

IMPERIAL COLLEGE LONDON

DOCTORAL THESIS

---

**Embedding high-level quantum  
mechanical approaches within  
linear-scaling density functional theory**

---

*Author:*

Robert John CHARLTON

*Supervisors:*

Prof. Peter HAYNES  
& Prof. Andrew HORSFIELD

A thesis submitted for the degree of Doctor of Philosophy  
at Imperial College London

Theory and Simulation of Materials Centre for Doctoral Training  
Department of Materials

March 11, 2020

## Declaration of Originality

I declare that the work presented here is my own.

Unless otherwise stated, any and all work contained in this thesis is my own.

Appropriate references are used in all cases where work is that of others.

This thesis has not been submitted for award or examination elsewhere.

Chapter 5 contains work originally published as part of an earlier publication, and is reprinted here with the permission of AIP Publishing.

Chapter 7 contains work originally published as part of an earlier publication, and is reprinted here with the permission of the American Chemical Society.

## Declaration of Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-NonCommercial 4.0 International Licence (CC BY-NC).

Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose.

When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes.

Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

# Acknowledgements

The research carried out as part of this thesis was part-financed by the EPSRC Centre for Doctoral Training in Theory and Simulation of Materials (No. EP/L0155179/1) and a President's PhD Scholarship from Imperial College London.

I would like to thank my supervisors Peter Haynes and Andrew Horsfield, for their boundless wisdom and patience throughout my PhD studies.

I am grateful to all the talented researchers I had the pleasure to work with at various stages in my time at Imperial. A special thanks in particular goes to Joseph Prentice and Richard Fogarty, without whose help my research would never have come to this. I owe a debt of gratitude to all the members of the ONETEP developers' group, particularly Nicholas Hine, James Womack and Jacek Dziedzic for their advice and thorough code reviews.

I am grateful to all members of the TSM-CDT, who helped to make my time at Imperial such an enjoyable experience. My thanks in particular go to Miranda Smith and all the administrative staff for making life easy, and to all the members of Cohort VI, with whom I share many fond memories.

Finally, I thank my family for all their love and support throughout my life to this stage. Tá an tráchtas seo scríofa i ndíl chuimhne m'athair. Ní bheadh mé anseo gan a anam agus a ghrá.

# Abstract

Advances in computational methods in recent decades have significantly expanded the range of problems in condensed matter physics that can be tackled from first principles. Linear-scaling density functional theory methods enable quantum mechanical calculations to be performed on systems containing tens of thousands of atoms, with modern approaches capable of reproducing the accuracy of plane wave DFT approaches. This opens up the possibility of treating highly complex molecular systems such as doped organic molecular crystals that require the dopant molecule to be contained within a large periodic structure. One example of such a system is pentacene in *p*-terphenyl, a system that finds use as a room-temperature maser. Understanding the maser mechanism requires both a highly accurate description of the pentacene molecule and a computationally efficient approach that can correctly capture the impact of the *p*-terphenyl host on the active pentacene subsystem. Quantum embedding allows an accurate but expensive hybrid functional to be embedded within a cheaper semi-local functional, for maximum combination of accuracy and efficiency in a DFT-in-DFT framework. In this dissertation we consider the implementation of embedded mean-field theory (EMFT) in the linear-scaling DFT software package ONETEP, enabling hybrid functionals to be used on selected subsystems within a cheaper DFT environment. This approach is validated for several types of molecular systems, including a crystalline structure containing several thousand atoms, demonstrating the potential of the EMFT approach when combined with linear-scaling and verifying the importance of using a large explicit host environment for accurate calculations.

# Contents

<b>Declaration of Originality</b>	<b>i</b>
<b>Declaration of Copyright</b>	<b>i</b>
<b>Acknowledgements</b>	<b>ii</b>
<b>Abstract</b>	<b>iii</b>
<b>List of Figures</b>	<b>viii</b>
<b>List of Tables</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Dissertation outline . . . . .	4
<b>2 Background Theory —</b>	
<b>From Schrödinger to Kohn-Sham</b>	<b>9</b>
2.1 The many-electron problem . . . . .	9
2.1.1 The equations of motion . . . . .	9
2.1.2 The Born-Oppenheimer Approximation . . . . .	11
2.1.3 Variational principle . . . . .	12
2.1.4 Hellmann-Feynman theorem . . . . .	12
2.2 Trial wavefunction solutions . . . . .	13
2.2.1 Hartree-Fock Theory . . . . .	14
2.2.2 Correlation effects — post Hartree-Fock . . . . .	15
2.3 Density functional theory . . . . .	16
2.3.1 Hohenberg-Kohn theorems . . . . .	17
2.3.2 Kohn-Sham equations . . . . .	18
2.4 The exchange-correlation problem . . . . .	20
2.4.1 Properties of the exact functional . . . . .	20
2.4.2 Constructing approximate functionals . . . . .	23
2.4.3 More advanced approaches: range-separated hybrids . . . . .	25
2.5 Density matrix approach to electronic structure . . . . .	25
2.6 Periodic systems . . . . .	27

2.6.1	Bloch's theorem . . . . .	28
2.6.2	Boundary conditions . . . . .	30
2.6.3	Plane waves . . . . .	31
2.6.4	Wannier functions . . . . .	32
2.7	Summary . . . . .	33
<b>3</b>	<b>Multi-level Embedding</b>	<b>37</b>
3.1	Pseudopotentials . . . . .	37
3.1.1	Phillips-Kleinman pseudopotentials . . . . .	38
3.1.2	Norm-conservation and the Kleinman-Bylander expansion . . . .	40
3.2	Quantum embedding . . . . .	41
3.2.1	Frozen Density Embedding (FDE) . . . . .	41
	Freeze-and-thaw iterative optimisation . . . . .	44
3.2.2	Embedding potential and non-orthogonality . . . . .	44
3.3	Approximate embedding schemes . . . . .	45
3.3.1	Continuum solvation models . . . . .	45
3.3.2	Molecular mechanical embedding . . . . .	46
3.3.3	DFT-in-DFT . . . . .	47
3.3.4	Wavefunction-in-DFT . . . . .	48
3.4	Projector embedding . . . . .	48
3.4.1	Embedded Hartree-Fock Equations . . . . .	49
3.4.2	Many projector embedding . . . . .	50
3.4.3	Huzinaga embedding . . . . .	51
3.5	Embedded Mean-Field Theory (EMFT) . . . . .	51
3.5.1	Block orthogonalisation . . . . .	53
3.6	Summary . . . . .	55
<b>4</b>	<b>Linear-Scaling DFT and ONETEP</b>	<b>59</b>
4.1	Nearsightedness and linear-scaling . . . . .	60
4.1.1	Density matrices with non-orthogonal basis sets . . . . .	61
4.1.2	Constraining the density matrix . . . . .	63
	Purification transformations and penalty functionals . . . . .	63
	LNV method . . . . .	64
	Normalisation . . . . .	65
	Optimising the support functions . . . . .	65
4.2	The ONETEP approach . . . . .	66
4.2.1	Linear-scaling with ONETEP . . . . .	67
4.2.2	Total Energy Optimisation in ONETEP . . . . .	68
4.2.3	NGWF and density kernel gradients . . . . .	68

4.3	Relevant ONETEP functionalities . . . . .	71
4.3.1	Forces and geometry optimisation . . . . .	71
4.3.2	Conduction NGWF optimisation . . . . .	71
4.3.3	Hartree-Fock exchange . . . . .	73
4.3.4	Excitation energies from ONETEP . . . . .	73
4.4	Summary . . . . .	75
<b>5</b>	<b>Influence of the crystal host on excitons in pentacene derivatives</b>	<b>80</b>
5.1	Background and Motivation . . . . .	80
5.2	Chromophores and the crystal host . . . . .	82
5.2.1	Choosing the cluster . . . . .	83
5.2.2	Pentacene-based derivatives . . . . .	84
5.3	Excited state methods . . . . .	86
5.4	Vacuum results . . . . .	87
5.4.1	Natural bond orbital (NBO) analysis . . . . .	87
5.4.2	Vacuum excitations . . . . .	89
5.5	Host effects . . . . .	90
5.5.1	Discussion . . . . .	91
5.6	Maser candidates . . . . .	94
5.7	Summary . . . . .	95
<b>6</b>	<b>Freeze-and-thaw embedding in ONETEP</b>	<b>100</b>
6.1	Embedding and linear-scaling DFT . . . . .	100
6.1.1	Embedding in ONETEP . . . . .	101
6.1.2	NGWF gradient for subsystem optimisation . . . . .	103
6.2	Results — iterative subsystem NGWF optimisation . . . . .	106
6.2.1	Frozen density embedding (FDE) . . . . .	108
6.2.2	Excitation energies with FDE . . . . .	111
6.3	Summary . . . . .	113
<b>7</b>	<b>EMFT in ONETEP</b>	<b>116</b>
7.1	Combining EMFT and ONETEP . . . . .	117
7.1.1	Block-orthogonalised EMFT . . . . .	118
7.2	Results — initial tests of EMFT in ONETEP . . . . .	119
7.2.1	NGWF optimisation with EMFT . . . . .	120
7.2.2	Kernel-only optimisation . . . . .	122
7.2.3	Block orthogonalisation . . . . .	123
7.2.4	Hybrid functionals . . . . .	125
7.3	Results — Hydrogenation of pentacene . . . . .	127

7.4	Excitation energies with pentacene derivatives . . . . .	129
7.5	EMFT for large systems . . . . .	132
7.6	Summary . . . . .	135
<b>8</b>	<b>Conclusions</b>	<b>138</b>
<b>A</b>	<b>Derivation of the Hartree-Fock Equations</b>	<b>140</b>
<b>B</b>	<b>Derivation of the Embedded Hartree-Fock Equations</b>	<b>144</b>
B.1	Generalised Embedded Hartree-Fock Equations . . . . .	146
<b>C</b>	<b>Embedding data structures in ONETEP</b>	<b>148</b>
C.1	New data types . . . . .	148
C.2	New module . . . . .	150
C.3	Major datatype modifications . . . . .	151
C.4	Parallel strategies . . . . .	151
C.4.1	Data structure changes for parallel strategies . . . . .	152
C.5	Embedding algorithm . . . . .	153



# List of Figures

2.1	Schematic illustration of the piecewise linear form of the exact exchange-correlation functional. . . . .	22
2.2	Schematic illustration of the behaviour of various exchange-correlation functionals for fractional electron number. . . . .	24
3.1	Schematic illustration of the pseudopotential approximation. . . . .	38
3.2	Schematic illustration of a system divided into two subsystems, an active subregion A and an inactive environment B. . . . .	41
4.1	Sinc function (left) and the periodic bandwidth limited delta function (psinc, right). . . . .	67
4.2	Self-consistent routine for optimisation of total energy in ONETEP. . . .	69
5.1	Pentacene excitation spectrum as part of the room-temperature maser mechanism cycle. . . . .	81
5.2	Various views of the pentacene in <i>p</i> -terphenyl cluster used in this work.	82
5.3	Contribution to HOMO and LUMO states of pentacene by each of its 12 nearest neighbour <i>p</i> -terphenyls in the periodic crystal. . . . .	84
5.4	Dopant molecules considered as part of this study. For brevity, each diazapentacene molecule will be referred to by the shorthand “m,n-dap” throughout this work. . . . .	85
5.5	Changes in the electron populations associated with each atom and acene ring for the diazapentacene molecules compared to pentacene using natural population analysis (NPA). . . . .	88
5.6	Singlet and triplet excitation energies $\Delta E_{S_0 \rightarrow S_1}$ and $\Delta E_{S_0 \rightarrow T_1}$ for each chromophore in vacuum with PBE and OT-LC $\omega$ PBE exchange-correlation functionals. . . . .	90
5.7	Solvatochromic shifts in $\Delta E_{S_0 \rightarrow S_1}$ and $\Delta E_{S_0 \rightarrow T_1}$ by placing each chromophore in a conductor-like polarisable continuum model and an explicit <i>p</i> -terphenyl cluster . . . . .	92
5.8	Proportion of the $S_1$ excited state that consists of the HOMO $\rightarrow$ LUMO Kohn-Sham transition vector for each chromophore in both implicit (CPCM) and explicit (cluster) host. . . . .	93

5.9	Local density of states (LDOS) analysis of the various chromophores in explicit host using OT-LC $\omega$ PBE. . . . .	94
6.1	Left: Water dimer molecule used for embedding tests. The distance between the oxygen atoms is used as a measure of the separation between the molecules. Right: NGWF gradient for water dimer for optimisation of donor and acceptor molecules in freeze-and-thaw (F+T) cycles at a separation of 2.95 Å. . . . .	107
6.2	Absolute difference in the final energies obtained from iterative optimisation of NGWFs in freeze-thaw (F+T) cycles compared to full optimisation of all NGWFs simultaneously for a water dimer. . . . .	108
6.3	Interaction potential for water dimer using pseudoatomic orbitals (PAOs), molecular orbitals (MOs) and non-orthogonal generalised Wannier functions (NGWFs). . . . .	109
6.4	Interaction potential for water dimer via selective optimisation of subsystem NGWFs, using either pseudoatomic orbitals (PAOs, left) or molecular orbitals (MOs, right) as the starting point for the dimer calculations. . . . .	110
6.5	Interaction potential for water dimer via selective optimisation of subsystem NGWFs. . . . .	111
7.1	Interaction potential for a water dimer using EMFT, treating either the acceptor (A) or donor (D) as the active subsystem. . . . .	120
7.2	Behaviour of the NGWF gradient and electron populations during a self-consistent optimisation of the NGWFs and density kernel with EMFT, treating the O $\cdots$ H hydrogen bond as the active system. . . . .	121
7.3	Interaction potential for water dimer. For EMFT calculations, only the density kernel is optimised at the high-level of theory, with the NGWFs fixed at the lower level — see text for more details. Active regions: acceptor (A), donor (D), hydrogen bond (O $\cdots$ H). . . . .	122
7.4	Interaction potential for water dimer using block orthogonalised (BO) and corrected (BO+) NGWFs, treating the O $\cdots$ H bond as the active subsystem. . . . .	123
7.5	PBE-in-LDA interaction potential for water dimer using block orthogonalised (BO) NGWFs, treating the O $\cdots$ H bond as the active subsystem. . . . .	124

7.6	Interaction potential for a water dimer. For EMFT calculations, B3LYP is used for the active subsystem, with the environment treated at either the LDA (left) or PBE (right) levels of theory. All EMFT were performed by optimising the NGWFs at the low-level of theory, before optimising the density kernel with EMFT. Also shown is the effect of block-orthogonalising (BO) the environment NGWFs, with an additional correction for the block-orthogonalisation error (BO+) also included. . . . .	125
7.7	Block-orthogonalised embedded mean-field theory (BO-EMFT) algorithm as implemented in ONETEP. Compare with Figure 4.2 . . . . .	126
7.8	Embedding regions used for the various calculations on pentacene and terminally-hydrogenated pentacene. . . . .	127
7.9	Terminal hydrogenation energy, $\Delta E_{\text{hyd}}$ , of a pentacene molecule as a function of the number of C atoms treated with B3LYP, with all NGWFs optimised at the PBE level of theory. . . . .	128
7.10	Singlet and triplet excitation energies of chromophores in vacuum. Both PBE and B3LYP calculations make use of NGWFs optimised at the PBE level of theory. . . . .	130
7.11	Solvatochromic shift of $T_1$ and $S_1$ states for various dopant molecules in <i>p</i> -terphenyl cluster relative to vacuum. . . . .	131
7.12	The two configurations used in this study for the pentacene in <i>p</i> -terphenyl system. . . . .	132

# List of Tables

5.1	Maximum root mean square (RMS) forces on each chromophore for crystal-optimised geometry in crystal host and vacuum. . . . .	85
5.2	Singlet and triplet excitation energies for pentacene in vacuum from experimental and <i>ab initio</i> studies. . . . .	89
6.1	Vacuum energy differences between the $S_0$ singlet ground state and the $T_1$ triplet ground state for pentacene using non-orthogonal generalised Wannier functions (NGWFs) and pseudoatomic orbitals (PAOs). . . . .	112
6.2	Energy differences between the $S_0$ singlet ground state and the $T_1$ triplet ground state for the pentacene in <i>p</i> -terphenyl cluster host for various basis functions. . . . .	113
7.1	Energy differences between the $S_0$ singlet ground state and the $T_1$ triplet ground state for pentacene, both in vacuum and in the explicit <i>p</i> -terphenyl hosts shown in Figure 7.12. . . . .	133
7.2	Energy differences between the $S_0$ singlet ground state and the $S_1$ singlet excited state for pentacene, both in vacuum and in the explicit <i>p</i> -terphenyl hosts shown in Figure 7.12. . . . .	134

# 1 Introduction

There are very few parts of the modern world that have not been touched by the laser, with applications found in fields far beyond the imaginations of the early researchers who developed the device in the 1950s [1]. The theory behind the laser dates back to the earliest days of quantum mechanics, but the first demonstration of the phenomenon of stimulated emission in 1953 was via microwave emission with ammonia i.e. a microwave laser (maser) [2, 3]. The earliest masers operated at temperatures near absolute zero and required very large magnetic fields, a state of affairs that has scarcely improved and has greatly restricted the applications of maser technology compared to its younger sibling, the laser. In 2012, Oxborrow *et al.* demonstrated the first maser to operate at room temperature and in Earth's magnetic field using a *p*-terphenyl crystal doped with pentacene [4]. Although this represents a significant step forward in maser technology, there are challenges that remain to be overcome before such a device can have any practical applications, such as the need for continuous operation and improvements in efficiency [5]. The driving principle behind the current device is the alignment of the excited singlet and triplet states of pentacene. This facilitates an intersystem crossing which gives rise to a population inversion within the triplet state, enabling a photon with microwave frequency to be emitted. An understanding of this process from first principles thus requires one to accurately describe the excitation energies and the transition rates arising from the intersystem crossing. Experimental and computational studies have shown that the *p*-terphenyl host is not passive and instead has a significant impact on the excitation energies of pentacene [6, 7], especially the singlet energies, due to the delocalisation of these excited states [8]. Therefore a theoretical understanding of the maser requires an *ab initio* description of the excited states of pentacene, as well as the ability to evaluate the impact of the wider crystal.

Since the advent of quantum mechanics a century ago, scientists have made incredible progress in advancing our understanding of the fundamental building blocks of our world. Beginning with the hydrogen atom, our understanding of materials has expanded to an extraordinary degree, with methods since developed to tackle the challenges faced by theorists at many different length scales, such as individual molecules, periodic crystals, point defects in solid and more. However, the fundamental problem in exploiting the laws of quantum mechanics is the difficulty in applying them to anything other than the most simple systems to obtain analytic solutions. The rise

of computational science as a field has facilitated the first principles study of systems much larger than would ever be possible by analytic means. Within the field of electronic structure, this has come about by a combination of theoretical insights that make the solution of quantum systems viable and extensive work aimed at exploiting the increasing power of computers. The modern quantum chemist or condensed matter physicist has a vast array of methods at their disposal, depending on the properties that they wish to probe. For excited states, wavefunction based approaches such as configuration interaction [9] and coupled cluster [10] provide highly accurate results for individual molecules. The *GW* approximation uses an alternative formulation based on the single particle Green's function [11], a method that has become highly popular in recent years due to its accuracy and theoretical rigour [12]. The above approaches, however, are restricted in their applicability by the cost of their calculations. Wavefunction based quantum chemical methods, in particular, suffer from highly inefficient scaling with respect to the number of atoms in the system. Typically, quantum chemical methods based on correlated wavefunctions exhibit scaling of  $N^4$ - $N^7$ , where  $N$  is the number of atoms.

Density functional theory (DFT) [13, 14] has become the method of choice for many computational materials science applications due to its combination of good accuracy and low computational cost [15]. Standard DFT, as formulated by Kohn and Sham in 1965, reduces the complexity of the quantum system to an effective independent electron problem, where all properties of the system can be obtained by knowing a single quantity, the ground state electron density. This approach scales cubically with the number of atoms in the system, a drastic improvement over wavefunction theory (WFT) approaches. Many DFT codes now exploit the power of modern supercomputers, enabling large-scale parallel calculations to be performed on structures consisting of hundreds of atoms.

DFT has been a highly successful method for several decades, a fact encapsulated by the rapid expansion of its use [16]. In the process, it has found use alongside experimental methods to provide a theoretical explanation for observed phenomena, as well as providing a predictive tool to aid experimental research in the search for new materials. Amongst the topics that have been probed using DFT are materials discovery for renewable energy production [17], biological processes [18] and even the composition of the Earth's core [19]. In such cases, it can be used on individual molecules, small composites such as proteins with hundreds of atoms or on infinite periodic crystals, though standard implementations of Kohn-Sham DFT struggle for systems beyond a few hundred atoms as the cubic scaling quickly makes larger calculations impossible. For the last few decades, extensive research has gone into the development of computational methods that can bypass the cubic scaling barrier of DFT and achieve

even greater efficiency which, by exploiting the phenomenon of locality in quantum systems [20, 21], leads us towards the goal of linear-scaling density functional theory (LS-DFT) [22–24]. A variety of formulations now exist that can give rise to LS-DFT, with several electronic structure software programmes now available that can boast this claim (see [25] and references therein). Consequently, the scientific bounds of DFT have expanded to include systems consisting of tens of thousands of atoms, with on-going work to push this limit even higher.

While the success of DFT is beyond question, it has its fair share of well-documented drawbacks [15, 16, 26, 27]. As any electronic structure theorist will proudly declare, the theoretical foundations for DFT, as encapsulated in the Hohenberg-Kohn theorems and the Kohn-Sham equations, provide an exact treatment of the quantum system. However, practical uses of the method require approximations to the treatment of exchange and correlation effects, such as the local density approximation (LDA) whereby it is assumed that these interactions depend solely on the value of the density at a given point in space (and not, for instance, derivatives of the density), which can severely impact the accuracy of DFT calculations. Nowhere is this more prominently seen than in band structure calculations, in particular the determination of the band gap, a central quantity in determining whether a given solid will be a metal, insulator or semiconductor [15]. The gold standard *ab initio* method in this field is *GW*, which frequently captures band gaps that are in line with experimental observation, while DFT with LDA habitually underestimates the band gap (if it predicts one at all). More sophisticated approximations to exchange-correlation effects than LDA do exist [28–30], but these also have their drawbacks and systematic improvements are not always forthcoming. The result is that DFT by itself can give qualitatively incorrect behaviour for electronic structure properties due to the severity of the approximations to exchange-correlation.

In theory, excited state properties can be obtained with DFT if one knows the correct functional form, but in practice approximations to exchange-correlation effects render these beyond the scope of ground state DFT calculations. One approach to deal with this is to allow time evolution of the Kohn-Sham DFT equations, giving rise to time-dependent DFT (TDDFT) [31], which preserves several of the benefits of DFT, such as connection to the electron density and moderate computational cost. Indeed, linear scaling approaches to TDDFT have recently been implemented, allowing for the treatment of excited states in systems with thousands of atoms [32–34]. However, many common implementations of TDDFT exclude history dependence from the system which has further ramifications in terms of accuracy, which for small molecules often lags behind WFT approaches. Methods do exist for the use of exact exchange in DFT/TDDFT calculations, but these can be expensive and restrict the size of the system

that one can simulate.

In recent years, there has been growing interest in quantum embedding methods that seek to combine different levels of theory to solve quantum mechanical problems [35, 36]. The theoretical foundations for this have been well established [37, 38] and current research focuses on optimising the effectiveness of such techniques. In theory, quantum embedding enables one to combine a high level wavefunction-based method to examine part of a system, while using a more efficient approach such as DFT to treat the remaining environment (WFT-in-DFT) [39]. Similar approaches have long been exploited for the study of solvents by combining quantum mechanical methods (QM) for studying the solute with classical molecular mechanical (MM) treatments of the solvent, known as QM/MM [40]. Quantum embedding takes this to the next level by allowing an explicit quantum mechanical description of the environment. Using LS-DFT, this would enable calculations to be performed on systems with thousands of atoms while still providing a very accurate description of phenomena in the embedded subsystem. In particular, embedded mean-field theory (EMFT) [41] is a form of quantum embedding that is well-suited to the density matrix approach to linear-scaling [42]. Examples of potential applications include doped molecular crystals, such as the pentacene in *p*-terphenyl structure used for the maser, and large proteins with excitations that are highly localised in the structure.

