, 2022, **105**, 184114.

[147] R. Hu, Y. Li, Q. Zeng and J. Shang, *Applied Surface Science*, 2020, **525**, 146588.

[148] Liu Jin-Cheng, Xiao Hai, Zhao Xiao-Kun, Zhang Nan-Nan, Liu Yuan, Xing Deng-Hui, Yu Xiaohu, Hu Han-Shi and Li Jun, *CCS Chemistry*, 2022, **5**, 152–163.

[149] Z. Hou, X. Wang, T. Ikeda, K. Terakura, M. Oshima and M.-a. Kakimoto, *Physical Review B*, 2013, **87**, 165401.

[150] S. Liu, Z. Li, C. Wang, W. Tao, M. Huang, M. Zuo, Y. Yang, K. Yang, L. Zhang, S. Chen, P. Xu and Q. Chen, *Nature Communications*, 2020, **11**, 938.

[151] S. Hong and G. Mpourmpakis, *Catalysis Science & Technology*, 2021, **11**, 6390–6400.

[152] J. Zhao, Q. Du, S. Zhou and V. Kumar, *Chemical Reviews*, 2020, **120**, 9021–9163.

[153] B. Weinert and S. Dehnen, in *Clusters – Contemporary Insight in Structure and Bonding*, ed. S. Dehnen, Springer International Publishing, Cham, 2017, pp. 99–134.

[154] D. M. P. Mingos, in *50th Anniversary of Electron Counting Paradigms for Polyhedral Molecules : Bonding in Clusters, Intermetallics and Intermetalloids*, ed. D. Mingos, Springer International Publishing, Cham, 2021, pp. 1–67.

[155] F. Gam, J. Wei, S. Kahlal, J.-Y. Saillard and J.-F. Halet, in *50th Anniversary of Electron Counting Paradigms for Polyhedral Molecules : Bonding in Clusters, Intermetallics and Intermetalloids*, ed. D. Mingos, Springer International Publishing, Cham, 2021, pp. 69–102.

[156] W. Klein, A. Schier and T. F. Fässler, in *50th Anniversary of Electron Counting Paradigms for Polyhedral Molecules : Bonding in Clusters, Intermetallics and Intermetalloids*, ed. D. Mingos, Springer International Publishing, Cham, 2021, pp. 149–195.

[157] F. Pan, B. Weinert and S. Dehnen, in *50th Anniversary of Electron Counting Paradigms for Polyhedral Molecules : Bonding in Clusters, Intermetallics and Intermetalloids*, ed. D. Mingos, Springer International Publishing, Cham, 2021, pp. 103–148.

[158] S. M. Kauzlarich, *Chemistry of Materials*, 2023, **35**, 7355–7362.

[159] S. M. Kauzlarich, Z. Ju, E. Tseng and J. Lundervold, *Chemical Society Reviews*, 2021, **50**, 13236–13252.

[160] S. Fang, J. Li, K. Zou, H. Shuai, L. Xu, W. Deng, G. Zou, H. Hou and X. Ji, *Chemical Engineering Journal*, 2022, **433**, 133841.

[161] R. Q. Cabrera, A. Salamat, O. I. Barkalov, O. Leynaud, P. Hutchins, D. Daisenberger, D. Machon, A. Sella, D. W. Lewis and P. F. McMillan, *Journal of Solid State Chemistry*, 2009, **182**, 2535–2542.

[162] C. Lorenz, S. Gärtner and N. Korber, *Crystals*, 2018, **8**, 276.

[163] J. T. Vaughey, G. J. Miller, S. Gravelle, E. Alejandro Leon-Escamilla and J. D. Corbett, *Journal of Solid State Chemistry*, 1997, **133**, 501–507.

[164] J. Evers, G. Oehlinger and H. R. Ott, *Journal of the Less Common Metals*, 1980, **69**, 389–391.

[165] R. Castillo, A. I. Baranov, U. Burkhardt, Y. Grin and U. Schwarz, *Zeitschrift für anorganische und allgemeine Chemie*, 2015, **641**, 355–361.