This thesis details the investigation of environment effects on pentacene and subsequently the implementation of quantum embedding within the linear-scaling ONETEP programme [43, 44], to enable quantum embedding approaches to be applied in systems containing thousands of atoms.

## 1.1 Dissertation outline

In Chapter 2 we begin by outlining the treatment of many electron quantum mechanical systems and in particular some of the computational approaches used to tackle these systems, building up basic approximations to the wavefunction before introducing the principles of DFT. We discuss the challenges that exist with performing DFT calculations for real systems, followed by extensions of the DFT approach to periodic systems and density matrices.

In Chapter 3 we discuss the basics of pseudopotential theory, which provides the context behind embedding methods. This is followed by a discussion of embedding from a DFT perspective, including the challenges therein and a variety of methods for embedding systems, including EMFT.

Chapter 4 provides an introduction to LS-DFT, with a focus on density matrix methods to achieve linear-scaling and their implementation within ONETEP. Also included



here are details of the relevant pre-existing functionalities within ONETEP that we will exploit later. This is the final literature review chapter, with all material presented after this being produced as part of this work.

Chapter 5 gives a detailed discussion of our studies of the pentacene in *p*-terphenyl and related systems for the maser using both ground state methods and TDDFT. We outline the challenges faced with providing reliable results for such systems and discuss the limitations of the pre-existing methods for *ab initio* studies of such systems.

Chapter 6 explores our early work with an embedding toy model within ONETEP, examining the theoretical implementation of freeze-and-thaw and frozen density embedding approaches, as well as the viability of a full quantum embedding approach.

Chapter 7 introduces the complete embedding structures as currently implemented in the ONETEP package using embedded mean-field theory (EMFT), with corresponding tests of the method and application to systems containing thousands of atoms.

Finally, Chapter 8 concludes the results presented earlier and outlines proposals for future work.

## References

- [1] A. L. Schawlow and C. H. Townes. In: *Phys. Rev.* 112 (6 1958), pp. 1940–1949. DOI: [10.1103/PhysRev.112.1940](https://doi.org/10.1103/PhysRev.112.1940).
- [2] J. P. Gordon, H. J. Zeiger, and C. H. Townes. In: *Phys. Rev.* 95 (1 1954), pp. 282–284. DOI: [10.1103/PhysRev.95.282](https://doi.org/10.1103/PhysRev.95.282).
- [3] J. P. Gordon, H. J. Zeiger, and C. H. Townes. In: *Phys. Rev.* 99 (4 1955), pp. 1264–1274. DOI: [10.1103/PhysRev.99.1264](https://doi.org/10.1103/PhysRev.99.1264).
- [4] Mark Oxborrow, Jonathan D. Breeze, and Neil M. Alford. In: *Nature* 488.7411 (2012), pp. 353–356. ISSN: 0028-0836. DOI: [10.1038/nature11339](https://doi.org/10.1038/nature11339).
- [5] Jonathan Breeze et al. In: *Nat. Commun.* 6.6215 (2015). DOI: [10.1038/ncomms7215](https://doi.org/10.1038/ncomms7215).
- [6] J. Köhler, A.C.J. Brouwer, E.J.J. Groenen, and J. Schmidt. In: *Chemical Physics Letters* 250.1 (1996), pp. 137–144. ISSN: 0009-2614. DOI: [http://dx.doi.org/10.1016/0009-2614\(95\)01393-8](http://dx.doi.org/10.1016/0009-2614(95)01393-8).
- [7] E. Heinecke, D. Hartmann, R. Müller, and A. Hese. In: *The Journal of Chemical Physics* 109.3 (1998), pp. 906–911. DOI: <http://dx.doi.org/10.1063/1.476631>.
- [8] Paul M. Zimmerman, Zhiyong Zhang, and Charles B. Musgrave. In: *Nat Chem* 2.8 (2010), pp. 648–652.
- [9] Péter G. Szalay et al. In: *Chemical Reviews* 112.1 (2012). PMID: 22204633, pp. 108–181. DOI: [10.1021/cr200137a](https://doi.org/10.1021/cr200137a).
- [10] Rodney J. Bartlett and Monika Musiał. In: *Rev. Mod. Phys.* 79 (1 2007), pp. 291–352. DOI: [10.1103/RevModPhys.79.291](https://doi.org/10.1103/RevModPhys.79.291).
- [11] Lars Hedin. In: *Phys. Rev.* 139 (3A 1965), A796–A823. DOI: [10.1103/PhysRev.139.A796](https://doi.org/10.1103/PhysRev.139.A796).
- [12] Giovanni Onida, Lucia Reining, and Angel Rubio. In: *Rev. Mod. Phys.* 74 (2 2002), pp. 601–659. DOI: [10.1103/RevModPhys.74.601](https://doi.org/10.1103/RevModPhys.74.601).
- [13] P. Hohenberg and W. Kohn. In: *Phys. Rev.* 136 (3B 1964), B864–B871. DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- [14] W. Kohn and L. J. Sham. In: *Phys. Rev.* 140 (4A 1965), A1133–A1138. DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133).
- [15] Feliciano Giustino. *Materials Modelling using Density Functional Theory*. Department of Materials, Oxford University: Oxford University Press, 2014.
- [16] K. Burke. In: 109.1 (), pp. 9982–9985. ISSN: 00219606. DOI: [10.1002/qua](https://doi.org/10.1002/qua).

- [17] Anubhav Jain, Yongwoo Shin, and Kristin A Persson. In: *Nature Reviews Materials* 1 (2016), p. 15004. DOI: [10.1038/natrevmats.2015.4](https://doi.org/10.1038/natrevmats.2015.4).
- [18] Shaun D Wong et al. In: *Nature* 499.7458 (2013), pp. 320–323.
- [19] D Alfè, M.J Gillan, and G.D Price. In: *Earth and Planetary Science Letters* 195.1 (2002), pp. 91 –98. ISSN: 0012-821X. DOI: [https://doi.org/10.1016/S0012-821X\(01\)00568-4](https://doi.org/10.1016/S0012-821X(01)00568-4).
- [20] W. Kohn. In: *Phys. Rev. Lett.* 76 (17 1996), pp. 3168–3171. DOI: [10.1103/PhysRevLett.76.3168](https://doi.org/10.1103/PhysRevLett.76.3168).
- [21] E. Prodan. In: *Physical Review B - Condensed Matter and Materials Physics* 73.8 (2006), pp. 0–3. ISSN: 10980121. DOI: [10.1103/PhysRevB.73.085108](https://doi.org/10.1103/PhysRevB.73.085108).
- [22] Giulia Galli and Michele Parrinello. In: *Phys. Rev. Lett.* 69 (24 1992), pp. 3547–3550. DOI: [10.1103/PhysRevLett.69.3547](https://doi.org/10.1103/PhysRevLett.69.3547).
- [23] Giulia Galli. In: *Current Opinion in Solid State and Materials Science* 1.6 (1996), pp. 864 –874. ISSN: 1359-0286. DOI: [http://dx.doi.org/10.1016/S1359-0286\(96\)80114-8](http://dx.doi.org/10.1016/S1359-0286(96)80114-8).
- [24] Stefan Goedecker. In: *Rev. Mod. Phys.* 71 (4 1999), pp. 1085–1123. DOI: [10.1103/RevModPhys.71.1085](https://doi.org/10.1103/RevModPhys.71.1085).
- [25] D R Bowler and T Miyazaki. In: *Reports on Progress in Physics* 75.3 (2012), p. 036503. DOI: [10.1088/0034-4885/75/3/036503](https://doi.org/10.1088/0034-4885/75/3/036503).
- [26] Aurora Pribram-Jones, David A. Gross, and Kieron Burke. In: *Annual Review of Physical Chemistry* 66.1 (2015). PMID: 25830374, pp. 283–304. DOI: [10.1146/annurev-physchem-040214-121420](https://doi.org/10.1146/annurev-physchem-040214-121420).
- [27] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge, U.K.: Cambridge University Press, 2004.
- [28] John P. Perdew. In: *Phys. Rev. B* 33 (12 1986), pp. 8822–8824. DOI: [10.1103/PhysRevB.33.8822](https://doi.org/10.1103/PhysRevB.33.8822).
- [29] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. In: *Phys. Rev. Lett.* 77 (18 1996), pp. 3865–3868. DOI: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [30] Axel D Becke. In: *J. Chem. Phys* 98 (1993), pp. 5648–5652.
- [31] Erich Runge and E. K. U. Gross. In: *Phys. Rev. Lett.* 52 (12 1984), pp. 997–1000. DOI: [10.1103/PhysRevLett.52.997](https://doi.org/10.1103/PhysRevLett.52.997).
- [32] ChiYung Yam, Satoshi Yokojima, and GuanHua Chen. In: *Phys. Rev. B* 68 (15 2003), p. 153105. DOI: [10.1103/PhysRevB.68.153105](https://doi.org/10.1103/PhysRevB.68.153105).

- [33] T. J. Zuehlsdorff et al. In: *The Journal of Chemical Physics* 139.6, 064104 (2013). DOI: <http://dx.doi.org/10.1063/1.4817330>.
- [34] T. J. Zuehlsdorff, N. D. M. Hine, M. C. Payne, and P. D. Haynes. In: *The Journal of Chemical Physics* 143.20 (2015), p. 204107. DOI: [10.1063/1.4936280](https://doi.org/10.1063/1.4936280).
- [35] Andre Severo Pereira Gomes and Christoph R. Jacob. In: *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* 108 (1 2012), pp. 222–277. DOI: [10.1039/C2PC90007F](https://doi.org/10.1039/C2PC90007F).
- [36] P. Huang and E.M. Carter. In: *Annu. Rev. Phys. Chem.* 59 (2008), pp. 261–290.
- [37] Tomasz Adam Wesolowski and Arie Warshel. In: *The Journal of Physical Chemistry* 97.30 (1993), pp. 8050–8053. DOI: [10.1021/j100132a040](https://doi.org/10.1021/j100132a040).
- [38] Yuriy G. Khait and Mark R. Hoffmann. In: *The Journal of Chemical Physics* 133.4, 044107 (2010). DOI: <http://dx.doi.org/10.1063/1.3460594>.
- [39] N. Govind, Y.A. Wang, A.J.R. da Silva, and E.A. Carter. In: *Chemical Physics Letters* 295.1–2 (1998), pp. 129–134. ISSN: 0009-2614. DOI: [http://dx.doi.org/10.1016/S0009-2614\(98\)00939-7](http://dx.doi.org/10.1016/S0009-2614(98)00939-7).
- [40] Hans Martin Senn and Walter Thiel. In: *Angew. Chem., Int. Ed.* 48 (2009), pp. 1198–1229.
- [41] Mark E. Fornace et al. In: *Journal of Chemical Theory and Computation* 11.2 (2015), pp. 568–580. ISSN: 15499626. DOI: [10.1021/ct5011032](https://doi.org/10.1021/ct5011032).
- [42] D. R. Bowler and M. J. Gillan. In: *Chemical Physics Letters* 355.3-4 (2002), pp. 306–310. ISSN: 00092614. DOI: [10.1016/S0009-2614\(02\)00273-7](https://doi.org/10.1016/S0009-2614(02)00273-7).
- [43] Chris-Kriton Skylaris, Peter D. Haynes, Arash A. Mostofi, and Mike C. Payne. In: *The Journal of Chemical Physics* 122.8, 084119 (2005). DOI: <http://dx.doi.org/10.1063/1.1839852>.
- [44] URL: [www.onetep.org](http://www.onetep.org).

## 2 Background Theory — From Schrödinger to Kohn-Sham

Although the laws of quantum mechanics have been known for nearly a century, the key problem in making use of these laws is the difficulty of applying them to large systems. Analytic solutions are only forthcoming for the most simple systems, such as the hydrogen atom, so to gain any knowledge of real systems we must consider numerical solutions. The rise of computational science as a field has facilitated the first principles study of systems much larger than would ever be possible by analytic means. However, even with the power of modern computers it is extremely difficult to apply the laws of quantum mechanics in all their glory to systems consisting of more than a few atoms as the computational cost becomes prohibitive.

In this chapter, we will briefly revise the fundamentals of quantum theory [1] before seeing how we can construct approximate solutions to the many-body Schrödinger equation. This will lead us naturally into density functional theory (DFT) [2, 3], where we will focus on the challenges in obtaining accurate solutions for real systems. From there we will explore the application of DFT to periodic systems via Bloch's theorem.

### 2.1 The many-electron problem

#### 2.1.1 The equations of motion

The fundamental physics of the interaction of  $N$  electrons and  $M$  nuclei is encapsulated within the time-dependent Schrödinger equation (TDSE),

$$\hat{\mathcal{H}} |\Psi\rangle = i\hbar \frac{\partial |\Psi\rangle}{\partial t}, \quad (2.1)$$

where  $\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}, t)$  is the many-body wavefunction, which in general can depend on electron position  $\mathbf{r}$ , ion position  $\mathbf{R}$  and time  $t$ , and  $\hat{\mathcal{H}}$  is the many-body Hamiltonian

operator for the system,

$$\begin{aligned} \hat{\mathcal{H}} = & - \sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{I=1}^M \frac{\hbar^2}{2M_I} \nabla_I^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{e^2}{4\pi\epsilon_0} \frac{1}{|r_i - r_j|} \\ & - \sum_{i=1}^N \sum_{I=1}^M \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|r_i - R_I|} + \sum_{I=1}^M \sum_{J>I}^M \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|R_I - R_J|}, \end{aligned} \quad (2.2)$$

where  $Z_I$  is the atomic number for atom  $I$ ,  $M_I$  is its nuclear mass and all other physical constants have their usual meaning. In this equation the first two terms are the kinetic energies of the electrons and nuclei respectively, while the remainder are electron-electron, electron-nucleus and nucleus-nucleus interaction terms. Note that the Hamiltonian itself does not depend on time.

Throughout this work we employ Dirac notation for representation of the wavefunction, such that  $|\Psi\rangle$  is the state vector corresponding to the many-body wavefunction  $\Psi$ , which can be retrieved in the position representation,

$$\Psi(\mathbf{r}, \mathbf{R}) = \langle \mathbf{r}, \mathbf{R} | \Psi \rangle. \quad (2.3)$$

For simplicity throughout this work we will refer to  $\Psi(\mathbf{r}, \mathbf{R})$  and its state vector  $|\Psi\rangle$  interchangeably as the wavefunction, swapping between them as convenient. We also use Hartree atomic units, whereby  $\hbar = m_e = e = 4\pi\epsilon_0 = 1$ , and drop explicit reference to these physical constants, which will make the notation much tidier. The reader is warned, however, that these physical constants are still present implicitly throughout this work, along with the corresponding units. Our task now is to find solutions to (2.1). We begin by performing a separation of variables upon the wavefunction to partition it into position- and time-dependent functions,

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}, t) = \Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) T(t). \quad (2.4)$$

Recalling that the Hamiltonian (2.2) is a time-independent quantity, substituting this result into (2.1) gives us two separate equations for  $\Phi$  and  $T$ ,

$$\hat{\mathcal{H}}\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = \epsilon\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}), \quad (2.5)$$

$$i\frac{\partial T(t)}{\partial t} = \omega T(t), \quad (2.6)$$

The time-dependent equation has solutions of the form  $e^{-i\omega t}$ , such that the wavefunction can be written as

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}, t) = \Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) e^{-i\omega t}. \quad (2.7)$$

The task of solving the many-body problem has thus been reduced down to finding solutions to the time-independent Schrödinger equation (TISE, 2.5). However, analytical solutions to this set of equations are only forthcoming for the most basic of physical systems, such as the hydrogen atom and other single-electron ions such as  $\text{He}^+$  and  $\text{Li}^{2+}$ . In order to determine solutions for more complex systems, we must consider approximations to the many-body picture.

### 2.1.2 The Born-Oppenheimer Approximation

Given that the nuclei are considerably more massive than electrons in an atom, it is reasonable to consider the motion of the nuclei to be slow compared to that of the electrons. We thus assume that the electrons respond instantaneously to any change in the nuclear configuration and neglect the kinetic energy of the nuclei. This is known as the Born-Oppenheimer Approximation [4] and we will apply it throughout this work.

As a consequence, the electronic wavefunction is assumed to depend only parametrically on the nuclear positions, which enables us to perform yet another separation of variables,

$$\Phi(\{\mathbf{r}_i\}, \{\mathbf{R}_I\}) = \psi(\{\mathbf{r}_i\}; \{\mathbf{R}_I\}) \chi(\{\mathbf{R}_I\}), \quad (2.8)$$

where  $|\psi\rangle$  is the many-electron wavefunction, which satisfies the equation

$$\hat{H} |\psi\rangle = E |\psi\rangle, \quad (2.9)$$

where the Born-Oppenheimer Hamiltonian, in Hartree units, is

$$\hat{H} = -\sum_{i=1}^N \frac{\nabla_i^2}{2} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|r_i - r_j|} - \sum_{i=1}^N \sum_{J=1}^M \frac{Z_I}{|r_i - R_J|} + \sum_{I=1}^N \sum_{J>I}^M \frac{Z_I Z_J}{|R_I - R_J|}. \quad (2.10)$$

A similar equation can be written down for the nuclear wavefunction  $|\chi\rangle$  [5]. The role of the nuclei in the Born-Oppenheimer Approximation can be thought of as an external potential acting on the electrons, a concept which becomes clearer if we rewrite the Hamiltonian as

$$\hat{H} = \hat{T}(\mathbf{r}) + \hat{V}_{ee}(\mathbf{r}) + \hat{V}_{\text{ext}}(\mathbf{r}; \mathbf{R}) + E_{II}(\mathbf{R}), \quad (2.11)$$

where  $\hat{T}$  is the kinetic energy operator,  $\hat{V}_{ee}$  is the electron-electron interaction term,  $\hat{V}_{\text{ext}}$  is the external potential resulting from the nuclei acting on the electrons and  $E_{II}$  is the (constant) ion-ion interaction energy.

### 2.1.3 Variational principle

The average or expectation value  $a$  of any quantum operator  $\hat{A}$  can be expressed in the form

$$a[\Psi] = \frac{\langle \Psi | \hat{A} | \Psi \rangle}{\langle \Psi | \Psi \rangle}, \quad (2.12)$$

where the square brackets indicate that  $a[\Psi]$  is a *functional* of the wavefunction [6] i.e.  $a[\Psi]$  maps a function  $|\Psi\rangle$  onto a number, just as the function  $f(x)$  maps a number  $x$  onto another number. The mathematical details of functionals are left to the interested reader; for our purposes we merely note that the rules of variational calculus using functionals apply in much the same way as for functions.

Now consider a Hamiltonian  $\hat{H}$  with eigenstates  $\{\Psi_i\}$  and corresponding eigenvalues  $\{E_i\}$ , such that  $|\Psi_0\rangle$  is the true ground state wavefunction with energy  $E_0$ . Then the variational principle states that for a given state  $|\Psi\rangle$  the expectation value of the Hamiltonian must satisfy [6]

$$E[\Psi] = \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0, \quad (2.13)$$

where the equality holds if and only if  $|\Psi\rangle = |\Psi_0\rangle$ . In other words, the ground state energy is a lower bound to the energy expectation value for a trial wavefunction  $|\Psi\rangle$ . This provides a useful method for finding the ground state solution to the Schrödinger equation as we can start from an initial guess for  $|\Psi\rangle$  and minimise  $E[\Psi]$  with respect to  $|\Psi\rangle$ . By the variational principle, the wavefunction that yields the ground state energy  $E_0$  will be the true ground state many-body wavefunction  $|\Psi_0\rangle$ .

### 2.1.4 Hellmann-Feynman theorem

Consider the wavefunction  $\Psi(\lambda)$  that has an implicit dependence on the parameter  $\lambda$ . Then the derivative of the energy  $E$  with respect to  $\lambda$  is given as

$$\frac{dE(\lambda)}{d\lambda} = \langle \Psi | \frac{d\hat{H}}{d\lambda} | \Psi \rangle, \quad (2.14)$$

i.e. the derivative of the energy with respect to an external parameter can be calculated as simply the expectation value of the derivative of the Hamiltonian — the dependence of  $|\Psi\rangle$  on  $\lambda$  does not feature. This is known as the Hellmann-Feynman theorem. Most commonly, this theorem is invoked in the calculation of ionic forces. By the



Born-Oppenheimer approximation, the electronic wavefunction depends only parametrically on the ionic coordinates, and so the forces can be calculated as

$$\mathbf{F}_I = -\frac{dE}{d\mathbf{R}_I} = -\langle \Psi | \frac{d\hat{H}}{d\mathbf{R}_I} | \Psi \rangle. \quad (2.15)$$

## 2.2 Trial wavefunction solutions

So far in our discussion we have neglected relativistic effects, which are not included in the TISE. A rigorous treatment of relativistic effects within quantum mechanics would require invocation of the Dirac equation, which is beyond the remit of this work. Nonetheless, one relativistic property of the wavefunction which is not evident from the TISE but is essential for understanding the quantum nature of matter is *spin*. Spin is a property of fundamental particles that arises naturally from the Dirac equation and has crucial experimental consequences such as the Zeeman effect. All particles can be classified based on their spin as either bosons (integer spin) or fermions (half-integer spin). The former includes photons, while the latter includes protons, neutrons and electrons. In order to capture this in our discussion, we must insert spin as an additional quantity upon which the wavefunction depends. Often in the literature the dependence of the wavefunction on position  $\mathbf{r}$  and spin  $\sigma$  is wrapped into a 4-coordinate variable  $\mathbf{x}$ ,

$$\mathbf{x} = \{\mathbf{r}, \sigma\}. \quad (2.16)$$

From the perspective of identifying solutions to the many-body problem, one consequence of this property of electrons is that the wavefunction is antisymmetric with respect to exchange of the space and spin coordinates of any two electrons,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots) = -\Psi(\mathbf{x}_2, \mathbf{x}_1, \dots). \quad (2.17)$$

Any trial solutions to the many-body problem must therefore satisfy this property. This gives rise to the celebrated Pauli exclusion principle for fermions, whereby no two fermions can occupy the same quantum state, and demonstrates that the motion of electrons is correlated [7].

A trial solution for the many-body problem can be constructed from a product of one-electron orbitals  $\{\phi_i\}$ , known as a Hartree product. While such a wavefunction clearly does not satisfy the antisymmetry principle, it can be easily shown that a Slater determinant composed of one-electron orbitals as a linear combination of Hartree products is antisymmetric and thus a valid trial solution. For example, for a two-electron

system, the trial wavefunction would be

$$\begin{aligned}\Psi(\mathbf{x}_1, \mathbf{x}_2) &= \frac{1}{\sqrt{2}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}} (\phi_1(\mathbf{x}_1)\phi_2(\mathbf{x}_2) - \phi_2(\mathbf{x}_1)\phi_1(\mathbf{x}_2)), \\ &= -\Psi(\mathbf{x}_2, \mathbf{x}_1)\end{aligned}\tag{2.18}$$

where the prefactor  $1/\sqrt{2}$  is included to ensure the wavefunction is normalised.

### 2.2.1 Hartree-Fock Theory

As a Slater determinant built from one-electron orbitals satisfies the Pauli exclusion principle, the challenge now is to construct a suitable set of orbitals  $\{\phi_i\}$  that will yield a wavefunction that is optimal in the variational sense. The advance is provided by Hartree-Fock theory [7–9], which involves constructing a single Slater determinant from a set of single electron orbitals,

$$|\Phi_{\text{HF}}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\mathbf{x}_1) & \phi_2(\mathbf{x}_1) & \dots & \phi_N(\mathbf{x}_1) \\ \phi_1(\mathbf{x}_2) & \phi_2(\mathbf{x}_2) & \dots & \phi_N(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\mathbf{x}_N) & \phi_2(\mathbf{x}_N) & \dots & \phi_N(\mathbf{x}_N) \end{vmatrix}.\tag{2.19}$$

The individual orbitals  $\{\phi_i\}$  are solutions to the Hartree-Fock equations,

$$\hat{F}|\phi_i\rangle = \epsilon_i|\phi_i\rangle,\tag{2.20}$$

where  $\hat{F}$  is the Fock operator i.e. the effective Hamiltonian for the single-particle equation

$$\hat{F} \equiv \hat{H}^{\text{core}} + \sum_{j=1}^N (2\hat{J}_j - \hat{K}_j),\tag{2.21}$$

where  $\hat{H}^{\text{core}}$  is the one-electron core Hamiltonian consisting of the kinetic energy contribution and nuclear-nuclear interactions,  $\hat{J}_j$  is the Coulomb operator and  $\hat{K}_j$  is the exchange operator,

$$\hat{J}_j|\phi_i\rangle \equiv |\phi_i\rangle\langle\phi_j|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}|\phi_j\rangle,\tag{2.22}$$

$$\hat{K}_j|\phi_i\rangle \equiv |\phi_j\rangle\langle\phi_j|\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}|\phi_i\rangle,\tag{2.23}$$

where the bracket indicates that integration is only done over one of the position variables,

$$\langle \phi_j | \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} | \phi_i \rangle = \int \frac{\phi_j^*(\mathbf{r}_2) \phi_i(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_2. \quad (2.24)$$

The summation in (2.21) runs over all orbitals since the self-interaction term  $i = j$  in the Coulomb operator is exactly cancelled by the equivalent term in the exchange operator. A detailed derivation of these equations is provided in Appendix A. Notice that the solutions to each of the equations (2.20) implicitly depend on all other orbitals, the effects of which are gathered into a mean-field approach in the form of the Coulomb and exchange operators. The equations (2.20) can be solved self-consistently using the variational principle to iteratively improve the quality of the Hartree-Fock wavefunction  $|\Phi_{\text{HF}}\rangle$  until the total energy converges.

Although the Hartree-Fock method provides an approach to finding approximate solutions to the many-body problem, it has its drawbacks. Most notably, there is no accounting for correlation effects within the Hartree-Fock method beyond the Pauli principle [7]. Consequently, the level of accuracy obtained with this approach is insufficient for understanding chemical properties — for instance, bond energies are severely underestimated by Hartree-Fock compared to experiment [10].

### 2.2.2 Correlation effects — post Hartree-Fock

The Hartree-Fock orbitals are constructed as the ground state of the Fock operator, such that  $|\Phi_{\text{HF}}\rangle$  contains only occupied orbitals constructed from the lowest  $N$  eigenstates of  $\hat{F}$ . One possible extension is to consider the inclusion of higher unoccupied orbitals via the use of creation and annihilation operators on the Hartree-Fock wavefunction. The exact many-body wavefunction can be rigorously expressed as a linear combination of all permutations of single, double, triple etc. excitations [7],

$$|\Psi_{\text{FCI}}\rangle = c_0 |\Phi_{\text{HF}}\rangle + \sum_{i,a} c_i^a |\Phi_i^a\rangle + \sum_{\substack{i>j, \\ a>b}} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{\substack{i>j>k, \\ a>b>c}} c_{ijk}^{abc} |\Phi_{ijk}^{abc}\rangle + \dots, \quad (2.25)$$

$$= \left( c_0 + \sum_{p=1}^n \hat{C}_p \right) |\Phi_{\text{HF}}\rangle, \quad (2.26)$$

where  $i, j, k, \dots$  denote the occupied orbitals replaced by the virtual orbitals  $a, b, c, \dots$  by the action of the operators  $\hat{C}_p$  on the Slater determinant  $|\Phi_{\text{HF}}\rangle$ , and  $c_0 = \langle \Phi_{\text{HF}} | \Psi_{\text{FCI}} \rangle$ . The determinants  $|\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle$  etc. represent single, double etc. excitations from the ground state. This is known as full configuration interaction (FCI) and is the most general wavefunction-based solution possible to the many-body problem. FCI generally yields excellent agreement with experiment and is often treated as an exact result with which

to compare other trial wavefunction methods constructed from equivalent one-electron orbitals. This is due to the inclusion of additional contributions to the correlation energy (beyond the Pauli principle) that are missing from Hartree-Fock, and indeed we can define the correlation energy as

$$E_c \equiv E_{\text{FCI}} - E_{\text{HF}}. \quad (2.27)$$

However, the number of possible configurations scales factorially with the number of electrons in the system, making FCI computationally unfeasible for all but the simplest of physical systems. Instead, quantum chemists often make use of approximate methods whereby the Hartree-Fock solution acts as the zeroth order term in an expansion to include lower-order excited state transitions or perturbative external potentials. Collectively these methods are referred to as *post Hartree-Fock* approaches, and include coupled clusters (CC) [11], which make use of an exponential parametrization of the operators in (2.26), and multi-configurational self consistent field (MCSCF) approaches [12, 13], including its offspring complete active space second-order perturbation theory (CASPT2) [14]. Typically, quantum chemical methods based on correlated wavefunctions exhibit scaling of  $N^4 - N^7$ , where  $N$  is the number of electrons in the system. This renders large-scale calculations on systems including thousands of atoms beyond the reach of such wavefunction-based methods.

## 2.3 Density functional theory

So far, we have only discussed approaches to solving the many-body Schrödinger that involve approximations to the wavefunction. An alternative viewpoint comes courtesy of density functional theory (DFT) [2, 3, 6], whereby all quantities of interest in a quantum system can be expressed in terms of the electronic charge density,

$$n(\mathbf{r}) = \langle \Psi | \hat{n} | \Psi \rangle = N \int \prod_{i=2}^N d\mathbf{r}_i |\Psi(\mathbf{r}, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N)|^2. \quad (2.28)$$

where  $\hat{n}$  is the density operator. Equation (2.28) illustrates that, while the many-body wavefunction is a function of  $3N$  coordinates,  $n(\mathbf{r})$  depends on only three, greatly reducing the complexity of the problem. The first density functional approach was provided by the Thomas-Fermi model [15, 16], whereby the kinetic energy is calculated as a local functional of the electron density, modelled using the density of a set of non-interacting electrons in a homogeneous gas [5]. Although this approach is too crude to be used in practical calculations, this idea of describing energy quantities as functionals of the electron density will in fact be fundamental for much of the work to come in

the remainder of this thesis.

### 2.3.1 Hohenberg-Kohn theorems

The theoretical basis required for DFT is motivated by the two Hohenberg-Kohn theorems [2]. To begin, we write the Hamiltonian in terms of an external potential  $V_{\text{ext}}(\mathbf{r})$ ,

$$\begin{aligned}\hat{H} &= - \sum_{i=1}^N \frac{\nabla_i^2}{2} + \frac{1}{2} \sum_{i=1}^N \sum_{j \neq i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i), \\ &= \hat{T} + \hat{V}_{ee} + \hat{V}_{\text{ext}}.\end{aligned}\tag{2.29}$$

The first theorem states that  $V_{\text{ext}}(\mathbf{r})$  is determined uniquely by the ground state density  $n_0(\mathbf{r})$ . Consequently, the Hamiltonian  $\hat{H}$  is also uniquely determined, which yields the exact many-body wavefunction for the system, from which it is possible to obtain all ground and excited state properties of the quantum system. Thus all properties of the system can be determined from knowledge of the ground state electron density  $n_0(\mathbf{r})$  alone. As a result, we can define the ground state many-body wavefunction to be itself a functional of the density  $\Psi[n]$ , since it is determined as an eigenstate of the Hamiltonian (2.29).

The second theorem states that for any  $V_{\text{ext}}(\mathbf{r})$ , there exists a functional  $E[n]$ , the global minimum of which is precisely the ground state energy of the system, and the electron density that minimises this functional is the ground state density  $n_0(\mathbf{r})$ . From (2.29), we can write the total energy functional as

$$\begin{aligned}E[n] &= \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle + \int V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}, \\ &= F_{\text{HK}}[n] + \int V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r}.\end{aligned}\tag{2.30}$$

where  $F_{\text{HK}}[n]$  is the Hohenberg-Kohn universal functional,

$$F_{\text{HK}}[n] \equiv \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle.\tag{2.31}$$

This quantity is universal in the sense that for any system it is always possible to define a functional in this way, independent of  $V_{\text{ext}}$ , though of course this does not reveal anything about the nature of the functional itself. The second Hohenberg-Kohn theorem provides a DFT equivalent of the variational principle discussed in Section 2.1.3.

### 2.3.2 Kohn-Sham equations

While the Hohenberg-Kohn theorems provide an academic insight into the nature of the electronic structure problem, they do not in and of themselves form a practical method for solving the many body problem, since the form of the universal functional  $F_{\text{HK}}[n]$  is unknown. The critical leap was provided by Kohn and Sham [3] in 1965. They showed that the interacting many-body problem can be rigorously reduced to a set of non-interacting independent electron equations acting in an external field  $V_{\text{KS}}[n]$ , the Kohn-Sham potential. In this section we give a brief derivation of the Kohn-Sham equations as this will help to clarify many key concepts that we will encounter later.

We seek to apply the variational principle to the Hohenberg-Kohn total energy functional (2.30),

$$\frac{\delta}{\delta n(\mathbf{r})} \left[ F_{\text{HK}}[n] + \int V_{\text{ext}}(\mathbf{r})n(\mathbf{r})d\mathbf{r} - \mu \left( \int n(\mathbf{r})d\mathbf{r} - N \right) \right] = 0, \quad (2.32)$$

where  $\mu$  is a Lagrange multiplier introduced to maintain normalisation. In order to progress from this, we must consider the form of the universal functional  $F_{\text{HK}}[n]$ . From (2.31), we have

$$F_{\text{HK}}[n] = \langle \Psi[n] | \hat{T} + \hat{V}_{ee} | \Psi[n] \rangle, \quad (2.33)$$

$$= \langle \hat{T} \rangle + \langle \hat{V}_{ee} \rangle, \quad (2.34)$$

$$= (T_s[n] + \Delta T[n]) + (E_{\text{H}}[n] + \Delta V_{ee}[n]). \quad (2.35)$$

where  $T_s[n]$  is the non-interacting electron kinetic energy,  $E_{\text{H}}[n]$  is the self-interacting Hartree energy of  $n(\mathbf{r})$ ,

$$E_{\text{H}}[n] \equiv \frac{1}{2} \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}', \quad (2.36)$$

and  $\Delta T[n]$  and  $\Delta V_{ee}[n]$  contain the many-body effects that are not captured within the non-interacting picture, which it is hoped will be considerably smaller in magnitude than the non-interacting components. These last two terms consist of the exchange and correlation effects that we briefly discussed in the previous section, collectively called the exchange-correlation functional,

$$E_{\text{XC}}[n] = \Delta T[n] + \Delta V_{ee}[n]. \quad (2.37)$$

We shall explore the nature of  $E_{\text{XC}}$  in the next section. For now, we simply note that it is a functional of the density whose exact form is unspecified.

Substituting these terms into  $F_{\text{HK}}[n]$ , we have

$$\frac{\delta F_{\text{HK}}[n]}{\delta n(\mathbf{r})} = \frac{\delta}{\delta n(\mathbf{r})} \left\{ T_s[n] + E_{\text{H}}[n] + E_{\text{XC}}[n] \right\}, \quad (2.38)$$

$$= \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' + \frac{\delta E_{\text{XC}}[n]}{\delta n(\mathbf{r})}, \quad (2.39)$$

$$= \frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_{\text{H}}(\mathbf{r}) + V_{\text{XC}}(\mathbf{r}). \quad (2.40)$$

The final two terms in (2.40) are defined as the Hartree and exchange-correlation potentials respectively,

$$V_{\text{H}}(\mathbf{r}) = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}', \quad V_{\text{XC}}(\mathbf{r}) = \frac{\delta E_{\text{XC}}[n]}{\delta n(\mathbf{r})}. \quad (2.41)$$

Now substituting this result into (2.32), we obtain

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_{\text{H}}(\mathbf{r}) + V_{\text{XC}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r}) = \mu. \quad (2.42)$$

We now define the *Kohn-Sham potential*,

$$V_{\text{KS}}(\mathbf{r}) \equiv V_{\text{H}}(\mathbf{r}) + V_{\text{XC}}(\mathbf{r}) + V_{\text{ext}}(\mathbf{r}) \quad (2.43)$$

such that

$$\frac{\delta T_s[n]}{\delta n(\mathbf{r})} + V_{\text{KS}}(\mathbf{r}) = \mu. \quad (2.44)$$

But (2.44) is simply the result we would get for a set of non-interacting electrons moving in an effective potential  $V_{\text{KS}}(\mathbf{r})$  with kinetic energy  $T_s[n]$ . Thus we can write down a set of independent electron Schrödinger equations,

$$\left[ -\frac{1}{2} \nabla^2 + V_{\text{KS}}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}), \quad (2.45)$$

where  $\{\psi_i\}$  are the Kohn-Sham orbitals from which the density can be constructed,

$$n(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2. \quad (2.46)$$

These are the Kohn-Sham equations, which can be solved by a self-consistent loop to obtain the true ground state density  $n_0(\mathbf{r})$  that minimises the functional  $E[n]$ , via the second Hohenberg-Kohn theorem. The brevity of this derivation may mask the profound nature of the result — we have shown that it is possible to obtain the ground state density  $n_0(\mathbf{r})$  for a many-body interacting system by solving an equivalent set of

independent electron equations for the fictitious Kohn-Sham system. By the Hohenberg-Kohn theorems, we have thus established a method from which it is possible to derive all the fundamental properties of a quantum system without direct reference to the wavefunction.

## 2.4 The exchange-correlation problem

It is worth noting at this stage that the Kohn-Sham equations (2.45) and (2.46) provide an exact theoretical description of all the properties of the ground state of the quantum system. The only part of the universal functional whose form we have not yet specified is the exchange-correlation functional  $E_{\text{XC}}[n]$ . In order to gain some more insight into the nature of the exchange-correlation energy, let us combine (2.35) and (2.37) and rewrite in terms of  $E_{\text{XC}}[n]$  [5],

$$E_{\text{XC}}[n] = F_{\text{HK}}[n] - (T_s[n] + E_{\text{H}}[n]) \quad (2.47)$$

$$= \left( \langle \hat{T} \rangle - T_s[n] \right) + \left( \langle \hat{V}_{\text{int}} \rangle - E_{\text{H}}[n] \right). \quad (2.48)$$

The latter form illustrates that the exchange-correlation functional is exactly defined as the difference between the true many-body interacting system and the independent electron Kohn-Sham system. If we knew the exact form of this functional, then we would be able to obtain the true ground state of the many-body interacting system via solution of the Kohn-Sham equations. In practice, the form of  $E_{\text{XC}}[n]$  is unknown in general and we must make approximations to this functional in order to obtain useful results with DFT. Nonetheless, decades of investigation into the design of approximate functionals have revealed the importance of fulfilling the known properties of the exact functional, which we shall briefly explore in the next section.

### 2.4.1 Properties of the exact functional

Due to effects such as Pauli repulsion, in real materials electrons tend to be spatially separated, resulting in a region (or hole) around an electron where other electrons are unlikely to be found. The exchange-correlation energy can be represented as an interaction between the electron density and an exchange-correlation hole [17],

$$E_{\text{XC}}[n] = \int \frac{n(\mathbf{r})h_{\text{XC}}(\mathbf{r}, \mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' \quad (2.49)$$

where  $h_{\text{XC}}(\mathbf{r}, \mathbf{r}')$ , the exchange-correlation hole density, accounts for the non-classical exchange and correlation interactions between two electrons simultaneously at  $\mathbf{r}$  and



$\mathbf{r}'$ . This quantity is normalised as

$$\int h_{\text{XC}}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' = -1. \quad (2.50)$$

In effect, the exchange-correlation hole has a net charge deficit of one electron. Equation (2.50) is known as the sum rule, and it is essential that any approximate exchange-correlation functionals satisfy this condition.

The exchange-correlation hole can be split into distinct Kohn-Sham exchange and correlation holes [6],

$$h_{\text{XC}}(\mathbf{r}, \mathbf{r}') = h_{\text{X}}(\mathbf{r}, \mathbf{r}') + h_{\text{C}}(\mathbf{r}, \mathbf{r}'), \quad (2.51)$$

which, in turn, satisfy their own normalisation conditions,

$$\int h_{\text{X}}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' = -1, \quad \int h_{\text{C}}(\mathbf{r}, \mathbf{r}') d\mathbf{r}' = 0. \quad (2.52)$$

Since the hole densities are additive, from (2.49) it is trivial to partition  $E_{\text{XC}}[n]$  into exchange and correlation contributions,

$$E_{\text{XC}}[n] = E_{\text{X}}[n] + E_{\text{C}}[n]. \quad (2.53)$$

$E_{\text{X}}$  is the exchange energy obtained from the Slater determinant of the Kohn-Sham orbitals, while  $E_{\text{C}}$  accounts for all remaining non-classical effects [10]. In general, the exchange energy is the dominant contribution.

Throughout our discussion so far, we have implicitly assumed an integer electron number, though there is no compelling reason for this (aside from physical sanity). The Kohn-Sham DFT setup can be easily extended to consider fractional particle number [18, 19]. The resulting minimum energy curve is a series of straight line segments, with derivative discontinuities at integer  $Z$ , as illustrated in Figure 2.1. For  $Z - 1 < N < Z$ , the slope corresponds to the ionisation energy of the  $Z$ -electron system  $-I_Z$ . This derivative discontinuity is another fundamental property of the exact energy functional.

One consequence of the use of a fictitious Kohn-Sham picture is the loss of physical meaning for the Kohn-Sham eigenvalues. Nonetheless, the eigenvalues do have a distinct meaning within DFT in the form of Janak's theorem [21], which states that the energy eigenvalues can be obtained by variation of the total energy functional with respect to the occupancies  $\{f_i\}$ ,

$$\epsilon_i = \frac{\partial E[n]}{\partial f_i}. \quad (2.54)$$

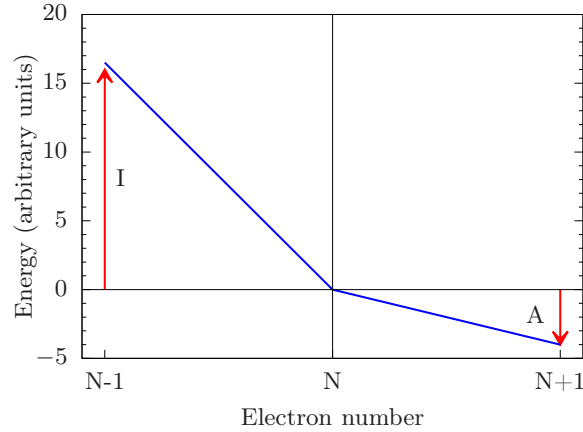


FIGURE 2.1: Schematic illustration of the piecewise linear form of the exact exchange-correlation functional. The slopes of the  $[N - 1, N]$  and  $[N, N + 1]$  segments are the ionisation energy  $I$  and the electron affinity  $A$  for the  $N$  electron system respectively. Based on Figure 2 in [20].

Combined with our previous discussion of the derivative discontinuity, we immediately see that the highest occupied molecular orbital (HOMO) energy is the negative of the ionisation energy,

$$\epsilon_{\text{HOMO}} = -I_Z. \quad (2.55)$$

Unfortunately, the corresponding result for the lowest unoccupied molecular orbital (LUMO) is not quite so simple. Due to the discontinuity in the energy derivatives at integer  $Z$ , there is a step change in the exchange-correlation potential as the number of electrons changes from  $Z - \delta$  to  $Z + \delta$ , which yields a LUMO energy of [22]

$$\epsilon_{\text{LUMO}} = -I_{Z+1} - \Delta_{\text{XC}}, \quad (2.56)$$

where  $\Delta_{\text{XC}}$  is the derivative discontinuity. Consequently, the Kohn-Sham gap is

$$E_{\text{gap}} = I_Z - I_{Z+1} = \epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}} + \Delta_{\text{XC}}. \quad (2.57)$$

Finally, it is worth considering the asymptotic behaviour of the exchange-correlation potential  $V_{\text{XC}}(\mathbf{r})$ . The long-range behaviour of this potential is dominated by the exchange term, leading to an  $r \rightarrow \infty$  limit for finite systems of [23]

$$V_{\text{XC}}(\mathbf{r}) \rightarrow -\frac{1}{r}. \quad (2.58)$$

## 2.4.2 Constructing approximate functionals

In their original paper [3], Kohn and Sham modelled the exchange-correlation functional based on a homogeneous electron gas. This is known as the local density approximation (LDA), whereby the XC functional can be expressed as

$$E_{\text{XC}}^{\text{LDA}}[n] = \int \epsilon^{\text{LDA}}(n(\mathbf{r})) n(\mathbf{r}) d\mathbf{r}, \quad (2.59)$$

where  $\epsilon^{\text{LDA}}(n(\mathbf{r}))$  is the energy density calculated from an equivalent homogeneous electron gas at point  $\mathbf{r}$ . Despite the simple nature of the LDA functional, it has proved highly successful for a variety of chemical systems and properties [24], though it does have its flaws such as strongly overbinding molecules and vastly underestimating band gaps (if it predicts one at all). A seemingly sensible next step for constructing better approximations to the XC functional would be to include higher order corrections to the density, for instance by performing a Taylor expansion of  $n(\mathbf{r})$  [3]. However, tests of such functionals yielded poor results even in comparison to LDA [10]. Fundamentally, this is due to the fact that such functionals violate the sum rule (2.50) [25], which the LDA satisfies, an observation that highlights the importance of satisfying this criterion when constructing new functionals.

More sophisticated approaches to XC approximations follow the strategy of Perdew and Schmidt [26], climbing the rungs of Jacob’s ladder of exchange-correlation functionals from “Hartree World” to the paradise of chemical accuracy. The next rung above LDA is the generalised gradients approximation (GGA), whereby the energy density is now assumed to depend on the first-order derivative of the density [27],

$$E_{\text{XC}}^{\text{GGA}}[n] = \int \epsilon^{\text{GGA}}([n, \nabla n], \mathbf{r}) n(\mathbf{r}) d\mathbf{r}. \quad (2.60)$$

where  $\epsilon^{\text{GGA}}([n, \nabla n], \mathbf{r})$  is the energy density for the GGA functional in question, which is explicitly constructed to ensure that the sum rule and normalisation conditions are satisfied. Generally GGA functionals yield better results than LDA, in particular correcting the severe overbinding of molecules that plagues LDA, though this improvement is not uniform across molecular systems [28].

Climbing another rung of Jacob’s ladder leads us into the realm of meta-GGAs, which introduce a dependence on the kinetic energy density  $\tau$  [29],

$$E_{\text{XC}}^{\text{mGGA}}[n] = \int \epsilon^{\text{mGGA}}([n, \nabla n, \tau], \mathbf{r}) n(\mathbf{r}) d\mathbf{r}, \quad (2.61)$$

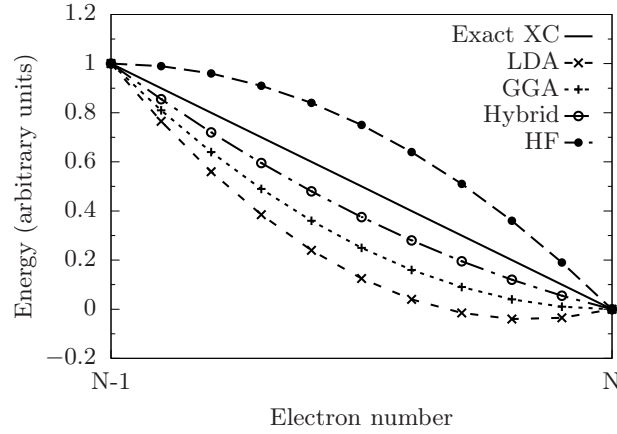


FIGURE 2.2: Schematic illustration of the behaviour of various exchange-correlation functionals for fractional electron number. Based on Figure 2 in [20] and Figure 1 in [31].

where  $\tau(\mathbf{r})$  can be expressed as,

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_i^{N_{\text{occ}}} |\nabla \psi_i(\mathbf{r})|^2, \quad (2.62)$$

in order to reduce the self-correlation error typical of LDA and GGA functionals. While this does generally improve accuracy, typically the benefits are modest as the improvement is primarily contained within the correlation energy  $E_C$ , which, as mentioned earlier, is usually smaller in magnitude than the exchange energy. Importantly, meta-GGAs satisfy the condition that the correlation energy of any one-electron density is zero [30].