[166] U. Schwarz, R. Castillo, J. M. Hübner, A. Wosylus, Y. Prots, M. Bobnar and Y. Grin, *Zeitschrift für Naturforschung B*, 2020, **75**, 209–216.

[167] J.-T. Wang, C. Chen and Y. Kawazoe, *Physical Review B*, 2015, **91**, 054107.

[168] M. Pani and A. Palenzona, *Journal of Alloys and Compounds*, 2008, **462**, L9–L11.

[169] R. Benhafid, A. Belgacem Bouzida, Y. Djaballah, A. Candan, A. İyigör and G. Uğur, *Journal of Phase Equilibria and Diffusion*, 2019, **40**, 195–205.

[170] J. Shi, W. Cui, J. A. Flores-Livas, A. San-Miguel, S. Botti and M. A. L. Marques, *Physical Chemistry Chemical Physics*, 2016, **18**, 8108–8114.

[171] B. Eisenmann, K. H. Janzon, H. Schäfer and A. Weiss, *Zeitschrift für Naturforschung B*, 1969, **24**, 457–458.

[172] U. Aydemir, A. Ormeci, H. Borrmann, B. Böhme, F. Zürcher, B. Uslu, T. Goebel, W. Schnelle, P. Simon, W. Carrillo-Cabrera, F. Haarmann, M. Baitinger, R. Nesper, H. G. von Schnering and Y. Grin, *Zeitschrift für anorganische und allgemeine Chemie*, 2008, **634**, 1651–1661.

[173] F. Zürcher and R. Nesper, *Angewandte Chemie International Edition*, 1998, **37**, 3314–3318.

[174] F. Zürcher, S. Leoni and R. Nesper, *Zeitschrift für Kristallographie - Crystalline Materials*, 2003, **218**, 171–177.

[175] W. D. Stohrer and R. Hoffmann, *Journal of the American Chemical Society*, 1972, **94**, 1661–1668.

[176] W. D. Stohrer and R. Hoffmann, *Journal of the American Chemical Society*, 1972, **94**, 779–786.

[177] V. C. Gibson and M. McPartlin, *Journal of the Chemical Society, Dalton Transactions*, 1992, 947–956.

[178] M.-M. Rohmer and M. Bénard, *Chemical Society Reviews*, 2001, **30**, 340–354.

[179] J. E. McGrady, *Angewandte Chemie International Edition*, 2000, **39**, 3077–3079.

[180] J. A. Labinger, *Comptes Rendus Chimie*, 2002, **5**, 235–244.

[181] B. Hammann, C. Chen, U. Flörke, R. Hauptmann, E. Bill, S. Sinnecker and G. Henkel, *Angewandte Chemie International Edition*, 2006, **45**, 8245–8249.

[182] J. M. Mayer, *Angewandte Chemie International Edition*, 1992, **31**, 286–287.

[183] P. R. Remya and C. H. Suresh, *Journal of Computational Chemistry*, 2017, **38**, 1704–1711.

[184] F. Breher, *Coordination Chemistry Reviews*, 2007, **251**, 1007–1043.

[185] P. v. R. Schleyer, A. F. Sax, J. Kalcher and R. Janoschek, *Angewandte Chemie International Edition*, 1987, **26**, 364–366.

[186] R. Koch, T. Bruhn and M. Weidenbruch, *Journal of Molecular Structure: THEOCHEM*, 2004, **680**, 91–97.

[187] R. Koch, T. Bruhn and M. Weidenbruch, *Journal of Molecular Structure: THEOCHEM*, 2005, **714**, 109–115.

[188] G. Román-Pérez and J. M. Soler, *Physical Review Letters*, 2009, **103**, 096102.

[189] J. Klimeš, D. R. Bowler and A. Michaelides, *Physical Review B*, 2011, **83**, 195131.

[190] J. Sun, A. Ruzsinszky and J. P. Perdew, *Physical Review Letters*, 2015, **115**, 036402.

[191] R. Sabatini, T. Gorni and S. de Gironcoli, *Physical Review B*, 2013, **87**, 041108.

[192] J. Ning, M. Kothakonda, J. W. Furness, A. D. Kaplan, S. Ehlert, J. G. Brandenburg, J. P. Perdew and J. Sun, *Physical Review B*, 2022, **106**, 075422.