All of the functionals discussed so far fail to correct for self-interaction errors, which can only be truly accounted for with exact exchange (i.e. Hartree-Fock) [31]. This motivates the use of hybrid functionals [32], which replace a portion of the exchange-correlation contribution from a semi-local functional with the exact-exchange contribution. For example, the popular B3LYP functional consists of contributions from LDA, GGA and Hartree-Fock [33],

$$\begin{aligned} E_{\text{XC}}^{\text{B3LYP}} = & E_{\text{x}}^{\text{LDA}} + a_0 (E_{\text{x}}^{\text{HF}} - E_{\text{x}}^{\text{LDA}}) + a_x (E_{\text{x}}^{\text{GGA}} - E_{\text{x}}^{\text{LDA}}) \\ & + E_{\text{c}}^{\text{LDA}} + a_c (E_{\text{c}}^{\text{GGA}} - E_{\text{c}}^{\text{LDA}}). \end{aligned} \quad (2.63)$$

where  $a_0 = 0.20$ ,  $a_x = 0.72$  and  $a_c = 0.81$ . Hybrid functionals yield accuracy that is comparable to post-Hartree-Fock methods, though still within the framework of DFT and at a significantly reduced cost [10]. The reason for the success of hybrid functionals can be understood from Figure 2.2, which illustrates the variation of the total energy

for fractional electron number of B3LYP compared to some of the other functionals we have discussed so far. We see that the piecewise linear form of the exact functional is not reproduced with approximate functionals and indeed there is a general tendency for DFT functionals to follow a convex energy curve, while pure Hartree-Fock gives the opposite result, an overestimation of the energy barrier to be crossed in transitioning from  $N$  to  $N + 1$  electrons. Hybrid functionals generally lie in between the extrema of exact exchange and semi-local functionals, and so provide a better fit to the exact functional behaviour. However, like semi-local functionals hybrids still suffer from self-interaction error [31], which can only be truly eradicated with 100% exact exchange.

### 2.4.3 More advanced approaches: range-separated hybrids

One of the most dramatic failures of LDA is the long-range asymptotic behaviour of the exchange-correlation potential, which decays exponentially with distance for finite systems [23], in contrast to the expected  $1/r$  behaviour from (2.58). This motivates the use of range-separated hybrids (RSHs) [34, 35], wherein the short-range exchange interaction is treated using a semi-local functional and the long-range using exact exchange. The length-scale for this changeover is controlled by a range-separation parameter  $\lambda$  with units of inverse length. RSHs show significantly reduced curvature compared to other functionals, bringing them closer to the piecewise linearity of the exact functional and minimising the impact of the derivative discontinuity [22]. This corrects for the asymptotic failures of semi-local functionals, but introduces an arbitrariness in the choice of  $\lambda$  — often RSHs are fitted to empirical data, the result of which is a system-dependence in the choice of functional that can involve many tunable parameters [10], undermining the universality of  $E_{\text{XC}}$ .

Kronik and coworkers [20, 36, 37] have developed a formalism whereby the tuning is performed such that  $\epsilon_{\text{HOMO}}$  is the negative of the ionisation energy, the condition (2.55) of the exact functional. This optimally-tuned RSH (OT-RSH) provides a more solid physical motivation for the tuning of the exchange interaction [38].

## 2.5 Density matrix approach to electronic structure

So far we have discussed the principles of quantum mechanics from the perspective of the wavefunction and density functional theory, whereby it is possible to extract all the properties of a quantum system by treating the electron density as the primary quantity. In this section we will show that there is another perspective from which we can understand a quantum system, namely the one-particle density matrix [39]. The

density matrix formalism will be central to the work laid out in this thesis, so we will review the important concepts here.

To begin, we define the density operator for an  $N$  electron system in terms of the Kohn-Sham eigenstates  $\{\psi_n\}$  [39, 40],

$$\hat{\rho} = \sum_n f_n |\psi_n\rangle \langle \psi_n|, \quad (2.64)$$

where  $\{f_n\}$  are the occupancies of the Kohn-Sham states. Recall that the Kohn-Sham orbitals are orthonormal i.e.

$$\langle \psi_i | \psi_j \rangle = \delta_{ij}. \quad (2.65)$$

We shall now seek to illustrate the relationship between the formulation of DFT in terms of the Kohn-Sham eigenstates outlined in Section 2.3 and the density matrix formalism by examining the properties of this object. Consider the square of the density operator,

$$\begin{aligned} \hat{\rho}^2 &= \sum_{m,n} f_m f_n |\psi_m\rangle \langle \psi_m | \psi_n \rangle \langle \psi_n|, \\ &= \sum_{m,n} f_m f_n |\psi_m\rangle \delta_{mn} \langle \psi_n|, \\ &= \sum_n f_n^2 |\psi_n\rangle \langle \psi_n|, \end{aligned} \quad (2.66)$$

which is simply  $\hat{\rho}$  if  $f_n^2 = f_n$ , which is trivially true if  $f_n = 1$  (occupied state below the Fermi level) or  $f_n = 0$  (unoccupied, above the Fermi level). In practice, this state of affairs is only achievable for computational methods if there is a bandgap in the system. The property

$$\hat{\rho}^2 = \hat{\rho}, \quad (2.67)$$

is known as *idempotency* and is equivalent to orthonormality of the Kohn-Sham orbitals with occupancies  $f_n = \{0, 1\}$ .  $\hat{\rho}$  is thus a projector onto the space of occupied states. Similarly, the trace of (2.64) can be written as

$$\text{Tr}(\rho) = \sum_i \langle \psi_i | \hat{\rho} | \psi_i \rangle, \quad (2.68)$$

$$\begin{aligned} &= \sum_{i,n} f_n \langle \psi_i | \psi_n \rangle \langle \psi_n | \psi_i \rangle, \\ &= \sum_i f_i = N, \end{aligned} \quad (2.69)$$

i.e. the density matrix is normalised such that

$$\text{Tr}(\rho) = N. \quad (2.70)$$

Next, we will consider the trace of the product of  $\hat{\rho}$  and a quantum operator  $\hat{A}$

$$\text{Tr}(\hat{\rho}\hat{A}) = \sum_n \langle \psi_n | \hat{\rho} \hat{A} | \psi_n \rangle, \quad (2.71)$$

$$= \sum_{n,m} f_n \langle \psi_n | \psi_m \rangle \langle \psi_m | \hat{A} | \psi_n \rangle, \quad (2.72)$$

$$= \sum_{n,m} f_n \delta_{nm} \langle \psi_m | \hat{A} | \psi_n \rangle, \quad (2.73)$$

$$= \sum_{n,m} f_n \langle \psi_n | \hat{A} | \psi_n \rangle, \quad (2.74)$$

$$= \sum_n f_n a_n = \langle \hat{A} \rangle. \quad (2.75)$$

Indeed, this is a general result: the expectation value of any operator in a statistical ensemble can be written as

$$\text{Tr}(\hat{\rho}\hat{A}) = \langle \hat{A} \rangle. \quad (2.76)$$

Finally, we note that the density matrix can be written in the position representation as

$$\rho(\mathbf{r}, \mathbf{r}') = \sum_n f_n \psi_n(\mathbf{r}) \psi_n^*(\mathbf{r}'), \quad (2.77)$$

the diagonal elements of which are

$$\rho(\mathbf{r}, \mathbf{r}) = \sum_n f_n |\psi_n(\mathbf{r})|^2 = n(\mathbf{r}), \quad (2.78)$$

i.e. the electron density can be easily extracted from the diagonal elements of the density matrix. By the Hohenberg-Kohn theorems, all data required to fully describe our system is thus encapsulated within the one-particle density matrix, which we can construct by virtue of the Kohn-Sham mapping from the many-body system to a fictitious independent electron picture.

## 2.6 Periodic systems

Conceptually when we speak of the interactions of electrons it is natural to focus on the case of individual atoms or molecules, however DFT can also be applied to periodic systems. A comprehensive discussion of the theory of periodic systems is provided by

Ashcroft & Mermin [41], but here we will only introduce some of the key concepts that are needed. A crystal may (in an idealistic sense) be represented as a primitive unit cell repeated on an infinite lattice, with the vertices of the primitive unit cell defined by the lattice vectors  $\{\mathbf{a}_i\}$ . Then we define the Bravais lattice to be the set of all points with position vectors of the form

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (2.79)$$

where  $\{n_i\}$  are integers. The reciprocal lattice is defined as the set of wave vectors  $\{\mathbf{K}\}$  that produce plane waves with the periodicity of the Bravais lattice,

$$e^{i\mathbf{K}\cdot(\mathbf{r}+\mathbf{R})} = e^{i\mathbf{K}\cdot\mathbf{r}}, \quad (2.80)$$

$$\Rightarrow e^{i\mathbf{K}\cdot\mathbf{R}} = 1. \quad (2.81)$$

From this it can be seen that the primitive vectors of the reciprocal lattice  $\{\mathbf{b}_i\}$  satisfy

$$\mathbf{b}_i \cdot \mathbf{a}_j = 2\pi\delta_{ij}. \quad (2.82)$$

Now we can construct any wave vector  $\mathbf{k}$  in reciprocal space as

$$\mathbf{k} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3, \quad (2.83)$$

where, in general,  $\{k_i\} \in \mathbb{R}$ .

Crystalline systems can be simulated by choosing a large unit cell (or supercell) to account for the bulk properties of the material, using periodic boundary conditions (PBCs) imposed to avoid spurious edge effects. In practice PBCs are often applied in place of open boundary conditions (OBCs) even for isolated systems, with the simulation cell being sufficiently large to avoid spurious self-interactions between images in neighbouring cells. The supercell approach may require extremely large supercells to accurately capture the behaviour of the material in bulk. Fortunately, we can utilise the power of reciprocal space to greatly reduce the complexity of the problem via Bloch's theorem.

### 2.6.1 Bloch's theorem

Here we will follow the derivation of Ashcroft and Mermin [41]. Let us consider an electron in a perfect lattice as defined above, with Bravais lattice vector  $\mathbf{R}$ . Then the



independent electron Schrödinger equation for this system may be written as

$$\hat{H} |\psi\rangle = \left( -\frac{1}{2} \nabla^2 + V(\mathbf{r}) \right) |\psi\rangle = E |\psi\rangle, \quad (2.84)$$

where  $V(\mathbf{r})$  is a potential with the periodicity of the Bravais lattice,

$$V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}). \quad (2.85)$$

Clearly the Hamiltonian for this system is also periodic. Now let us consider a translation operator  $\hat{T}_{\mathbf{R}}$  which acts on any function that is periodic in the Bravais lattice such that

$$\hat{T}_{\mathbf{R}} f(\mathbf{r}) = f(\mathbf{r} + \mathbf{R}). \quad (2.86)$$

Then

$$\hat{T}_{\mathbf{R}} \hat{H}(\mathbf{r}) \psi(\mathbf{r}) = \hat{H}(\mathbf{r} + \mathbf{R}) \psi(\mathbf{r} + \mathbf{R}) = \hat{H}(\mathbf{r}) \hat{T}_{\mathbf{R}} \psi(\mathbf{r}). \quad (2.87)$$

Thus the Hamiltonian and translation operator commute i.e.  $[\hat{T}_{\mathbf{R}}, \hat{H}] = 0$ . As a consequence they possess a simultaneous set of eigenstates,

$$\hat{H} |\psi\rangle = E |\psi\rangle, \quad (2.88)$$

$$\hat{T}_{\mathbf{R}} |\psi\rangle = c(\mathbf{R}) |\psi\rangle. \quad (2.89)$$

Now

$$\hat{T}_{\mathbf{R}} \hat{T}_{\mathbf{R}'} |\psi\rangle = c(\mathbf{R}) c(\mathbf{R}') |\psi\rangle = c(\mathbf{R} + \mathbf{R}') |\psi\rangle, \quad (2.90)$$

which implies that  $c$  can be represented as an exponential. Defining the eigenvalues of the lattice vectors  $\{\mathbf{a}_i\}$  with corresponding complex numbers  $\{\mathbf{x}_i\}$ ,

$$c(\mathbf{a}_i) = e^{2i\pi x_i}. \quad (2.91)$$

Then, remembering that the lattice vectors can be expressed in the form of (2.79), we have

$$c(\mathbf{R}) = c(\mathbf{a}_1)^{n_1} c(\mathbf{a}_2)^{n_2} c(\mathbf{a}_3)^{n_3} = e^{i\mathbf{k} \cdot \mathbf{R}}, \quad (2.92)$$

where  $\mathbf{k}$  is a reciprocal lattice vector in the form of (2.83),

$$\mathbf{k} = x_1 \mathbf{b}_1 + x_2 \mathbf{b}_2 + x_3 \mathbf{b}_3. \quad (2.93)$$

such that  $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi \delta_{ij}$ . Finally substituting this result into (2.89) we arrive at the first form of Bloch's theorem

$$\hat{T}_{\mathbf{R}} \psi(\mathbf{r}) = \psi(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k} \cdot \mathbf{R}} \psi(\mathbf{r}). \quad (2.94)$$

Let us now write down an ansatz for the wavefunction as a product of a plane wave and a function  $u_{n\mathbf{k}}(\mathbf{r})$  with the periodicity of the lattice,

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \quad (2.95)$$

where we have introduced the subscripts  $n$  and  $\mathbf{k}$  to highlight the dependence on the band index and wave vector respectively. Then

$$\psi_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{r}} e^{i\mathbf{k}\cdot\mathbf{R}} u_{n\mathbf{k}}(\mathbf{r} + \mathbf{R}) = e^{i\mathbf{k}\cdot\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{R}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (2.96)$$

which agrees with (2.94). Thus (2.95) is simply another statement of Bloch's theorem.

### 2.6.2 Boundary conditions

At this point we must consider the boundary conditions for the wavefunction in the crystal. We apply generalised PBCs, also known as Born-von Karman boundary conditions [41],

$$\psi(\mathbf{r} + N_i \mathbf{a}_i) = \psi(\mathbf{r}), \quad i = \{1, 2, 3\}, \quad (2.97)$$

where  $N_i$  is the number of cells in the  $\mathbf{a}_i$  direction and  $N_{\text{cells}} = N_1 N_2 N_3$  is the total number of primitive cells in the crystal i.e. the “ends” of the wavefunction are connected at opposite surfaces. By Bloch's theorem, we have

$$\psi_{n\mathbf{k}}(\mathbf{r} + N_i \mathbf{a}_i) = e^{iN_i \mathbf{k}\cdot\mathbf{a}_i} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (2.98)$$

which, applying (2.97) and (2.83), implies

$$e^{iN_i \mathbf{k}\cdot\mathbf{a}_i} = 1 \quad \Rightarrow \quad e^{2\pi i N_i x_i} = 1, \quad (2.99)$$

which is fulfilled if

$$x_i = \frac{m_i}{N_i}, \quad m_i \in \mathbb{Z}. \quad (2.100)$$

Thus the Bloch wave vectors must be of the form

$$\mathbf{k} = \sum_{i=1}^3 \frac{m_i}{N_i} \mathbf{b}_i. \quad (2.101)$$

For each such wave vector there is associated a volume  $\Delta\mathbf{k}$  defined as

$$\Delta\mathbf{k} = \frac{\mathbf{b}_1}{N_1} \cdot \left( \frac{\mathbf{b}_2}{N_2} \times \frac{\mathbf{b}_3}{N_3} \right) = \frac{1}{N} [\mathbf{b}_1 \cdot (\mathbf{b}_2 \times \mathbf{b}_3)]. \quad (2.102)$$

The quantity in square brackets is the volume of the first Brillouin zone (FBZ), and thus the number of allowed wave vectors is equal to the number of sites in the crystal. Therefore we can restrict our treatment of  $\mathbf{k}$ -space to sampling points within the FBZ. As we approach the limit of a perfect crystal  $N_i \rightarrow \infty$ ,  $\mathbf{k}$  becomes continuous in the FBZ, however the wavefunctions vary smoothly in  $\mathbf{k}$ -space [42], such that the sampling of  $\mathbf{k}$ -space can be discretised via methods such as the Monkhorst-Pack approach [43].

The size of the FBZ is inversely proportional to the size of the real-space unit cell,

$$\Omega_{\text{FBZ}} \propto \frac{1}{\Omega_{\text{cell}}}. \quad (2.103)$$

Thus a larger supercell reduces the size of the corresponding FBZ in reciprocal space, which concurrently reduces the number of  $\mathbf{k}$ -points required. For the purposes of our work, which we seek to apply to very large systems, our supercells will be sufficiently large that we can restrict ourselves to a single  $\mathbf{k}$ -point  $\Gamma$ , where  $\mathbf{k} = 0$ .

### 2.6.3 Plane waves

Since it is possible to Fourier expand any periodic function, such as  $u_{n\mathbf{k}}(\mathbf{r})$ , in terms of a plane wave basis, we can rewrite (2.95) as

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{G}} c_{n\mathbf{G}} e^{i\mathbf{G}\cdot\mathbf{r}}, \quad (2.104)$$

where  $\mathbf{G}$  is the reciprocal lattice vector  $\mathbf{G} = y_1 \mathbf{b}_1 + y_2 \mathbf{b}_2 + y_3 \mathbf{b}_3$ . Thus Bloch's theorem illustrates that plane waves are a natural choice of basis for solving the Kohn-Sham equations in a periodic system. In theory, the sum in (2.104) is over all possible  $\mathbf{G}$ -points, though in practice it is necessary to truncate this expansion. The highest  $\mathbf{G}$ -vector included in the expansion is specified by the kinetic energy cutoff,

$$E_{\text{cut}} = \frac{|\mathbf{k} + \mathbf{G}|^2}{2}. \quad (2.105)$$

This cutoff energy is a critical quantity for the quality of a Kohn-Sham DFT calculation. Traditionally calculations using plane wave implementations of DFT must be converged with respect to this parameter.

Although plane waves follow naturally from our preceding description of Bloch's theorem, they do have their drawbacks. For spatially localised systems a large number of plane waves are required to construct an accurate picture of the system. Additionally, there is a vacuum penalty to be paid for isolated systems since the plane waves expand into the vacuum even though there is, in principle, little or no density outside the system of interest.

### 2.6.4 Wannier functions

An alternative but equivalent approach to the Bloch orbital formalism can be obtained by performing a Fourier transform of the Bloch wavefunctions [44],

$$|\mathbf{R}n\rangle = \frac{V}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} e^{-i\mathbf{k}\cdot\mathbf{R}} |\psi_{n\mathbf{k}}\rangle, \quad (2.106)$$

where  $\mathbf{R}$  is a real-space lattice vector and the integral is computed over the Brillouin zone. The resulting functions are highly localised in space and orthonormal,

$$\langle \mathbf{R}n | \mathbf{R}'n' \rangle = \delta_{\mathbf{R}\mathbf{R}'} \delta_{nn'}, \quad (2.107)$$

and are known as Wannier functions. The  $\exp(-i\mathbf{k}\cdot\mathbf{R})$  term in (2.106) provides a generalised phase factor which shifts the Wannier function by  $\mathbf{R}$  in real space. An inverse Fourier transform can be constructed to produce the Bloch orbitals from the Wannier functions,

$$|\psi_{n\mathbf{k}}\rangle = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} |\mathbf{R}n\rangle. \quad (2.108)$$

Combined with (2.106) this forms a unitary transformation between the Bloch and Wannier functions. Thus the Wannier functions provide an equally valid description of the periodic system as the Bloch orbitals. The price for this spatial localisation is that the Wannier functions are not eigenstates of the Hamiltonian; in effect, we have swapped localisation in energy space (via the Bloch orbitals) for localisation in position space.

It is worth noting that there is a gauge freedom to the Wannier functions that arises from the relationship to the Bloch orbitals [45]. If we replace the Bloch orbital  $|\psi_{n\mathbf{k}}\rangle$  with

$$|\tilde{\psi}_{n\mathbf{k}}\rangle = e^{i\varphi_n(\mathbf{k})} |\psi_{n\mathbf{k}}\rangle, \quad (2.109)$$

where  $\varphi_n(\mathbf{k})$  is any real function that is periodic in reciprocal space, there is no resulting change in our description of the system in the Bloch orbital formalism. Transforming to the Wannier function basis,

$$|\mathbf{R}n\rangle = \frac{V}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} e^{i(\varphi_n(\mathbf{k}) - \mathbf{k}\cdot\mathbf{R})} |\psi_{n\mathbf{k}}\rangle, \quad (2.110)$$

we see that the transformation is not unique due to the presence of the gauge freedom in  $\varphi_n(\mathbf{k})$ , the choice of which will determine the form of the resulting Wannier functions. Consequently, it is important to establish a suitable gauge when attempting

to construct Wannier functions, though no particular choice of gauge can be considered superior. This freedom has given rise to many different schemes for constructing Wannier functions for applications, such as maximally localised Wannier functions (MLWFs) which involve a more general mixing of multiple bands rather than a single band as discussed above. A comprehensive review of Wannier functions and their applications is provided by Marzari et al. [46].

## 2.7 Summary

In this chapter we have provided a brief discussion of the challenges for solving the many-body problem for materials applications. Some of the most accurate approaches for treating ground and excited states of atoms and molecules fall within the realm of post-Hartree-Fock methods, which seek to construct approximate wavefunction solutions. However, the cost of such methods renders them unusable for large-scale calculations consisting of hundreds or thousands of atoms. Density functional theory (DFT) provides an alternative approach by treating the electron density  $n(\mathbf{r})$  as the fundamental quantity of interest, while the Kohn-Sham equations provide a theoretically exact description of the quantum system within an effective independent electron framework. The key challenge in DFT involves constructing good approximation to exchange and correlation effects as encapsulated within the energy functional  $E_{XC}$ , some of which we have discussed in this chapter. Following on from this, we saw how DFT can also be applied to periodic systems via Bloch's theorem, which highlights the suitability of using a plane wave basis to construct the Kohn-Sham states. Equivalently, localised Wannier functions give an equally valid description of periodic systems, a point which we shall return to in later chapters.

## References

- [1] Claude Cohen-Tannoudji, Bernard. Diu, and Franck. Laloe. *Quantum mechanics, volume 1*. Wiley, 1978.
- [2] P. Hohenberg and W. Kohn. In: *Phys. Rev.* 136 (3B 1964), B864–B871. DOI: [10.1103/PhysRev.136.B864](https://doi.org/10.1103/PhysRev.136.B864).
- [3] W. Kohn and L. J. Sham. In: *Phys. Rev.* 140 (4A 1965), A1133–A1138. DOI: [10.1103/PhysRev.140.A1133](https://doi.org/10.1103/PhysRev.140.A1133).
- [4] M. Born and R. Oppenheimer. In: *Annalen der Physik* 389.20 (1927), pp. 457–484. ISSN: 00033804. DOI: [10.1002/andp.19273892002](https://doi.org/10.1002/andp.19273892002).
- [5] Richard M. Martin. *Electronic Structure: Basic Theory and Practical Methods*. Cambridge, U.K.: Cambridge University Press, 2004.
- [6] Robert G. Parr and Weitao Yang. *Density-Functional Theory of Atoms and Molecules*. New York: Oxford University Press, 1989.
- [7] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Revised 1st edition. Mineola, New York: Dover Publications, Inc., 1996.
- [8] Ira N. Levine. *Quantum Chemistry*. 5th. City University of New York, Brooklyn: Prentice Hall, 2000.
- [9] John P. Lowe. *Quantum Chemistry*. 2nd. Department of Chemistry, Pennsylvania State University: Academic Press, Inc., 1993.
- [10] Axel D. Becke. In: *The Journal of Chemical Physics* 140.18 (2014), 18A301. ISSN: 0021-9606. DOI: [10.1063/1.4869598](https://doi.org/10.1063/1.4869598).
- [11] Rodney J. Bartlett and Monika Musiał. In: *Rev. Mod. Phys.* 79 (1 2007), pp. 291–352. DOI: [10.1103/RevModPhys.79.291](https://doi.org/10.1103/RevModPhys.79.291).
- [12] B. Swirles D.R. Hartree, W. Hartree. In: *Phil. Trans. R. Soc. London Ser.* 238.790 (1939), p. 299. DOI: [10.1098/rsta.1939.0008](https://doi.org/10.1098/rsta.1939.0008).
- [13] Björn O. Roos, Peter R. Taylor, and Per E.M. Siegbahn. In: *Chemical Physics* 48.2 (1980), pp. 157 –173. ISSN: 0301-0104. DOI: [http://dx.doi.org/10.1016/0301-0104\(80\)80045-0](http://dx.doi.org/10.1016/0301-0104(80)80045-0).
- [14] Peter Pulay. In: *International Journal of Quantum Chemistry* 111.13 (2011), pp. 3273–3279. ISSN: 1097-461X. DOI: [10.1002/qua.23052](https://doi.org/10.1002/qua.23052).
- [15] L H Thomas. In: *Proc. Cambridge Philos. Soc.* 23.1 (1927), pp. 542–548. ISSN: 0008-1981.

- [16] Enrico Fermi. In: *Rend. Accad. Naz. Lincei* (1927). ISSN: 1434-6001. DOI: [10.1007/BF01351576](https://doi.org/10.1007/BF01351576).
- [17] O. Gunnarsson and B. I. Lundqvist. In: *Physical Review B* 13.10 (1976), pp. 4274–4298. ISSN: 0556-2805. DOI: [10.1103/PhysRevB.13.4274](https://doi.org/10.1103/PhysRevB.13.4274).
- [18] John P. Perdew, Robert G. Parr, Mel Levy, and Jose L. Balduz. In: *Physical Review Letters* 49.23 (1982), pp. 1691–1694. ISSN: 0031-9007. DOI: [10.1103/PhysRevLett.49.1691](https://doi.org/10.1103/PhysRevLett.49.1691).
- [19] John P. Perdew and Mel Levy. In: *Phys. Rev. B* 56.24 (1997), p. 16021. ISSN: 1550-235X. DOI: [10.1103/PhysRevB.56.16021](https://doi.org/10.1103/PhysRevB.56.16021).
- [20] Leeor Kronik, Tamar Stein, Sivan Refaely-Abramson, and Roi Baer. *Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals*. 2012. DOI: [10.1021/ct2009363](https://doi.org/10.1021/ct2009363).
- [21] J.F. Janak. In: *Physical Review B* 18.2 (1978), pp. 7165–7168. ISSN: 0163-1829. DOI: [10.1103/PhysRevB.18.7165](https://doi.org/10.1103/PhysRevB.18.7165).
- [22] Tamar Stein et al. In: *The Journal of Physical Chemistry Letters* 3.24 (2012), pp. 3740–3744. ISSN: 1948-7185. DOI: [10.1021/jz3015937](https://doi.org/10.1021/jz3015937).
- [23] Roi Baer, Ester Livshits, and Ulrike Salzner. In: *Annual Review of Physical Chemistry* 61.1 (2010), pp. 85–109. ISSN: 0066-426X. DOI: [10.1146/annurev.physchem.012809.103321](https://doi.org/10.1146/annurev.physchem.012809.103321).
- [24] R. O. Jones and O. Gunnarsson. In: *Reviews of Modern Physics* 61.3 (1989), pp. 689–746. ISSN: 0034-6861. DOI: [10.1103/RevModPhys.61.689](https://doi.org/10.1103/RevModPhys.61.689).
- [25] John P. Perdew. In: *Physical Review Letters* 55.16 (1985), pp. 1665–1668. ISSN: 00319007. DOI: [10.1103/PhysRevLett.55.1665](https://doi.org/10.1103/PhysRevLett.55.1665).
- [26] John P. Perdew and Karla Schmidt. In: *AIP Conference Proceedings* 577.1 (2001), pp. 1–20. DOI: [10.1063/1.1390175](https://doi.org/10.1063/1.1390175).
- [27] John P. Perdew and Wang Yue. In: *Physical Review B* 33.12 (1986), pp. 8800–8802. ISSN: 01631829. DOI: [10.1103/PhysRevB.33.8800](https://doi.org/10.1103/PhysRevB.33.8800).
- [28] Narbe Mardirossian and Martin Head-Gordon. In: *Molecular Physics* 115.19 (2017), pp. 2315–2372. ISSN: 13623028. DOI: [10.1080/00268976.2017.1333644](https://doi.org/10.1080/00268976.2017.1333644).
- [29] Narbe Mardirossian and Martin Head-Gordon. In: *Journal of Chemical Physics* 142.7 (2015). ISSN: 00219606. DOI: [10.1063/1.4907719](https://doi.org/10.1063/1.4907719).
- [30] James C. Womack, Narbe Mardirossian, Martin Head-Gordon, and Chris-Kriton Skylaris. In: *The Journal of Chemical Physics* 145.20 (2016), p. 204114. ISSN: 0021-9606. DOI: [10.1063/1.4967960](https://doi.org/10.1063/1.4967960).

- [31] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. In: *Journal of Chemical Physics* 125.20 (2006). ISSN: 00219606. DOI: [10.1063/1.2403848](https://doi.org/10.1063/1.2403848).
- [32] Axel D. Becke. In: *The Journal of Chemical Physics* 98.2 (1993), pp. 1372–1377. ISSN: 00219606. DOI: [10.1063/1.464304](https://doi.org/10.1063/1.464304).
- [33] Axel D. Becke. In: *The Journal of Chemical Physics* 98.7 (1993), pp. 5648–5652. ISSN: 0021-9606. DOI: [10.1063/1.464913](https://doi.org/10.1063/1.464913).
- [34] Andreas Savin and Heinz-Jürgen Flad. In: *International Journal of Quantum Chemistry* 56.4 (), pp. 327–332. DOI: [10.1002/qua.560560417](https://doi.org/10.1002/qua.560560417).
- [35] Thierry Leininger, Hermann Stoll, Hans-Joachim Werner, and Andreas Savin. In: *Chemical Physics Letters* 275.3-4 (1997), pp. 151–160. ISSN: 0009-2614. DOI: [10.1016/S0009-2614\(97\)00758-6](https://doi.org/10.1016/S0009-2614(97)00758-6).
- [36] Natalia Kuritz, Tamar Stein, Roi Baer, and Leeor Kronik. In: *Journal of Chemical Theory and Computation* 7.8 (2011), pp. 2408–2415. ISSN: 1549-9618. DOI: [10.1021/ct2002804](https://doi.org/10.1021/ct2002804).
- [37] Tamar Stein, Leeor Kronik, and Roi Baer. In: *Journal of the American Chemical Society* 131.8 (2009), pp. 2818–2820. ISSN: 0002-7863. DOI: [10.1021/ja8087482](https://doi.org/10.1021/ja8087482).
- [38] C. O. Almbladh and U. Von Barth. In: *Physical Review B* 31.6 (1985), pp. 3231–3244. ISSN: 01631829. DOI: [10.1103/PhysRevB.31.3231](https://doi.org/10.1103/PhysRevB.31.3231).
- [39] R McWeeny. In: *Rev. Mod. Phys.* 32.2 (1960), pp. 335–369. ISSN: 0034-6861. DOI: [10.1103/RevModPhys.32.335](https://doi.org/10.1103/RevModPhys.32.335).
- [40] Karl Blum. *Density matrix theory and applications; 3rd ed.* Springer Series on Atomic Optical and Plasma Physics. Berlin: Springer, 2012. DOI: [10.1007/978-3-642-20561-3](https://doi.org/10.1007/978-3-642-20561-3).
- [41] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Philadelphia: Saunders College, 1976.
- [42] L. P. Bouckaert, R. Smoluchowski, and E. Wigner. In: *Phys. Rev.* 50 (1 1936), pp. 58–67. DOI: [10.1103/PhysRev.50.58](https://doi.org/10.1103/PhysRev.50.58).
- [43] Hendrik J Monkhorst and James D Pack. In: 13.12 (1976), pp. 5188–5192. ISSN: 01631829. DOI: [10.1103/PhysRevB.16.1748](https://doi.org/10.1103/PhysRevB.16.1748).
- [44] Gregory H. Wannier. In: *Phys. Rev.* 52 (3 1937), pp. 191–197. DOI: [10.1103/PhysRev.52.191](https://doi.org/10.1103/PhysRev.52.191).
- [45] Nicola Marzari, Ivo Souza, and David Vanderbilt. In: *Psi-K newsletter* 57 (2003), p. 129.
- [46] Nicola Marzari et al. In: *Rev. Mod. Phys.* 84 (4 2012), pp. 1419–1475. DOI: [10.1103/RevModPhys.84.1419](https://doi.org/10.1103/RevModPhys.84.1419).



## 3 Multi-level Embedding

In Chapter 2 we discussed the foundations of *ab initio* approaches to electronic structure theory. Throughout we considered solutions whereby all electrons are treated explicitly. Nonetheless when seeking to describe the properties of materials there are many cases where a smaller subsystem is sufficient to provide a substantial description of the properties of interest. For example, chemical bonding is determined by the configuration of valence electrons in atoms, with closed inner shell electrons and nuclei having little effect on the chemistry in question, while in large proteins photoexcitations tend to be localised to a specific part of the system. In both of these cases, the inactive environment provides essential information regarding the overall electronic structure, but for specific properties a high degree of accuracy is only required for the active subsystem. This principle forms the basis for the use of pseudopotential theory for the first principles description of the electronic structure of atoms and molecules [1–3], and similarly the use of quantum embedding [4, 5].

In this chapter, we begin by exploring the pseudopotential method, followed by an introduction to quantum embedding. From there we will discuss the challenges faced in the application of embedding and the methods that have been developed to enable the use of multiple levels of theory in first principles studies.

### 3.1 Pseudopotentials

In most systems the electronic structure is determined primarily by the valence electrons, with the core electrons providing a negligible contribution. The core wavefunctions are highly spatially localised around the atomic centre which, together with the need to maintain orthogonality with the valence electrons as they are all eigenfunctions of the Hamiltonian, causes the valence electron wavefunctions to oscillate rapidly around the nucleus. This can be extremely difficult to represent computationally in e.g. a plane wave basis, requiring the use of many basis functions and a high cutoff energy (2.105) in order to capture the core behaviour of the electronic wavefunctions.

Pseudopotential theory [1–3, 6] offers an alternative approach by replacing the core orbital terms with an effective potential that describes the effect of the core on the valence electrons. A new set of valence wavefunctions, known as pseudo-wavefunctions,

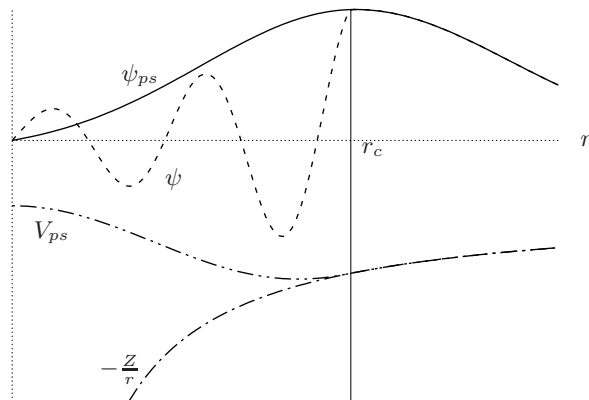


FIGURE 3.1: Schematic illustration of the pseudopotential approximation, based on Figure 5 in [3]. The pseudopotential  $V_{ps}$  converges to  $-Z/r$  beyond the cutoff radius  $r_c$ . Within this core, the pseudopotential wavefunction  $\psi_{ps}$  is smoothed compared to the all-electron wavefunction  $\psi$ , with the two being equivalent in the long-range outside the core.

can be constructed which have the same properties as the all-electron counterparts outside a core cutoff radius  $r_c$  but are much smoother and well-behaved within the core region — see Figure 3.1 for a schematic illustration. There are three key principles which any effective pseudopotential should satisfy:

- Accuracy — the pseudo-wavefunctions should exactly reproduce the properties of the atomic wavefunctions, such as eigenvalues and scattering properties.
- Smoothness — the pseudo-wavefunction should be free of nodes within the cutoff radius  $r_c$ , thus rendering the wavefunction easier to represent and lowering the cutoff energy.
- Transferability — ideally pseudopotentials should be valid for a range of chemical systems, since the differences between such systems should be controlled by the valence electrons.

### 3.1.1 Phillips-Kleinman pseudopotentials

The fundamental theory underlying the pseudopotential approximation was set out by Phillips and Kleinman [1], whereby the strong Coulomb interaction of the nucleus and core electrons are replaced with an approximate potential, based on the orthogonalised plane-wave method of Herring [7]. Consequently, we replace the all-electron valence wavefunctions with a set of pseudo-wavefunctions, which are smooth and do not oscillate rapidly in the core region. Since the core electrons are eigenstates of the

all-electron Hamiltonian in a potential  $V(\mathbf{r})$ ,

$$\hat{H} |\psi_v\rangle = \left( -\frac{\nabla^2}{2} + V(\mathbf{r}) \right) |\psi_v\rangle = E_v |\psi_v\rangle, \quad (3.1)$$

the all-electron valence orbitals  $|\psi_v\rangle$  must be orthogonal to the core states  $|\chi_i\rangle$ . We can then split  $|\psi_v\rangle$  into a smooth, non-orthogonalised pseudo-wavefunction  $|\tilde{\psi}_v\rangle$  plus a contribution to account for the orthogonality of the valence and core orbitals,

$$|\psi_v\rangle = |\tilde{\psi}_v\rangle + \sum_i^{\text{core}} a_i |\chi_i\rangle, \quad (3.2)$$

where  $a_i$  is an expansion coefficient. Since  $\langle \chi_i | \psi_v \rangle = 0$ , we have

$$a_i = -\langle \chi_i | \tilde{\psi}_v \rangle, \quad (3.3)$$

such that the all-electron orbital can be expressed as

$$|\psi_v\rangle = |\tilde{\psi}_v\rangle - \sum_i^{\text{core}} |\chi_i\rangle \langle \chi_i | \tilde{\psi}_v \rangle, \quad (3.4)$$

$$= (1 - \hat{P}_c) |\tilde{\psi}_v\rangle, \quad (3.5)$$

where  $\hat{P}_c$  is a projector onto the space of core orbitals. Then, substituting the expression for the all-electron orbital  $|\psi_v\rangle$  into the Schrödinger equation,

$$\hat{H} (1 - \hat{P}_c) |\tilde{\psi}_v\rangle = E_v (1 - \hat{P}_c) |\tilde{\psi}_v\rangle, \quad (3.6)$$

$$\left[ \hat{H} - \sum_i^{\text{core}} \epsilon_i |\chi_i\rangle \langle \chi_i| \right] |\tilde{\psi}_v\rangle = \left[ E_v - \sum_i^{\text{core}} E_v |\chi_i\rangle \langle \chi_i| \right] |\tilde{\psi}_v\rangle, \quad (3.7)$$

$$\Rightarrow \left[ \hat{H} + \sum_i^{\text{core}} (E_v - \epsilon_i) |\chi_i\rangle \langle \chi_i| \right] |\tilde{\psi}_v\rangle = E_v |\tilde{\psi}_v\rangle, \quad (3.8)$$

$$\Rightarrow [\hat{H} + \hat{V}_{\text{nl}}] |\tilde{\psi}_v\rangle = E_v |\tilde{\psi}_v\rangle, \quad (3.9)$$

where  $\hat{V}_{\text{nl}}$  is an additional non-local potential that accounts for the effect of the core orbitals,

$$\hat{V}_{\text{nl}} = \sum_i^{\text{core}} (E_v - \epsilon_i) |\chi_i\rangle \langle \chi_i|. \quad (3.10)$$

The result is a pseudo-wavefunction  $|\tilde{\psi}_v\rangle$  that is smoother than the all-electron valence orbital  $|\psi_v\rangle$  but possesses the same eigenvalue [2].

### 3.1.2 Norm-conservation and the Kleinman-Bylander expansion

An acceptable pseudopotential must satisfy the conditions set out at the beginning of this section, namely that the pseudo and atomic wavefunctions must match outside the cutoff radius and their eigenvalues must match. In addition, we require that outside the core region  $\psi_v$  and  $\tilde{\psi}_v$  must reproduce the same charge densities. This condition is guaranteed if [3]

$$\int_0^{r_c} \psi_v^*(\mathbf{r})\psi_v(\mathbf{r})d\mathbf{r} = \int_0^{r_c} \tilde{\psi}_v^*(\mathbf{r})\tilde{\psi}_v(\mathbf{r})d\mathbf{r}. \quad (3.11)$$

This is known as norm-conservation and generally pseudopotentials that satisfy this condition ensure high transferability to a range of chemical systems. To preserve scattering properties of the atomic system, norm-conserving pseudopotentials usually take on a non-local potential form which can vary across angular momentum components.

From the Phillips-Kleinman method, we see that a non-local form for the pseudopotential is a requirement to ensure an accurate representation of the core. In general we can construct the pseudopotential in semi-local form as the sum of local and non-local components,

$$\hat{V}_{ps} = \hat{V}_{local} + \hat{V}_{nl}, \quad (3.12)$$

where the non-local potential  $\hat{V}_{nl}$  is composed from available angular momentum channels,

$$\hat{V}_{nl} = \sum_l \sum_{m=-l}^l |lm\rangle \delta\hat{V}_l \langle lm|. \quad (3.13)$$

The cost of building the matrix elements of this semi-local pseudopotential scales with the square of the size of the basis set. For plane-wave basis calculations on complex systems this can be prohibitive since the number of basis functions can be many times greater than the number of electrons in the system. A common method to bypass this is provided by the Kleinman-Bylander representation [8],

$$\hat{V}_{KB} = \hat{V}_{local} + \sum_l \sum_{m=-l}^l \frac{|\phi_{lm}^0 \delta V_l\rangle \langle \delta V_l \phi_{lm}^0|}{\langle \phi_{lm}^0 | \delta V_l | \phi_{lm}^0 \rangle}, \quad (3.14)$$

where  $\{\phi_{lm}^0\}$  are the pseudoatomic wavefunctions and

$$\delta V_l = \hat{V}_{l,NL} - \hat{V}_{local}, \quad (3.15)$$

and  $V_{l,NL}$  is the  $l^{\text{th}}$  angular momentum channel of the non-local pseudopotential. This potential  $V_{KB}$  acts in the same way as the semi-local form  $\hat{V}_{ps}$  but scales only linearly with basis size. We will discuss this point in more detail in Chapter 4.

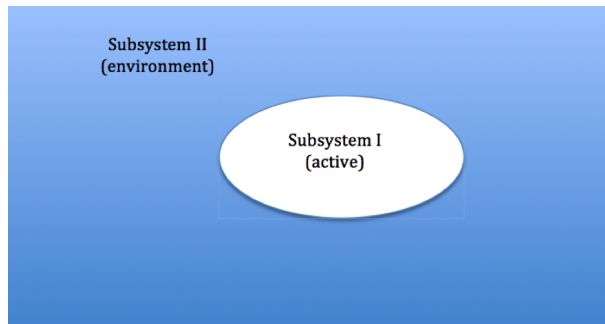


FIGURE 3.2: Schematic illustration of a system divided into two subsystems, an active subregion A and an inactive environment B. With embedding methods, the active subsystem density can be treated at a higher level of theory than the environment, allowing for a mixture of accuracy and efficiency in an *ab initio* calculation.

## 3.2 Quantum embedding

In Chapter 2 we discussed some of the approaches which are available for producing approximate solutions to the many-body problem. Density-functional theory (DFT), combined with the pseudopotential approach outlined above, provides an (in principle) exact treatment of the electronic structure of the system. However, the accuracy of any DFT calculation depends on the exchange-correlation functional that is utilised. Simple semi-local functionals such as LDA are efficient to calculate but, as outlined earlier, can drastically fail to reproduce the exact properties of the true exchange-correlation functional. Accuracy can be improved by using more elaborate hybrid functionals that combine semi-local functionals with Hartree-Fock exact exchange, such as B3LYP. Application of such functionals to complex quantum systems is restricted by the computation cost of such approaches, as the scaling of the exact exchange calculation can be prohibitive. In this section we will outline a theoretically exact approach to partition the system into subsystems that can be computed at different levels of theory [4, 5, 9], as shown in Figure 3.2.

### 3.2.1 Frozen Density Embedding (FDE)

We begin our discussion of quantum embedding from a density functional theory perspective, following the approach of Wesolowski and Warshel [9]. From our discussion in Section 2.3, we know that for a given electron density  $n_{\text{tot}}(\mathbf{r})$ , the total energy functional can be written as

$$E_{\text{tot}}[n_{\text{tot}}] = E_{\text{NN}} + \int n_{\text{tot}}(\mathbf{r})V_{\text{ext}}(\mathbf{r})d\mathbf{r} + E_{\text{H}}[n_{\text{tot}}] + E_{\text{XC}}[n_{\text{tot}}] + T_s[n_{\text{tot}}], \quad (3.16)$$

where  $E_{\text{NN}}$  is the ion-ion interaction energy,  $V_{\text{ext}}(\mathbf{r})$  is the external potential,  $E_{\text{H}}[n_{\text{tot}}]$  is the Hartree energy (2.36),  $E_{\text{XC}}[n_{\text{tot}}]$  is the exchange-correlation functional and  $T_s[n_{\text{tot}}]$  is the kinetic energy functional.

The first step in FDE is to partition the electronic density into two subsets A and B,

$$n_{\text{tot}}(\mathbf{r}) = n_{\text{A}}(\mathbf{r}) + n_{\text{B}}(\mathbf{r}). \quad (3.17)$$

Note that the density can always be partitioned into an arbitrary number of subsystems, but for illustrative purposes we shall restrict our discussion to just two. The total energy functional for the system can then be written as

$$\begin{aligned} E_{\text{tot}}[n_{\text{A}}, n_{\text{B}}] = & E_{\text{NN}} + \int (n_{\text{A}}(\mathbf{r}) + n_{\text{B}}(\mathbf{r})) (V_{\text{ext}}^{\text{A}}(\mathbf{r}) + V_{\text{ext}}^{\text{B}}(\mathbf{r})) d\mathbf{r} \\ & + \frac{1}{2} \int \frac{(n_{\text{A}}(\mathbf{r}) + n_{\text{B}}(\mathbf{r})) (n_{\text{A}}(\mathbf{r}') + n_{\text{B}}(\mathbf{r}'))}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' \\ & + E_{\text{XC}}[n_{\text{A}}] + E_{\text{XC}}[n_{\text{B}}] + E_{\text{XC}}^{\text{nad}}[n_{\text{A}}, n_{\text{B}}] \\ & + T_s[n_{\text{A}}] + T_s[n_{\text{B}}] + T_s^{\text{nad}}[n_{\text{A}}, n_{\text{B}}] \end{aligned} \quad (3.18)$$

where  $E_{\text{XC}}^{\text{nad}}[n_{\text{A}}, n_{\text{B}}]$  and  $T_s^{\text{nad}}[n_{\text{A}}, n_{\text{B}}]$  denote the non-additive kinetic and exchange-correlation energies, respectively,

$$E_{\text{XC}}^{\text{nad}}[n_{\text{A}}, n_{\text{B}}] \equiv E_{\text{XC}}[n_{\text{A}} + n_{\text{B}}] - E_{\text{XC}}[n_{\text{A}}] - E_{\text{XC}}[n_{\text{B}}], \quad (3.19)$$

$$T_s^{\text{nad}}[n_{\text{A}}, n_{\text{B}}] \equiv T_s[n_{\text{A}} + n_{\text{B}}] - T_s[n_{\text{A}}] - T_s[n_{\text{B}}], \quad (3.20)$$

where the first terms on the right-hand side of each equation are the exchange-correlation and kinetic energies for the full system respectively. Collectively these account for the non-linear interactions between the two subsystems [4].