[193] G. Pizzi, V. Vitale, R. Arita, S. Blügel, F. Freimuth, G. Géranton, M. Gibertini, D. Gresch, C. Johnson, T. Koretsune, J. Ibañez-Azpiroz, H. Lee, J.-M. Lihm, D. Marchand, A. Marrazzo, Y. Mokrousov, J. I. Mustafa, Y. Nohara, Y. Nomura, L. Paulatto, S. Poncé, T. Ponweiser, J. Qiao, F. Thöle, S. S. Tsirkin, M. Wierzbowska, N. Marzari, D. Vanderbilt, I. Souza, A. A. Mostofi and J. R. Yates, *Journal of Physics: Condensed Matter*, 2020, **32**, 165902.

[194] D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson and G. Henkelman, *The Journal of Chemical Physics*, 2012, **136**, 074103.

[195] A. Togo, L. Chaput, T. Tadano and I. Tanaka, *Journal of Physics: Condensed Matter*, 2023, **35**, 353001.

[196] F. Birch, *Physical Review*, 1947, **71**, 809–824.

[197] R. P. Stoffel, C. Wessel, M.-W. Lumey and R. Dronskowski, *Angewandte Chemie International Edition*, 2010, **49**, 5242–5266.

[198] V. L. Deringer, R. P. Stoffel and R. Dronskowski, *Physical Review B*, 2014, **89**, 094303.

[199] J. M. Skelton, S. C. Parker, A. Togo, I. Tanaka and A. Walsh, *Physical Review B*, 2014, **89**, 205203.

[200] J. Tsuchiya, T. Tsuchiya and R. M. Wentzcovitch, *Journal of Geophysical Research: Solid Earth*, 2005, **110**, B02204.

[201] R. M. Wentzcovitch, B. B. Karki, M. Cococcioni and S. de Gironcoli, *Physical Review Letters*, 2004, **92**, 018501.

[202] T. Tadano and S. Tsuneyuki, *Journal of the Physical Society of Japan*, 2018, **87**, 041015.

[203] B. Sadovyi, M. Wierzbowska, S. Stelmakh, S. Boccato, S. Gierlotka, T. Irifune, S. Porowski and I. Grzegory, *Physical Review B*, 2020, **102**, 235109.

[204] K. Tolborg, J. Klarbring, A. M. Ganose and A. Walsh, *Digital Discovery*, 2022, **1**, 586–595.

[205] I. Pallikara and J. M. Skelton, *Physical Chemistry Chemical Physics*, 2021, **23**, 19219–19236.

[206] M. T. Agne, S. Anand and G. J. Snyder, *Research*, 2022, **2022**, DOI: 10.34133/2022/9786705.

[207] F. Zürcher, R. Nesper, S. Hoffmann and T. F. Fässler, *Zeitschrift für anorganische und allgemeine Chemie*, 2001, **627**, 2211–2219.

[208] P. W. Anderson, *Physical Review*, 1961, **124**, 41–53.

[209] D. M. Edwards and D. M. Newns, *Physics Letters A*, 1967, **24**, 236–237.

[210] M. T. Greiner, T. E. Jones, S. Beeg, L. Zwiener, M. Scherzer, F. Girgsdies, S. Piccinin, M. Armbrüster, A. Knop-Gericke and R. Schlögl, *Nature Chemistry*, 2018, **10**, 1008–1015.

[211] B. Hammer and J. K. Nørskov, *Surface Science*, 1995, **343**, 211–220.

[212] A. Vojvodic, J. K. Nørskov and F. Abild-Pedersen, *Topics in Catalysis*, 2014, **57**, 25–32.

[213] H. Xin, A. Vojvodic, J. Voss, J. K. Nørskov and F. Abild-Pedersen, *Physical Review B*, 2014, **89**, 115114.

[214] S. Bhattacharjee, U. V. Waghmare and S.-C. Lee, *Scientific Reports*, 2016, **6**, 35916.

[215] S. Saini, J. Halldin Stenlid and F. Abild-Pedersen, *npj Computational Materials*, 2022, **8**, 1–12.