The total energy functional can thus be partitioned as

$$E_{\text{tot}}[n_{\text{A}}, n_{\text{B}}] = E[n_{\text{A}}] + E[n_{\text{B}}] + E_{\text{int}}[n_{\text{A}}, n_{\text{B}}], \quad (3.21)$$

where  $E[n_i]$  is the energy of each subsystem,

$$\begin{aligned} E[n_i] = & E_{\text{NN}}^i + \int n_i(\mathbf{r}) V_{\text{ext}}^i(\mathbf{r}) d\mathbf{r} \\ & + \frac{1}{2} \int \frac{n_i(\mathbf{r}) n_i(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r} d\mathbf{r}' + E_{\text{XC}}[n_i] + T_s[n_i], \end{aligned} \quad (3.22)$$

and  $E_{\text{int}}$  is the interaction energy,

$$E_{\text{int}}[n_A, n_B] = E_{\text{NN}}^{\text{int}} + \int [n_A(\mathbf{r})V_{\text{ext}}^{\text{B}}(\mathbf{r}) + n_B(\mathbf{r})V_{\text{ext}}^{\text{A}}(\mathbf{r})] d\mathbf{r} \\ + \int \frac{n_A(\mathbf{r})n_B(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{\text{XC}}^{\text{nad}}[n_A, n_B] + T_s^{\text{nad}}[n_A, n_B]. \quad (3.23)$$

Here  $E_{\text{NN}}^i$  and  $E_{\text{NN}}^{\text{int}}$  denote the nuclear-nuclear interaction energies within each subsystem  $i$  and between subsystems respectively.

Just as in the conventional DFT approach covered in Section 2.3, we seek to minimise the functional  $E_{\text{tot}}[n_A, n_B]$ . In the Wesolowski-Warshel frozen-density embedding (FDE) approach, we proceed by minimising (3.18) with respect to the electron density  $n_A(\mathbf{r})$  while keeping  $n_B(\mathbf{r})$  fixed (frozen), which yields the embedded Kohn-Sham equations for the orbitals of subsystem A [9]:

$$\left[ -\frac{\nabla^2}{2} + V_{\text{KS}}[n_A](\mathbf{r}) + V_{\text{emb}}^{\text{A}}[n_A, n_B](\mathbf{r}) \right] \psi_i^{\text{A}}(\mathbf{r}) = \epsilon_i^{\text{A}} \psi_i^{\text{A}}(\mathbf{r}), \quad (3.24)$$

where  $V_{\text{KS}}[n_A](\mathbf{r})$  is the usual KS effective potential for A (see (2.43)) and  $V_{\text{emb}}^{\text{A}}[n_A, n_B](\mathbf{r})$  is the effective embedding potential describing the interaction between A and B [4],

$$V_{\text{emb}}^{\text{A}}[n_A, n_B](\mathbf{r}) \equiv \frac{\delta E_{\text{int}}[n_A, n_B]}{\delta n_A(\mathbf{r})}, \quad (3.25)$$

$$= V_{\text{ext}}^{\text{B}}(\mathbf{r}) + V_{\text{H}}^{\text{B}}(\mathbf{r}) + V_T^{\text{A}}[n_A, n_B] + [V_{\text{XC}}^{\text{tot}}(\mathbf{r}) - V_{\text{XC}}^{\text{A}}(\mathbf{r})]. \quad (3.26)$$

This is a local potential that accounts for the impact of the environment on the active subsystem. All the terms contained within will look familiar from conventional DFT, except for the kinetic potential [9],

$$V_T^{\text{A}}[n_A, n_B] \equiv \frac{\delta T_s^{\text{nad}}[n_A, n_B]}{\delta n_A(\mathbf{r})}, \quad (3.27)$$

which arises due to the non-orthogonality of the subsystem orbitals [10]. An accurate description of the embedded system thus requires knowledge of (3.27), which we shall deal with in more detail in the next section. We thus see that using (3.24) we can solve the embedded Kohn-Sham equations for the A subsystem in the presence of a frozen B environment density as accounted for by the embedding potential  $V_{\text{emb}}^{\text{A}}[n_A, n_B](\mathbf{r})$ . The frozen density can be approximated to an acceptable level of accuracy rather than requiring a full self-consistent optimisation, thus providing a potential computational saving.

### Freeze-and-thaw iterative optimisation

There is, of course, nothing special about subsystem A in this derivation. We can repeat this for the B subsystem orbitals, enabling us to construct a pair of Kohn-Sham eigenvalue equations for the orbitals in each subsystem,

$$\left[ -\frac{\nabla^2}{2} + V_{\text{KS}}[n_{\text{B}}](\mathbf{r}) + V_{\text{emb}}^{\text{B}}[n_{\text{A}}, n_{\text{B}}](\mathbf{r}) \right] \psi_i^{\text{B}}(\mathbf{r}) = \epsilon_i^{\text{B}} \psi_i^{\text{B}}(\mathbf{r}), \quad (3.28)$$

where  $V_{\text{emb}}^{\text{B}}$  is defined as

$$V_{\text{emb}}^{\text{B}}[n_{\text{A}}, n_{\text{B}}](\mathbf{r}) = V_{\text{ext}}^{\text{A}}(\mathbf{r}) + V_{\text{H}}^{\text{A}}(\mathbf{r}) + V_T^{\text{B}}[n_{\text{A}}, n_{\text{B}}] + [V_{\text{XC}}^{\text{tot}}(\mathbf{r}) - V_{\text{XC}}^{\text{B}}(\mathbf{r})]. \quad (3.29)$$

Notice that the embedding and kinetic potentials in the coupled Kohn-Sham equations are distinct, and so these equations must be solved in a nested optimisation loop in order to reach the ground state total density  $n_{\text{tot}}(\mathbf{r})$ . This can be done via an iterative freeze-and-thaw process, whereby first we solve for the thawed  $\{\psi_i^{\text{A}}\}$  in the presence of frozen  $n_{\text{B}}(\mathbf{r})$ , then swap frozen and thawed subsystems until self-consistency is achieved. This provides an alternative route to obtaining the ground state density for the full system rather than the standard Kohn-Sham optimisation of all orbitals simultaneously.

### 3.2.2 Embedding potential and non-orthogonality

The key problem in the implementation of frozen-density embedding theory is the description of the embedding potential (3.26). The first two terms, consisting of the effective potential arising from the presence of the orbitals and the density of the B subsystem, are straightforward, accounting for the classical potential of the electrons and nuclei in the frozen environment. The non-classical contributions are contained in the exchange-correlation and kinetic potential terms. As ever, the exchange-correlation potential terms must be approximated since the functional form of this term is unknown. The kinetic potential  $V_T$  (3.27) is a new complication that arises in embedding theory due to the fact that the kinetic energy includes a non-additive contribution between the A and B subsystems, as stated in (3.20). Recall that the exact form of the kinetic energy functional for the many-body system is unknown, however, for a set of orthogonal closed-shell orbitals  $\{\psi_i\}$ , the kinetic energy functional can be written as [11]

$$T_s = \sum_i 2 \langle \psi_i | -\nabla^2/2 | \psi_i \rangle. \quad (3.30)$$



By splitting up the Kohn-Sham equations of A and B, we have made the KS orbitals of the two subsystems eigenstates of different Hamiltonians, thus  $\{\psi_i^A\}$  and  $\{\psi_i^B\}$  are not mutually orthogonal and we cannot write down the kinetic energy of the full system in the form of (3.30), though we can for the A and B subsystems separately. This gives rise to the non-additive contribution to the total kinetic energy, the primary difficulty in describing an accurate embedding potential.

Before continuing, it is worth noting the similarity between  $T_s^{\text{nad}}[n]$  and our discussion of exchange-correlation in Sections 2.3 and 2.4. Recall from (2.37) that  $E_{\text{XC}}[n]$  is defined as the difference between the true many-body kinetic and interaction energies and the non-interacting Kohn-Sham system,

$$E_{\text{XC}}[n] = \left( \langle \hat{T} \rangle - T_s[n] \right) + \left( \langle \hat{V}_{\text{int}} \rangle - E_{\text{H}}[n] \right). \quad (3.31)$$

The non-additive kinetic and exchange-correlation terms are another manifestation of the fact that we do not know the correct form for the exact functionals. Unlike  $E_{\text{XC}}[n]$ , little is known about the true form of  $T_s^{\text{nad}}[n]$ , which adds another layer of complexity to the process of constructing accurate kinetic potentials.

### 3.3 Approximate embedding schemes

Although the FDE method was derived in the realms of DFT, it is applicable for any method which produces an electronic density, from which the embedding potential can be generated. This allows for the rigorous treatment of a physical system at multiple levels of theory as suitable for the various subsystems. In this section we shall discuss some of the more widely used approaches to multi-level embedding, based on the classification presented by Gomes and Jacob [4].

#### 3.3.1 Continuum solvation models

Continuum models [12] ignore the explicit molecular structure of the environment and treat it instead as a solvent with a dielectric constant  $\epsilon^{\text{env}}$ . The active subsystem is embedded within this solvent by constructing a cavity within the dielectric continuum, with the transition from solvent to solute accounted for by electrostatic interactions resulting from induced surface charges surrounding the cavity. The value of the dielectric constant is usually taken from experimental measurements, while the size and shape of the cavity can be tuned to reproduce free-energies for the specific solute-solvent interaction. The charge induced on the cavity surface is known as the apparent surface charge  $\sigma(\mathbf{r})$ , giving rise to an electrostatic potential generated from the closed

surface of the cavity,

$$V_{\text{emb}}^{\text{SC}}(\mathbf{r}) = \int_{\text{Cavity surface}} \frac{\sigma(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^2\mathbf{r}', \quad (3.32)$$

where the integral runs over the cavity surface area. This can be solved computationally by discretising the cavity surface with point charges  $q_s$  at position  $\mathbf{r}_s$  that represent the apparent surface charge  $\sigma(\mathbf{r}_s)$  [4]. Then the embedding potential (3.26) can be represented as

$$V_{\text{emb}}^{\text{solvent}}(\mathbf{r}) \approx \sum_s \frac{q_s [\epsilon^{\text{env}}, n_A(\mathbf{r})]}{|\mathbf{r} - \mathbf{r}_s|}. \quad (3.33)$$

Such methods that make use of finite dielectric constants  $\epsilon^{\text{env}}$  are referred to as polarisable continuum models (PCM), due to the polarisation of the dielectric continuum in this picture. An alternative formulation involves setting the environment dielectric value to be infinite, such that the sum of the solute and solvent electrostatic potentials vanishes at the cavity surface [12]. The finite dielectric solvent result can then be reconstructed via the use of a scaling function that depends on the true  $\epsilon^{\text{env}}$ . This is known as a conductor-like PCM (CPCM), and is the basis of popular solvation methods such as the conductor-like screening model (COSMO) [12, 13]. Generally such methods are of use for large solute-solvent systems, providing an average picture where explicit molecular treatment would require the use of multiple configurations to provide a representative description of environment effects.

### 3.3.2 Molecular mechanical embedding

Instead of the solvent model used in continuum approaches, molecular mechanical (MM) embedding involves a discrete description of the environment structure, with active subsystem interactions accounted for by classical force field models. When combined with an explicit quantum mechanical (QM) description of the active subsystem, this is referred to as QM/MM embedding [14]. The use of classical mechanics for the solvent environment makes such methods very efficient for large systems such as biomolecular structures [15].

Unlike continuum methods, QM/MM enables the explicit consideration of molecular configurations in the environment. The embedding calculation can be performed either as a QM correction to an MM calculation on the full system, whereby no QM-MM coupling is included, or electrostatic interactions can be included directly in the model to account for the response of the active electron density to the presence of the environment. Other more sophisticated QM/MM embedding schemes, such as polarisable embedding and the use of link atoms to account for covalent bonds between the active and environment regions, can be performed and are reviewed extensively elsewhere [15].

### 3.3.3 DFT-in-DFT

As discussed in Chapter 2, practical calculations with DFT can only be performed by making approximations to the exchange-correlation functional, such as the local density approximation (LDA), generalised gradient approximation (GGA) or exact-exchange. Within the FDE formulation, the splitting up of this functional allows the possibility of performing multi-level calculations by constructing each subsystem using different levels of approximation. Both the active and environment subsystems are evaluated within DFT, using a high-level functional for the active region of interest and a low-level functional for the remainder of the system for a DFT-in-DFT embedding framework. The total exchange-correlation energy  $E_{\text{XC}}^{\text{emb}}$  can then be constructed from the subsystem and interaction terms,

$$E_{\text{XC}}^{\text{emb}}[n_{\text{A}} + n_{\text{B}}] = E_{\text{XC}}^{\text{high}}[n_{\text{A}}] + E_{\text{XC}}^{\text{low}}[n_{\text{B}}] + E_{\text{XC}}^{\text{nad}}[n_{\text{A}}, n_{\text{B}}]. \quad (3.34)$$

For semi-local functionals such as LDA or GGA, the non-additive term can be straightforwardly evaluated from (3.34). Embedding using hybrid functionals is conventionally done by treating the non-additive exchange term at the semi-local level [11]. Effectively this amounts to constructing the non-additive exchange-correlation potential at the lower-level of theory, which constitutes an additional approximation [4].

In addition to the exchange-correlation term, approximations must be made to the kinetic potential in order to perform useful calculations due to the non-orthogonality dilemma outlined in Section 3.2.2. The earliest FDE approximations were based on the Thomas-Fermi functional [9], which, as in conventional DFT, could be improved upon by appeal to GGA functionals [16, 17]. Nonetheless, such approximations were only found to be valid for weakly interacting systems such as hydrogen bonds, with covalently bonded systems generally yielding poor results [18, 19]. Much effort has been expended in attempting to construct good embedding potentials by more sophisticated means [4, 5]. In the Optimized Effective Potential formalism (OEP), one constructs a local potential for which the total energy remains stationary for small variations [20] [21]. While there exist many approaches to the embedding problem with OEPs [22] [23], the rigour and extendability of these approaches is questionable, as for finite basis sets the OEP problem is ill-posed [24]. Other methods have sought to establish a rigorous approach to construct a well-defined potential for a range of chemical systems (see [10, 25–27] for examples), but the fundamental problem of constructing an accurate and flexible potential remains elusive, particularly for systems where there is a strong overlap between the active and frozen subsystems.

### 3.3.4 Wavefunction-in-DFT

Govind *et al* extended the original DFT-in-DFT formulation to include the treatment of explicit correlation effects via the use of a quantum chemical approach within a DFT environment [28], giving rise to wavefunction theory in DFT (WFT-in-DFT) embedding. This allows the use of an accurate quantum chemical method such as the post Hartree-Fock formulations discussed in Section 2.2.2 embedded in a cheaper DFT environment.

So far, we have only considered the application of quantum embedding to a description of the ground state system, however the formalism presented above can also be applied to the treatment of excited states. This was confirmed for DFT-in-DFT by Khait and Hoffmann [29] following the same approach as Perdew and Levy to the treatment of excitations with ground-state DFT [30]. These excited states can be seen as the densities that satisfy (3.24) which do not yield the energy global minimum. Khait and Hoffmann also extended this formalism to WFT-in-DFT, showing that both the ground state and a subset of excited states (i.e. the stationary points of the energy functional) can be obtained in principle with embedding [29], opening the door for the treatment of excited states in the WFT region while treating the environment with DFT. Nonetheless, as with DFT-in-DFT the accuracy of such approaches depends critically on producing a decent description of the kinetic potential.

## 3.4 Projector embedding

Quantum mechanical embedding approaches suffer from the need to approximate the kinetic energy potential (3.27). An alternative approach is to consider an explicit orthogonalisation of the embedded subsystem orbitals with respect to the environment, which would eliminate the non-additive term in the kinetic energy and thus remove the need to construct an accurate kinetic potential. Multiple approaches have been developed which make use of this approach [31–34], motivated by the Phillips-Kleinman pseudopotential method discussed in Section 3.1. Indeed, pseudopotential theory can be construed as a type of embedding where the valence orbitals are calculated in the presence of an effective potential that describes the effect of the core orbitals. This connection can be seen in efforts to develop general theories for the separation of multi-electron systems, whereby the core-valence model is only a special case [6, 35, 36]. To illustrate the theory behind this approach, we shall briefly outline how this formalism can work within the context of Hartree-Fock theory.

### 3.4.1 Embedded Hartree-Fock Equations

We begin by considering an embedded system consisting of a set of orbitals  $\{\phi_i^A\}$ , representing the embedded subsystem, within the presence of a set of environment orbitals  $\{\phi_j^B\}$ . The bare Fock operator for this system is simply an extended version of (2.21),

$$\hat{H} = \hat{H}^{\text{core},A} + \sum_{j \in A} \left( 2\hat{J}_j^A - \hat{K}_j^A \right) + \sum_{j \in B} \left( 2\hat{J}_j^B - \hat{K}_j^B \right), \quad (3.35)$$

which includes Coulomb and exchange operators for each subsystem. Initially, we shall assume that the A and B subsystem orbitals are already orthogonal. Thus, a set of equations for the embedded A orbitals can be constructed (see Appendix B for details),

$$(1 - \hat{P}^B)\hat{F}|\phi_i^A\rangle = \epsilon_i |\phi_i^A\rangle. \quad (3.36)$$

where  $\hat{P}^B$  is a projector onto the B orbital subspace,

$$\hat{P}^B \equiv \sum_{j \in B} |\phi_j^B\rangle \langle \phi_j^B|. \quad (3.37)$$

The result is a projected Fock operator, which can easily be made Hermitian to give the Huzinaga equation [33, 36],

$$\left( \hat{F} - \hat{P}^B \hat{F} - \hat{F} \hat{P}^B \right) |\phi_i^A\rangle = \epsilon_i |\phi_i^A\rangle. \quad (3.38)$$

Following the Weeks-Rice method [37], we can generalise this result to the case where the A and B orbitals are not mutually orthogonal by performing a Gram-Schmidt orthogonalisation of the A orbitals,

$$|\tilde{\phi}_i^A\rangle = (1 - \hat{P}^B) |\phi_i^A\rangle. \quad (3.39)$$

Thus the Hartree-Fock equations can be generalised to (Appendix B)

$$(1 - \hat{P}^B)\tilde{F}|\tilde{\phi}_i^A\rangle = \epsilon_i |\tilde{\phi}_i^A\rangle, \quad (3.40)$$

where the tilde above the Fock operator indicates that it is composed of projected orbitals,

$$\tilde{F} = \hat{F}^{\text{core},A} + \sum_{j \in A} \left( 2\tilde{J}_j^A - \tilde{K}_j^A \right) + \sum_{j \in B} \left( 2\hat{J}_j^B - \hat{K}_j^B \right). \quad (3.41)$$

Expanding the  $\{\tilde{\phi}_i^A\}$  in (3.40), we get

$$(1 - \hat{P}^B) \tilde{F} (1 - \hat{P}^B) |\phi_i^A\rangle = \epsilon_i (1 - \hat{P}^B) |\phi_i^A\rangle. \quad (3.42)$$

This can be rewritten as

$$\left(\tilde{F} + \hat{V}_{WR}\right) |\phi_i^A\rangle = \epsilon_i |\phi_i^A\rangle, \quad (3.43)$$

where we have introduced the Weeks-Rice style pseudopotential [37],

$$\hat{V}_{WR} \equiv -\hat{P}^B \tilde{F} - \tilde{F} \hat{P}^B + \hat{P}^B \tilde{F} \hat{P}^B + \epsilon_i \hat{P}^B. \quad (3.44)$$

We thus see that the results from standard pseudopotential theory are easily extendable to produce a projector approach to embedding theory. At this stage we will discuss some of the practical implementations of the above concepts in the literature.

### 3.4.2 Manby projector embedding

Manby *et al.* [32] have developed a method that maintains the spirit of the projection operator embedding but seeks to reduce the complexity of the resulting potential such that it can be applicable to DFT-in-DFT or WFT-in-DFT calculations. Initially, a KS-DFT calculation is performed over the full system. The molecular orbitals are then localised and partitioned between the two subsystems,  $\{\phi_i^A\}$  and  $\{\phi_i^B\}$ , and the density matrices  $\rho_A$  and  $\rho_B$  are formed.

This embedding approach gives rise to a Fock operator of the form

$$\hat{F} = \left(1 - \hat{P}^B\right) \hat{H} \left(1 - \hat{P}^B\right), \quad (3.45)$$

where  $\hat{H}$  is the total Hamiltonian and  $\hat{P}^B$  is a projector onto the orbital subspace of B,

$$\hat{P}^B \equiv \sum_{j \in B} |\phi_j^B\rangle \langle \phi_j^B|. \quad (3.46)$$

The key conceptual leap by Manby *et al.* is the observation that the lower part of the spectrum of this Fock operator coincides with the spectrum of  $\hat{H} + \mu \hat{P}^B$  for  $\mu \rightarrow \infty$  [38]. Thus the Weeks-Rice style embedding pseudopotential arising due to the non-orthogonality of the A and B bases is replaced with a simple projector  $\mu \hat{P}^B$ , where  $\mu$  is a constant.

The effect of  $\mu \hat{P}^B$  is to elevate the energy of the  $i^{\text{th}}$  orbital  $\epsilon_i^B$  in subsystem B to  $\epsilon_i^B + \mu$  which, for  $\mu \rightarrow \infty$ , prevents the electrons in A from accessing the orbitals of B [39], thus enforcing orthogonality. The non-additive kinetic energy is eliminated, greatly simplifying the problem in principle. Tests of this method for a DFT-in-DFT system reproduce the exact DFT results for the full system for  $\mu \rightarrow \infty$  [40], providing a useful measure of the accuracy of any description of the embedding potential. In practice,

the level-shift parameter  $\mu$  must take on values in the range  $10^3 - 10^6 E_H$  to achieve accuracy, though for higher values numerical instability can be problematic [40].

### 3.4.3 Huzinaga embedding

Another projector approach was recently proposed by Hégely *et al.* [33]. Rather than using an *ad hoc* projector to enforce orthogonality, Hégely *et al.* utilise the Huzinaga equation (3.38). Similar to the Manby approach, the molecular orbitals are localised following a regular SCF calculation on the full system such that they can be assigned to individual subsystems. The density matrix (see Section 2.5) can then be straightforwardly written as the sum of the A and B subsystem matrices,

$$\rho^{AB} = \rho^A + \rho^B. \quad (3.47)$$

Then the Huzinaga equations are solved for the active subsystem A, which in matrix form can be expressed as

$$(\mathbf{F} - \mathbf{S}\rho^B\mathbf{F} - \mathbf{F}\rho^B\mathbf{S}) \mathbf{C}^A = \mathbf{S}\mathbf{C}^A\mathbf{E}^A, \quad (3.48)$$

where  $\mathbf{S}$  is the atomic orbital overlap matrix,  $\mathbf{C}^A$  is the matrix of molecular orbital coefficients and  $\mathbf{E}^A$  is a diagonal matrix of the Kohn-Sham orbital eigenvalues of the A subsystem.

It is worth noting that, while derived from different starting points, the Manby and Huzinaga projector methods are in effect simplifications of the general Weeks-Rice pseudopotential method — both can be constructed by removing terms from (3.44). Thus any implementation of a projector embedding approach would warrant some thought regarding the “correct” potential to use, since there appears to be a degree of ambiguity on this matter.

## 3.5 Embedded Mean-Field Theory (EMFT)

The Wesolowski-Warshel FDE picture which forms the basis for all the embedding methods we have discussed so far operates on a straightforward partitioning of the electronic density. Thus, whether we are considering embedding from the realms of Hartree-Fock theory or Kohn-Sham DFT, there is a clear distinction between the molecular orbitals (MOs) that are assigned to each subsystem, since they are eigenfunctions of different Hamiltonians. In practice the physical distinction between subsystems may not always be straightforward, and, as in the Manby and Huzinaga methods [32,



33], a localisation of the MOs may be necessary. This implies a degree of arbitrariness in the subsystem partitioning through the choice of a localisation scheme such as Mulliken charges. In addition, such approaches require that the subsystem particle number be fixed [41], which can be problematic for systems where electron transfer is possible across bonds that are partitioned between subsystems. Some embedding techniques avoid this problem by using the density matrix as their fundamental quantity for performing quantum embedding [42, 43], since the matrix can be easily divided into subsystem blocks and off-diagonal cross-overlap matrices,

$$\rho = \begin{pmatrix} \rho_{AA} & \rho_{AB} \\ \rho_{BA} & \rho_{BB} \end{pmatrix}. \quad (3.49)$$

The off-diagonal blocks facilitate electron transfer between the A and B subsystems.

From the perspective of performing electronic structure calculations, plane wave basis functions (see Section 2.6.3) that can extend throughout the system are not intuitively suitable for embedding. Indeed, an atom-centred basis set such as localised Wannier functions (see Section 2.6.4) would appear to be a much more appropriate basis for embedding calculations since each basis function can be easily assigned to a particular subsystem. In this section we will explore an approach based on a purely mean-field picture [41], whereby embedding will be performed via the density matrix.

In embedded mean-field theory (EMFT), the total energy can be written as

$$E[\rho] = \text{tr}[\rho H_0] + G[\rho], \quad (3.50)$$

where  $H_0$  contains the one-electron terms of the Hamiltonian and  $G[\rho]$  contains all two-electron terms (Hartree and exchange-correlation effects). In embedded mean-field theory (EMFT), the two-electron interaction for the active subsystem A is constructed at a higher level of theory to the rest of the system,

$$E^{\text{EMFT}}[\rho] = \text{tr}[\rho H_0] + G^{\text{low}}[\rho] + (G^{\text{high}}[\rho_{AA}] - G^{\text{low}}[\rho_{AA}]), \quad (3.51)$$

where  $G^{\text{low}}$  and  $G^{\text{high}}$  are the two-electron interaction energies at the lower and higher levels of theory, respectively. For example, the low level theory could be LDA while the higher level uses a hybrid functional such as B3LYP. We assume here that the core Hamiltonian  $H_0$  is the same at both levels of theory, though this need not necessarily be the case. The ground state of the embedded system can thus be obtained by minimising (3.51) with respect to the elements of the density matrix.



In general, we can use a hybrid functional for the active high-level subsystem, in which case the total energy becomes,

$$E^{\text{EMFT}}[\rho] = \text{tr}[\rho H_0] + E_J^{\text{low}}[\rho] + \left( E_J^{\text{high}}[\rho_{\text{AA}}] - E_J^{\text{low}}[\rho_{\text{AA}}] \right) + E_{\text{XC}}^{\text{low}}[\rho] + \left( E_{\text{XC}}^{\text{high}}[\rho_{\text{AA}}] + \alpha E_{\text{EE}}^{\text{A}}[\rho] - E_{\text{XC}}^{\text{low}}[\rho_{\text{AA}}] \right), \quad (3.52)$$

where we have split the two-electron energy into contributions from Coulomb, exchange-correlation and exact exchange components,  $\alpha E_{\text{EE}}^{\text{A}}$  is the exact exchange energy for subsystem A and  $\alpha$  is the portion of exact exchange to be included. The exact exchange contribution can thus be expressed purely in terms of the active subsystem density matrix,

$$E_{\text{EX},0}^{\text{A}} = -\frac{1}{4} \sum_{\alpha\beta\gamma\delta \in A} \rho_{\alpha\beta}^{\text{AA}} (\alpha\gamma|\beta\delta) \rho_{\gamma\delta}^{\text{AA}}. \quad (3.53)$$

This has a significant computation benefit, as the cost of evaluating this energy compared to a hybrid calculation of the full system is reduced by a factor of  $O((N_A/N)^4)$ , where  $N_A$  and  $N$  are the number of electrons in subsystem A and the full system respectively. Alternatively, one could include couplings between systems A and B,

$$E_{\text{EX},1}^{\text{A}} = -\frac{1}{4} \sum_{\alpha\beta \in A; \gamma\delta} \rho_{\alpha\beta}^{\text{AA}} (\alpha\gamma|\beta\delta) \rho_{\gamma\delta}, \quad (3.54)$$

Although this is more expensive than  $E_{\text{EX},0}^{\text{A}}$ , there is still a saving of  $O((N_A/N)^2)$  relative to an exact exchange calculation of the entire system. Fornace *et al.* have performed tests for the EMFT approach on a variety of chemical systems, such as partitioning covalent bonds, hydrogenation of conjugated systems, and defect formation in graphene.

### 3.5.1 Block orthogonalisation

The use of substantially different exchange-correlation functionals with EMFT can give rise to unphysical self-consistent solutions due to the collapse of the electron density into the cross-overlap terms of the density matrix [44]. This has been observed for localised atomic orbital basis sets by examining the normalisation term  $\text{Tr}[\rho\mathbf{S}]$ , which can be decomposed into subsystem components,

$$\text{Tr}[\rho\mathbf{S}] = N = \text{Tr}[\rho_{\text{AA}}\mathbf{S}_{\text{AA}}] + \text{Tr}[\rho_{\text{BB}}\mathbf{S}_{\text{BB}}] + \text{Tr}[\rho_{\text{AB}}\mathbf{S}_{\text{BA}}] + \text{Tr}[\rho_{\text{BA}}\mathbf{S}_{\text{AB}}]. \quad (3.55)$$

Although normalisation of the full density matrix is maintained, the existence of large, negative terms in the off-diagonal terms  $\text{Tr}[\rho_{\text{BA}}\mathbf{S}_{\text{AB}}]$  and  $\text{Tr}[\rho_{\text{AB}}\mathbf{S}_{\text{BA}}]$  results in the diagonal blocks becoming substantially larger than the actual electron count in those

subsystems. A self-consistent optimisation procedure can result in these terms becoming progressively larger, producing physically meaningless electron density. Ding *et al.* propose a solution to this by enforcing block orthogonality between the subsystem components of the density matrix. Similar to the Manby projector method outlined earlier [32], we begin by performing a Gram-Schmidt orthogonalisation of the environment orbital basis set  $\{\phi_i^B\}$  with respect to the active subsystem basis using a projector,

$$|\tilde{\phi}_i^B\rangle = (1 - \hat{P}^A) |\phi_i^B\rangle, \quad (3.56)$$

$$\hat{P}^A = \sum_{j,k \in A} |\phi_j^A\rangle (\mathbf{S}^{AA})_{jk}^{-1} \langle \phi_k^A|, \quad (3.57)$$

where  $(\mathbf{S}^{AA})^{-1}$  is the inverse overlap matrix constructed from the A subsystem atomic orbitals. This defines a transformation from the initial, non-orthogonal atomic orbital basis set  $\{\phi\}$  to the block-orthogonalised  $\{\tilde{\phi}\}$ ,

$$|\tilde{\phi}_\alpha\rangle = \sum_{i \in A,B} |\phi_i\rangle U_{i\alpha} \quad (3.58)$$

where  $\mathbf{U}$  is the transformation matrix,

$$\mathbf{U} = \begin{pmatrix} \mathbf{I}_{AA} & (\mathbf{S}^{AA})^{-1} \mathbf{S}_{AB} \\ 0 & \mathbf{I}_{BB} \end{pmatrix}. \quad (3.59)$$

Upon transforming the overlap matrix to the block-orthogonalised basis, the resulting matrix  $\tilde{\mathbf{S}}$  is a block-diagonal matrix,

$$\tilde{\mathbf{S}} = \mathbf{U}^T \mathbf{S} \mathbf{U} = \begin{pmatrix} \mathbf{S}_{AA} & 0 \\ 0 & \tilde{\mathbf{S}}_{BB} \end{pmatrix}, \quad (3.60)$$

where

$$\tilde{\mathbf{S}}_{BB} = \mathbf{S}_{BB} (\mathbf{S}_{AA})^{-1} \mathbf{S}_{AA} (\mathbf{S}_{AA})^{-1} \mathbf{S}_{BB}. \quad (3.61)$$

Consequently the normalisation condition is reduced to

$$\text{Tr} [\rho \mathbf{S}] = N = \text{Tr} [\rho_{AA} \mathbf{S}_{AA}] + \text{Tr} [\rho_{BB} \tilde{\mathbf{S}}_{BB}]. \quad (3.62)$$

Now all electrons are associated with the diagonal blocks of the density matrix. As these blocks must be positive semi-definite, the electron density contained within each diagonal sub-block must be less than the total electron density in the system, resulting

in a new, implicit constraint on the density matrix,

$$0 \leq \text{Tr}(\rho_{AA} \mathbf{S}_{AA}) \leq N, \quad 0 \leq \text{Tr}(\rho_{BB} \mathbf{S}_{BB}) \leq N. \quad (3.63)$$

### 3.6 Summary

In this chapter we have discussed the use of pseudopotentials in *ab initio* calculations, whereby the influence of the atomic core on the chemically significant valence electrons can be approximated with a single semi-local potential. This provides a theoretical and historical background to quantum embedding methods, both via the embedding potential that arises in the frozen density embedding (FDE) method, and the projector methods that seek to enforce orthogonality between subsystem Kohn-Sham orbitals. The former seeks to build reasonable approximations to the non-additive kinetic potential in order to accurately describe the subsystem coupling, while the latter attempt to bypass the difficulty associated with such approximations. Such approximate embedding methods enable the use of multiple levels of theory for multi-scale calculations, for example QM/MM techniques or combining different exchange-correlation functionals in different subsystems. We also reviewed the embedded mean-field theory (EMFT) method, whereby the one-electron density matrix is partitioned between subsystems which are treated using separate exchange-correlation functionals while permitting electron transfer between subsystems.

## References

- [1] James C. Phillips and Leonard Kleinman. In: *Physical Review* 116.2 (1959), pp. 287–294. ISSN: 0031899X. DOI: [10.1103/PhysRev.116.287](https://doi.org/10.1103/PhysRev.116.287).
- [2] Volker Heine. In: *Solid State Physics - Advances in Research and Applications* 24.C (1970), pp. 1–36. ISSN: 00811947. DOI: [10.1016/S0081-1947\(08\)60069-7](https://doi.org/10.1016/S0081-1947(08)60069-7).
- [3] M. C. Payne et al. In: *Reviews of Modern Physics* 64.4 (1992), pp. 1045–1097. ISSN: 0034-6861. DOI: [10.1103/RevModPhys.64.1045](https://doi.org/10.1103/RevModPhys.64.1045).
- [4] Andre Severo Pereira Gomes and Christoph R. Jacob. In: *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* 108 (1 2012), pp. 222–277. DOI: [10.1039/C2PC90007F](https://doi.org/10.1039/C2PC90007F).
- [5] Qiming Sun and Garnet Kin Lic Chan. In: *Accounts of Chemical Research* 49.12 (2016), pp. 2705–2712. ISSN: 15204898. DOI: [10.1021/acs.accounts.6b00356](https://doi.org/10.1021/acs.accounts.6b00356).
- [6] L. Szasz. *Pseudopotential Theory of Atoms and Molecules*. U.K.: John Wiley & Sons, Inc., 1985.
- [7] Conyers Herring. In: *Phys. Rev.* 57 (12 1940), pp. 1169–1177. DOI: [10.1103/PhysRev.57.1169](https://doi.org/10.1103/PhysRev.57.1169).
- [8] Leonard Kleinman and D M Bylander. In: *Phys. Rev. Lett.* 48.20 (1982), pp. 1425–1428. DOI: [10.1103/PhysRevLett.48.1425](https://doi.org/10.1103/PhysRevLett.48.1425).
- [9] Tomasz Adam Wesolowski and Arieh Warshel. In: *Journal of Physical Chemistry* 97.30 (1993), pp. 8050–8053. ISSN: 0022-3654. DOI: [10.1021/j100132a040](https://doi.org/10.1021/j100132a040).
- [10] J. D. Goodpaster, N. Ananth, F. R. Manby, and T. F. Miller III. In: *J. Chem Phys.* 133.084103 (2010).
- [11] S. Laricchia, E. Fabiano, and F. Della Sala. In: *Journal of Chemical Physics* 133.16 (2010). ISSN: 00219606. DOI: [10.1063/1.3494537](https://doi.org/10.1063/1.3494537).
- [12] Jacopo Tomasi, Benedetta Mennucci, and Roberto Cammi. In: *Chemical Reviews* 105.8 (2005), pp. 2999–3094. ISSN: 0009-2665. DOI: [10.1021/cr9904009](https://doi.org/10.1021/cr9904009).
- [13] A. Klamt and G. Schüürmann. In: *J. Chem. Soc., Perkin Trans. 2* (5 1993), pp. 799–805. DOI: [10.1039/P29930000799](https://doi.org/10.1039/P29930000799).
- [14] A. Warshel and M. Levitt. In: *Journal of Molecular Biology* 103.2 (1976), pp. 227–249. ISSN: 00222836. DOI: [10.1016/0022-2836\(76\)90311-9](https://doi.org/10.1016/0022-2836(76)90311-9).
- [15] Hans Martin Senn and Walter Thiel. In: *Angewandte Chemie - International Edition* 48.7 (2009), pp. 1198–1229. ISSN: 14337851. DOI: [10.1002/anie.200802019](https://doi.org/10.1002/anie.200802019).
- [16] Tomasz Adam Wesolowski, Henry Chermette, and Jacques Weber. In: *Journal of Chemical Physics* 105.20 (1996), pp. 9182–9190. ISSN: 00219606. DOI: [10.1063/1.472823](https://doi.org/10.1063/1.472823).

- [17] Tomasz Adam Wesolowski. In: *The Journal of Chemical Physics* 106.20 (1997), pp. 8516–8526. ISSN: 0021-9606. DOI: [10.1063/1.473907](https://doi.org/10.1063/1.473907).
- [18] Samuel Fux et al. In: *Chemical Physics Letters* 461.4-6 (2008), pp. 353–359. ISSN: 00092614. DOI: [10.1016/j.cplett.2008.07.038](https://doi.org/10.1016/j.cplett.2008.07.038).
- [19] S. Maya Beyhan, Andreas W. Götz, Christoph R. Jacob, and Lucas Visscher. In: *Journal of Chemical Physics* 132.4 (2010). ISSN: 00219606. DOI: [10.1063/1.3297886](https://doi.org/10.1063/1.3297886).
- [20] James D. Talman and William F. Shadwick. In: *Phys. Rev. A* 14 (1 1976), pp. 36–40. DOI: [10.1103/PhysRevA.14.36](https://doi.org/10.1103/PhysRevA.14.36).
- [21] R. T. Sharp and G. K. Horton. In: *Phys. Rev.* 90 (2 1953), pp. 317–317. DOI: [10.1103/PhysRev.90.317](https://doi.org/10.1103/PhysRev.90.317).
- [22] Samuel Fux et al. In: *The Journal of Chemical Physics* 132.16, 164101 (2010). DOI: <http://dx.doi.org/10.1063/1.3376251>.
- [23] Qin Wu and Weitao Yang. In: *The Journal of Chemical Physics* 118.6 (2003), pp. 2498–2509. DOI: <http://dx.doi.org/10.1063/1.1535422>.
- [24] Christoph R. Jacob. In: *The Journal of Chemical Physics* 135.24, 244102 (2011). DOI: <http://dx.doi.org/10.1063/1.3670414>.
- [25] O. Roncero et al. In: *Journal of Chemical Physics* 129.18 (2008). ISSN: 00219606. DOI: [10.1063/1.3007987](https://doi.org/10.1063/1.3007987).
- [26] Chen Huang and Emily A. Carter. In: *Journal of Chemical Physics* 135.19 (2011). ISSN: 00219606. DOI: [10.1063/1.3659293](https://doi.org/10.1063/1.3659293).
- [27] Alexander Zech, Francesco Aquilante, and Tomasz A. Wesolowski. In: *The Journal of Chemical Physics* 143.16, 164106 (2015). DOI: <http://dx.doi.org/10.1063/1.4933372>.
- [28] N. Govind, Y.A. Wang, A.J.R. da Silva, and E.A. Carter. In: *Chemical Physics Letters* 295.1–2 (1998), pp. 129 –134. ISSN: 0009-2614. DOI: [http://dx.doi.org/10.1016/S0009-2614\(98\)00939-7](http://dx.doi.org/10.1016/S0009-2614(98)00939-7).
- [29] Yuriy G. Khait and Mark R. Hoffmann. In: *The Journal of Chemical Physics* 133.4, 044107 (2010). DOI: <http://dx.doi.org/10.1063/1.3460594>.
- [30] John P. Perdew and Mel Levy. In: *Phys. Rev. B* 31 (10 1985), pp. 6264–6272. DOI: [10.1103/PhysRevB.31.6264](https://doi.org/10.1103/PhysRevB.31.6264).
- [31] Łukasz Rajchel, Piotr S. Zuchowski, Małgorzata M. Szcześniak, and Grzegorz Chałasiński. In: *Chemical Physics Letters* 486.4-6 (2010), pp. 160–165. ISSN: 00092614. DOI: [10.1016/j.cplett.2009.12.083](https://doi.org/10.1016/j.cplett.2009.12.083).

- [32] Frederick R. Manby, Martina Stella, Jason D. Goodpaster, and Thomas F. Miller. In: *Journal of Chemical Theory and Computation* 8.8 (2012), pp. 2564–2568. ISSN: 15499618. DOI: [10.1021/ct300544e](https://doi.org/10.1021/ct300544e).
- [33] Bence Hégely, Péter R. Nagy, György G. Ferenczy, and Mihály Kállay. In: *The Journal of Chemical Physics* 145.6 (2016), p. 064107. ISSN: 0021-9606. DOI: [10.1063/1.4960177](https://doi.org/10.1063/1.4960177).
- [34] F. Libisch, M. Marsman, J. Burgdörfer, and G. Kresse. In: *The Journal of Chemical Physics* 147.3 (2017), p. 034110. ISSN: 0021-9606. DOI: [10.1063/1.4993795](https://doi.org/10.1063/1.4993795).
- [35] John D. Weeks. In: *The Journal of Chemical Physics* 49.6 (1968), p. 2741. ISSN: 00219606. DOI: [10.1063/1.1670479](https://doi.org/10.1063/1.1670479).
- [36] S. Huzinaga and A. A. Cantu. In: *The Journal of Chemical Physics* 55.12 (1971), p. 5543. ISSN: 00219606. DOI: [10.1063/1.1675720](https://doi.org/10.1063/1.1675720).
- [37] John D. Weeks and Stuart A. Rice. In: *The Journal of Chemical Physics* 49.6 (1968), pp. 2741–2755. DOI: <http://dx.doi.org/10.1063/1.1670479>.
- [38] Simon J. Bennie, Martina Stella, Thomas F. Miller, and Frederick R. Manby. In: *The Journal of Chemical Physics* 143.2, 024105 (2015). DOI: <http://dx.doi.org/10.1063/1.4923367>.
- [39] Martina Stella, Simon J. Bennie, and Frederick R. Manby. In: *Molecular Physics* 113.13-14 (2015), pp. 1858–1864. DOI: [10.1080/00268976.2015.1018359](https://doi.org/10.1080/00268976.2015.1018359).
- [40] Frederick R. Manby, Martina Stella, Jason D. Goodpaster, and III Thomas F. Miller. In: *Journal of Chemical Theory and Computation* 8.8 (2012). PMID: 22904692, pp. 2564–2568. DOI: [10.1021/ct300544e](https://doi.org/10.1021/ct300544e).
- [41] Mark E. Fornace et al. In: *Journal of Chemical Theory and Computation* 11.2 (2015), pp. 568–580. ISSN: 15499626. DOI: [10.1021/ct5011032](https://doi.org/10.1021/ct5011032).
- [42] Gerald Knizia and Garnet Kin Lic Chan. In: *Journal of Chemical Theory and Computation* 9.3 (2013), pp. 1428–1432. ISSN: 15499618. DOI: [10.1021/ct301044e](https://doi.org/10.1021/ct301044e).
- [43] D. R. Bowler and M. J. Gillan. In: *Chemical Physics Letters* 355.3-4 (2002), pp. 306–310. ISSN: 00092614. DOI: [10.1016/S0009-2614\(02\)00273-7](https://doi.org/10.1016/S0009-2614(02)00273-7).
- [44] Feizhi Ding, Frederick R. Manby, and Thomas F. Miller. In: *Journal of Chemical Theory and Computation* (2017). ISSN: 15499626. DOI: [10.1021/acs.jctc.6b01065](https://doi.org/10.1021/acs.jctc.6b01065).

## 4 Linear-Scaling DFT and ONETEP

Traditionally, electronic structure calculations were performed by means of direct diagonalisation of the Hamiltonian matrix [1]. This has the advantage of yielding eigenstates that can be associated with discrete energy levels, useful for comparison with experiment, but diagonalisation is a cubic-scaling operation in CPU time [2]. Since the Hamiltonian is a square matrix whose size is proportional to the number of basis functions  $M$ , solving the Kohn-Sham equations by direct diagonalisation scales as  $\mathcal{O}(M^3)$ , which prohibits their application to large-scale calculations. The choice of basis functions was critical, as  $M$  could be much larger than the number of atoms  $N_a$  in the system when using plane waves instead of localised orbitals such as Gaussians. As computing power has increased, scientists have sought ways to avoid the  $\mathcal{O}(M^3)$  bottleneck, such as iterative diagonalisation methods [3] that reduce the cost to  $\mathcal{O}(N_a M^2)$ . The use of the fast Fourier transform method (FFT) in particular, with a computational scaling proportional to  $N_a M \ln(M)$ , has enabled plane-wave DFT calculations to be performed on systems with up to 1000 atoms, far beyond what would be possible with direct diagonalisation [2].

For systems of several hundred atoms, the need to enforce orthogonality between the Kohn-Sham orbitals becomes the bottleneck in CPU time. Cost-wise, increasing the number of electrons in the system requires the use of a larger basis set, giving a factor of  $N_a$ . Each plane wave basis function extends throughout the system, such that any manipulation of a single orbital is an  $\mathcal{O}(N_a)$  operation. Finally, each orbital must be orthogonalised with respect to all the others which adds another factor of  $N_a$  to the scaling of the calculation [4]. Thus, due to the need to impose orthogonality the scaling of a Kohn-Sham DFT calculation is  $\mathcal{O}(N_a^3)$  with a plane wave basis,

$$T_{\text{CPU}} \sim c_3 N_a^3, \quad (4.1)$$

where  $c_3$  is the prefactor for the calculation. While favourable compared to the wave-function-based quantum chemical methods discussed in Chapter 2, this  $\mathcal{O}(N_a^3)$  scaling restricts the applications of KS-DFT in terms of system size. In recent decades there has been widespread interest in linear-scaling DFT methods, whereby the computational cost can be represented as

$$T_{\text{CPU}} \sim c_1 N_a, \quad (4.2)$$

where  $c_1$  is the prefactor associated with the linear-scaling calculation. In general,  $c_1 > c_3$  [1], such that there is a crossover point where linear-scaling methods become more efficient than conventional cubic-scaling approaches, usually on the order of hundreds of atoms.

While there are many different approaches for achieving linear scaling (for reviews, see e.g. [1, 2, 5, 6]), in this chapter we will focus on methods based on the one-particle density matrix [7]. We will begin by discussing how the density-matrix formalism can facilitate the use of linear-scaling methods, with particular focus on the use of local-orbital basis sets and methods that avert the need to explicitly enforce orthonormality during the self-consistent optimisation process. This will lead us into a review of the implementation of such approaches within the ONETEP software package [8], which will form the basis for the remainder of the work presented in this thesis.

## 4.1 Nearsightedness and linear-scaling

An intuitive understanding of chemical bonding highlights the importance of locality in electronic structure — if such phenomena were determined by long-range effects in materials, our concept of chemical bonding and localised orbitals would be ill-defined [9]. This is an example of the nearsightedness of electronic structure [10], whereby the local electronic structure depends primarily on the external potential in the immediate vicinity. Theoretically, in systems with a band gap nearsightedness manifests itself via an exponential decay in the elements of the density matrix with distance,

$$\rho(\mathbf{r}, \mathbf{r}') \sim e^{-\gamma|\mathbf{r}-\mathbf{r}'|}, \quad (4.3)$$

where  $\gamma$  is a positive quantity that is proportional to the band gap [1]. From a computational perspective, nearsightedness highlights the fact that the cubic scaling of DFT is an algorithmic consequence rather than one inherent to the physics of many-body systems. This exponential decay forms the theoretical basis for linear-scaling methods in insulators, as it allows one to impose a localisation constraint on the density matrix. For metallic systems, the decay of the density matrix is slower, following a power law rather than exponential [1]. However, an exponential decay arises in the canonical ensemble at finite, non-zero temperatures [11, 12], facilitating the application of linear-scaling approaches to metallic systems. For the remainder of this work, we will assume that we are dealing with insulators and that the bandgap of the system is well-defined at zero temperature.

The earliest LS-DFT approach was Yang’s divide-and-conquer strategy [13], where



the system is split up into localisation regions, a process made possible by nearsightedness [9]. The electron densities from these subsystems are gathered together to obtain the total density. The divisions are made using the partition function for each subsystem [14]. The prefactor associated with this method can be very high if large localisation regions are used as much of the information in the density matrix is discarded during a calculation [1].

Another approach is to make explicit use of the nearsightedness of the density matrix  $\rho(\mathbf{r}, \mathbf{r}')$  by truncating the off-diagonal elements that exceed a cutoff radius  $R_c$  [15]. This has the advantage that one does not need to refer directly to the individual KS orbitals as all information regarding the physics of our system can be extracted by examining the density matrix. This latter approach is the key principle behind the ONETEP package, and so we will explore this in more detail as it is fundamental to the rest of this work.

#### 4.1.1 Density matrices with non-orthogonal basis sets

In Section 2.5 we reviewed the basic principles of density matrix theory [7, 16] within the context of Kohn-Sham density-functional theory. Now we shall expand the Kohn-Sham orbitals  $\{\psi_n\}$  in a set of basis functions  $\{\phi_\alpha\}$  known as support functions [17], which need not be orthogonal,

$$|\psi_n\rangle = |\phi_i\rangle M_n^i, \quad (4.4)$$

where  $M$  is the matrix of expansion coefficients and we use the Einstein convention of summing over repeated indices ( $i$  in this case). Throughout the remainder of this work, we will apply the Einstein convention to Latin indices when referring to support functions, while Greek characters will denote distinct orbitals or matrix elements and will not be summed unless explicitly stated. To account for the non-orthogonality of the  $\{\phi_\alpha\}$  basis, we have to consider the tensorial properties of the quantities in this equation [18, 19]. Subscripts refer to covariant quantities, while superscripts denote contravariant tensors. Thus the matrix element  $M_n^\alpha$  denotes a mixed contravariant-covariant quantity, though since the  $\{|\psi_n\rangle\}$  form an orthogonal basis, there is technically no tensorial distinction for the  $n$  subscript/superscript.

The basis functions  $\{\phi_\alpha\}$  may be any generalised basis set of real or complex functions, such as the Wannier functions discussed in Section 2.6.4. They form a covariant set of basis functions that are biorthogonal to a dual basis of contravariant functions  $\{\phi^\alpha\}$ , such that

$$\langle \phi_\alpha | \phi^\beta \rangle = \delta_\alpha^\beta. \quad (4.5)$$

The overlap matrices for these bases are defined as

$$S_{\alpha\beta} = \langle \phi_\alpha | \phi_\beta \rangle, \quad (4.6)$$

$$S^{\alpha\beta} = \langle \phi^\alpha | \phi^\beta \rangle = (\mathbf{S}^{-1})_{\alpha\beta}, \quad (4.7)$$

which form the metric tensors that can be used to convert covariant quantities to contravariant and vice versa. For more details on the use of tensors in electronic structure, see [18, 19].

Returning to (2.64), we now expand the density operator in terms of the support functions (4.4),

$$\hat{\rho} = \sum_n f_n |\psi_n\rangle \langle \psi_n|, \quad (4.8)$$

$$= \sum_n f_n |\phi_i\rangle M_n^i (M^\dagger)_n^j \langle \phi_j|, \quad (4.9)$$

$$= |\phi_i\rangle \left[ \sum_n M_n^i f_n (M^\dagger)_n^j \right] \langle \phi_j|, \quad (4.10)$$

$$= |\phi_i\rangle K^{ij} \langle \phi_j|, \quad (4.11)$$

where

$$K^{\alpha\beta} \equiv \sum_n M_n^\alpha f_n (M^\dagger)_n^\beta = \langle \phi^\alpha | \hat{\rho} | \phi^\beta \rangle \quad (4.12)$$

are the elements of the *density kernel* [20], the representation of the density matrix in the dual basis. In the position representation, the density matrix can thus be expressed as

$$\rho(\mathbf{r}, \mathbf{r}') = \phi_i(\mathbf{r}) K^{ij} \phi_j^*(\mathbf{r}'), \quad (4.13)$$

such that the electron density in the support function basis is given by

$$n(\mathbf{r}) = \phi_i(\mathbf{r}) K^{ij} \phi_j^*(\mathbf{r}) = K^{ij} [\phi_j^*(\mathbf{r}) \phi_i(\mathbf{r})]. \quad (4.14)$$

The properties of the density matrix established in Section 2.5 obviously still hold in the support function representation. For example, idempotency ( $\rho^2 = \rho$ ):

$$\rho^2(\mathbf{r}, \mathbf{r}') = \int \rho(\mathbf{r}, \tilde{\mathbf{r}}) \rho(\tilde{\mathbf{r}}, \mathbf{r}') d\tilde{\mathbf{r}}, \quad (4.15)$$

$$\begin{aligned} &= \int [\phi_i(\mathbf{r}) K^{ij} \phi_j^*(\tilde{\mathbf{r}})] [\phi_k(\tilde{\mathbf{r}}) K^{kl} \phi_l^*(\mathbf{r}')] d\tilde{\mathbf{r}}, \\ &= \phi_i(\mathbf{r}) K^{ij} \left[ \int \phi_j^*(\tilde{\mathbf{r}}) \phi_k(\tilde{\mathbf{r}}) d\tilde{\mathbf{r}} \right] K^{kl} \phi_l^*(\mathbf{r}), \\ &= \phi_i(\mathbf{r}) K^{ij} S_{jk} K^{kl} \phi_l^*(\mathbf{r}), \end{aligned} \quad (4.16)$$

Comparing with (4.13), we see that idempotency is satisfied if

$$\mathbf{K} = \mathbf{KSK}. \quad (4.17)$$

Similarly, the normalisation condition and expectation value for an operator corresponding to an observable of the system become

$$\text{Tr}(\rho) = \text{Tr}[\mathbf{KS}] = N, \quad (4.18)$$

$$\langle \hat{A} \rangle = \text{Tr}[\mathbf{KA}]. \quad (4.19)$$

### 4.1.2 Constraining the density matrix

The density matrix formalism of DFT provides a framework within which it is possible to reconstruct all information for the fictitious Kohn-Sham system without direct reference to the orbitals, enabling one to bypass the need for explicitly orthogonalisation. Linear-scaling approaches utilising density matrices must therefore seek to enforce (2.67) while avoiding expensive diagonalisation to write the DM explicitly in terms of the KS orbitals [15]. The ground state energy of the Kohn-Sham system can be obtained from explicit minimisation of the (non-zero) elements of the density matrix,

$$E_0 = \min_{\rho_{ij}} \text{Tr}(\rho \mathbf{H}), \quad (4.20)$$

where  $\mathbf{H}$  is the Hamiltonian matrix and we must ensure that idempotency (equation (2.67)) and normalisation (equation (2.70)) remain satisfied throughout this procedure. At the true ground state, the density matrix and the Hamiltonian will commute,

$$[\rho, \mathbf{H}] = \rho \mathbf{H} - \mathbf{H} \rho = 0. \quad (4.21)$$

### Purification transformations and penalty functionals

McWeeny [20] showed that an iterative steepest descent minimisation of a nearly idempotent matrix can be used to purify the matrix to idempotency,

$$\rho_{k+1} = 3\rho_k^2 - 2\rho_k^3, \quad (4.22)$$

which yields the following relation for the eigenvalues of the density matrices,

$$\lambda_{k+1} = 3\lambda_k^2 - 2\lambda_k^3. \quad (4.23)$$

Then if  $-\frac{1}{2} \leq \lambda_k \leq \frac{3}{2}$ , it follows that  $0 \leq \lambda_{k+1} \leq 1$ . Successive iterations will improve the idempotency of the matrix, with eigenvalues tending to 0 for unoccupied states and 1 for occupied states [15]. However, this merely yields weak idempotency [5], which ensures that the eigenvalues are bound in the range  $[0, 1]$  but does not guarantee that occupied states have an eigenvalue of 1, which can lead to numerical instabilities. An appropriate starting point  $\rho_0$  that will ensure that (4.21) is satisfied can be obtained by canonical purification [21].

The purification transformation can be derived as the result of a steepest descent optimisation of [22]

$$P[\rho] = \text{Tr} \left[ (\rho^2 - \rho)^2 \right] = \sum_n (\lambda_n^2 - \lambda_n)^2. \quad (4.24)$$

The term  $\rho^2 - \rho$  vanishes if and only if  $\rho$  is idempotent. Thus the functional  $P[\rho]$  can be used as an additional penalty functional term in a generalised total energy optimisation routine [10, 23],

$$Q[\rho] = E[\rho] + \alpha P[\rho], \quad (4.25)$$

where  $\alpha$  is a positive constant defined such that the ground state energy will be obtained when  $\rho$  is idempotent. Penalty functional approaches utilised in the literature are variations of the term  $\alpha P[\rho]$  designed in an effort to drive the density matrix to idempotency during the optimisation.

### LNV method

In the spirit of the purification transformation (4.22), Li, Nunes and Vanderbilt (LNV) defined the density matrix  $\rho$  in terms of an auxiliary matrix  $\sigma$ , [24, 25]

$$\rho = 3\sigma^2 - 2\sigma^3, \quad (4.26)$$

where  $\sigma$  can itself be written in terms of the support functions,

$$\sigma(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r}) L^{\alpha\beta} \phi_\beta^*(\mathbf{r}'), \quad (4.27)$$

where  $\mathbf{L}$  is the auxiliary kernel. The density kernel  $\mathbf{K}$  can thus be written as

$$\mathbf{K} = 3\mathbf{L}\mathbf{S}\mathbf{L} - 2\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{S}\mathbf{L}. \quad (4.28)$$

Then the Kohn-Sham total energy functional  $E[\rho]$  can be expressed instead as a functional of  $\mathbf{L}$ ,

$$E[\mathbf{L}] = \text{Tr} [\mathbf{H} (3\mathbf{L}\mathbf{S}\mathbf{L} - 2\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{S}\mathbf{L})], \quad (4.29)$$

which can be minimised variationally with respect to the elements of the auxiliary kernel  $\{L^{\alpha\beta}\}$  to obtain the ground state energy  $E_0$ . The cubic dependence of the energy functional on the auxiliary kernel avoids the problem of multiple minima during optimisation [26], and provided the occupancies remain within the range  $(-\frac{1}{2}, \frac{3}{2})$ , minimisation of the functional will yield a variational approximation to the ground state. However, should any of the occupancies fall outside the required range at any stage of the optimisation, the algorithm may become unstable. Variations of the LNV method, such as the Millam-Scuseria approach, have been developed to provide more stable convergence [27].

### Normalisation

There is also the question of normalisation of the density kernel. Thankfully this is somewhat more straightforward than the task of enforcing idempotency. We can define a grand potential of the form,

$$\Omega = E - \mu N, \quad (4.30)$$

where  $\mu$ , the chemical potential, acts as a Lagrange multiplier to ensure that normalisation is maintained. Now we seek to find minima of the functional  $\Omega$  instead of the total energy directly,

$$\Omega[\rho] = \text{Tr}[(\mathbf{H} - \mu\mathbf{I})\rho]. \quad (4.31)$$

Generally it can be more convenient to work at fixed chemical potential [17]. In terms of the density kernel, the functional to be minimised can now be expressed as

$$\Omega = \text{Tr}[\mathbf{H}\mathbf{K} - \mu\mathbf{S}\mathbf{K}]. \quad (4.32)$$

### Optimising the support functions

So far we have focused entirely on the density kernel and neglected the support functions. One can consider the support functions to be fixed during the calculation, such as in codes that utilise pseudoatomic orbitals or Gaussian basis sets [6], and only the density kernel is optimised. Improvements in accuracy can be obtained by increasing the size of the basis, although of course a larger basis implies a more costly calculation. In addition, fixed basis set methods are susceptible to basis set superposition error (BSSE) for binding energies, which requires counterpoise corrections to eliminate [28].

Other methods deal with the orbitals explicitly but remove the need for orthogonality to be directly imposed [29]. For instance, in the approaches of Mauri *et al.* and

Ordejon *et al.*, a Löwdin transformation is performed on the single particle wavefunctions [30],

$$|\psi_i\rangle = \left(\mathbf{S}^{-\frac{1}{2}}\right)^{ij} |\Phi_j\rangle, \quad (4.33)$$

The total energy functional is constructed such that orthonormality is implicitly included. Minimisation with respect to the orbitals  $\{\Phi_i\}$  proceeds without any orthogonalisation and these orbitals are orthonormal only at the global minimum of the energy functional [31], thus avoiding this costly procedure [5]. Meanwhile, the orbitals can be constrained to maintain strict locality such that the nearsightedness of the density matrix can still be exploited. Optimised local orbitals also do not suffer from BSSE, so no correction to interaction energies is required [22].

## 4.2 The ONETEP approach

Having established the fundamental theory required, we now move on to a description of the ONETEP linear-scaling DFT code [8, 32], which we will use throughout the remainder of this work. The density matrix is expressed in the form of (4.11), using an atom-centred local-orbital basis  $\{\phi_\alpha\}$ , referred to as non-orthogonal generalised Wannier functions (NGWFs) [33]. Given that we are mainly interested in calculations on large-scale systems, we can limit ourselves to performing calculations at the  $\Gamma$ -point in k-space (see Section 2.6.1) and the NGWFs are considered to be real. To facilitate efficient calculations on large systems, a minimal number of NGWFs per atoms are used during a calculation in ONETEP, such that the number of basis functions  $M$  is comparable to the number of electrons. These functions are generated as an expansion of periodic bandwidth limited delta functions, or psincs  $\{D_k(\mathbf{r})\}$ ,

$$\phi_\alpha(\mathbf{r}) = \sum_{k=0}^{N_1-1} \sum_{l=0}^{N_2-1} \sum_{m=0}^{N_3-1} C_{klm,\alpha} D_{klm}(\mathbf{r}), \quad (4.34)$$

$$D_{klm}(\mathbf{r}) = \frac{1}{N_1 N_2 N_3} \sum_{p=-J_1}^{J_1} \sum_{q=-J_2}^{J_2} \sum_{s=-J_3}^{J_3} e^{i(p\mathbf{b}_1 + q\mathbf{b}_2 + s\mathbf{b}_3) \cdot (\mathbf{r} - \mathbf{r}_{klm})}, \quad (4.35)$$

where  $N_i = 2J_i + 1$  is the number of grid points in the direction of lattice vector  $\mathbf{a}_i$ ,  $\mathbf{b}_i$  is the corresponding reciprocal lattice vector,  $\{\mathbf{r}_{klm}\}$  are the grid points of the periodic cell and the  $\{C_{klm,\alpha}\}$  are expansion coefficients. An example of a psinc function compared to a regular sinc (cardinal sine) function is shown in Figure 4.1, which highlights why they can be considered “periodic sinc functions” [32].

The power of this approach can be seen from the fact that the psincs are related to plane waves by a unitary Fourier transform, which makes the ONETEP approach

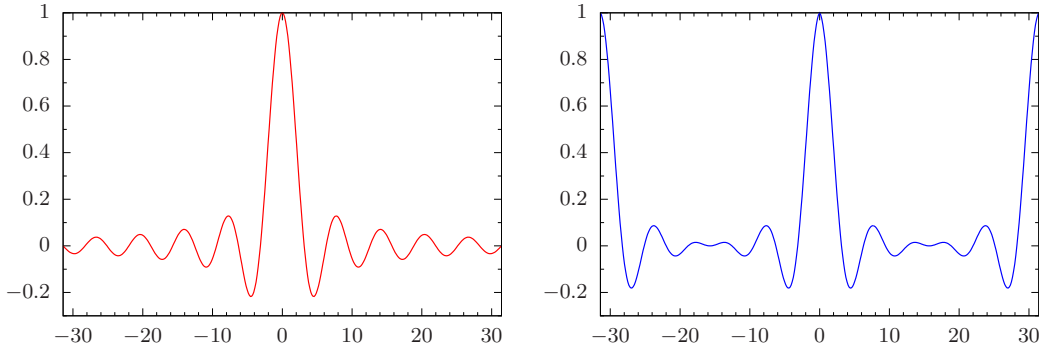


FIGURE 4.1: Sinc function (left) and the periodic bandwidth limited delta function (psinc, right).

effectively a linear-scaling reformulation of the plane wave method in terms of NGWFs [33]. Just as plane wave basis set quality is controlled by the kinetic energy cutoff, the quality of the psinc basis can be systematically improved using a single related parameter, the grid spacing [8], which allows ONETEP calculations to be converged to plane-wave accuracy [32].

#### 4.2.1 Linear-scaling with ONETEP

In order to exploit nearsightedness for linear-scaling, each NGWF is associated with a localisation region, outside of which they are forced to be zero [33]. For large-scale calculations this results in the overlap matrix  $\mathbf{S}$  and the Hamiltonian  $\mathbf{H}$  being sparse [4]. Additionally, the density kernel elements  $K^{\alpha\beta}$  are truncated if the separation between the centres of the corresponding NGWFs is greater than a specified cutoff distance  $|\mathbf{R}_\alpha - \mathbf{R}_\beta| > \mathbf{R}_c$  [8]. This enables us to use sparse matrix algebra to perform the required matrix operations during a calculation, which for sufficiently sparse matrices can be performed with  $\mathcal{O}(N)$  scaling.

There is one final ingredient required to ensure that ONETEP calculations are linear-scaling. As mentioned at the beginning of this chapter, fast Fourier transforms (FFTs) are widely used in *ab initio* methods for efficient computations. However, the  $N^2 \ln(N)$  scaling with system size is clearly inappropriate for linear-scaling. To bypass this bottleneck, ONETEP uses the FFT box technique [34], whereby a box is constructed around each NGWF that contains both it and its overlapping neighbours. The FFTs are then performed for each box individually, rather than over the whole system. Since the dimensions of the FFT box depends only on the NGWF radius, the size of the box is independent of system size (provided the simulation cell is sufficiently large) and the FFT operations can be performed in linear-scaling time.

### 4.2.2 Total Energy Optimisation in ONETEP

Total energy minimisation in ONETEP takes place within a nested two-stage loop as illustrated in Figure 4.2. In addition to optimising the elements of the auxiliary density kernel by means of the LNV method discussed in Section 4.1.2, we also optimise our NGWF basis by minimising the total energy with respect to the expansion coefficients of the underlying psinc basis. The total energy functional can then be written as a function of the elements of the auxiliary kernel and the psinc coefficients  $\Omega(\{L^{\alpha\beta}, \{C_{klm,\alpha}\}\})$ , and the optimisation takes place in a two-stage process [26],

$$E_0 = \min_{\{C_{klm,\alpha}\}} L(\{C_{klm,\alpha}\}), \quad (4.36)$$

$$L(\{C_{klm,\alpha}\}) = \min_{\{K^{\alpha\beta}\}} \Omega(\{K^{\alpha\beta}; \{C_{klm,\alpha}\}\}). \quad (4.37)$$

The initial NGWFs are generated from truncated pseudoatomic orbitals. Then the density kernel is initialised by canonical purification [21], refined using a penalty functional to ensure idempotency [23, 32] and optimised for this set of NGWFs via the LNV algorithm [24, 25]. This setup ensures that approximate idempotency of the kernel is maintained. Next, the NGWFs are updated by applying the NGWF gradient of the total energy functional (see below) and this process repeated until convergence is achieved [26].

### 4.2.3 NGWF and density kernel gradients

The total energy functional can be written as

$$\Omega = \text{Tr}[\mathbf{K}(\mathbf{H} - \mu\mathbf{S})] - E_{\text{DC}}[n] = K^{ij}(H_{ji} - \mu S_{ji}) - E_{\text{DC}}[n], \quad (4.38)$$

where we have included the  $-\mu\mathbf{S}$  term to ensure normalisation during the optimisation and  $E_{\text{DC}}[n]$  is the double-counting correction for the Hartree and exchange-correlation contributions to the DFT band-structure energy [35]. For simplicity, we will transform the Hamiltonian such that we do not need to explicitly refer to this term in the equations below,

$$\mathbf{H} \rightarrow \mathbf{H} - \mu\mathbf{S}. \quad (4.39)$$



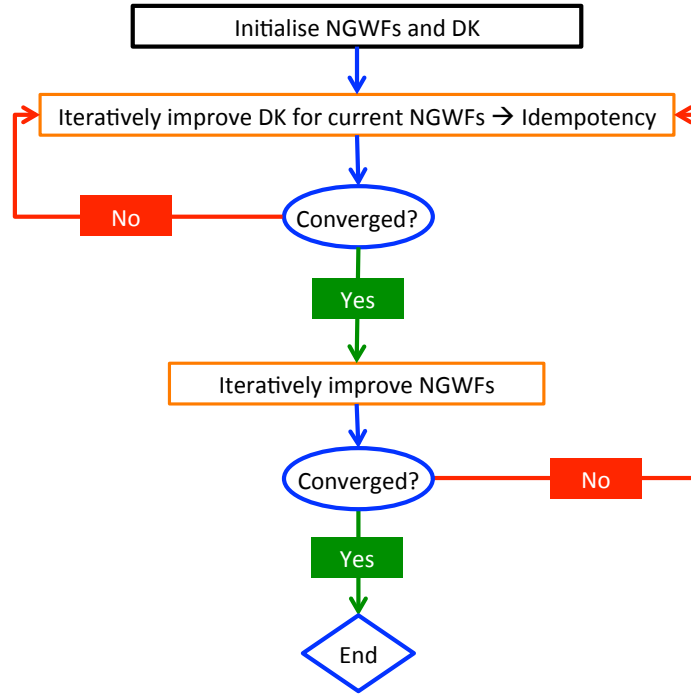


FIGURE 4.2: Self-consistent routine for optimisation of total energy in ONETEP.

In addition, the optimisation procedure including the double-counting term yields the same gradient as the band structure term  $\text{Tr}[\mathbf{KH}]$ , so for simplicity we omit the double-counting term below. Now expanding the density kernel as in equation (4.28) we have

$$\Omega = \text{Tr}[(3\mathbf{LSL} - 2\mathbf{LSLSL})\mathbf{H}] \quad (4.40)$$

$$= 3L^{ij}S_{jk}L^{km}H_{mi} - 2L^{ij}S_{jk}L^{km}S_{mn}L^{np}H_{pi}, \quad (4.41)$$

which can be rewritten explicitly in terms of the NGWFs as

$$\Omega = 3L^{ij}\langle\phi_j|\phi_k\rangle L^{km}\langle\phi_m|\hat{H}|\phi_i\rangle - 2L^{ij}\langle\phi_j|\phi_k\rangle L^{km}\langle\phi_m|\phi_n\rangle L^{np}\langle\phi_p|\hat{H}|\phi_i\rangle. \quad (4.42)$$

For clarity in this derivation we will treat the NGWFs directly as the variables to be optimised, rather than their expansion coefficients, though this is a straightforward addition.

Now differentiating with respect to the NGWFs yields the contravariant gradient,

$$\begin{aligned}
 |g^\alpha\rangle &= \frac{\partial\Omega}{\partial\langle\phi_\alpha|} \\
 &= |\phi_i\rangle (3\mathbf{LHL} - 2\mathbf{LSLHL} - 2\mathbf{LHLSL})^{i\alpha} \\
 &\quad + \hat{H} |\phi_j\rangle (3\mathbf{LSL} - 2\mathbf{LSLSL})^{j\alpha}, \\
 &= |\phi_i\rangle \mathbf{Q}^{i\alpha} + \hat{H} |\phi_j\rangle \mathbf{K}^{j\alpha},
 \end{aligned} \tag{4.43}$$

where

$$\mathbf{Q} = 3\mathbf{LHL} - 2\mathbf{LSLHL} - 2\mathbf{LHLSL}. \tag{4.44}$$

In order to obtain the correct gradient for the NGWF optimisation, we require the covariant term. This can be obtained by multiplying  $g^\alpha$  by the metric, the overlap matrix  $\mathbf{S}$  [18, 19]. The tensorially correct result is thus

$$|g_\alpha\rangle = |g^j\rangle S_{j\alpha}, \tag{4.45}$$

$$= |\phi_i\rangle (\mathbf{QS})^i_{\alpha} + \hat{H} |\phi_j\rangle (\mathbf{KS})^j_{\alpha}. \tag{4.46}$$

The NGWFs can then be updated by using the steepest descent method,

$$|\phi_\alpha^{(k+1)}\rangle = |\phi_\alpha^{(k)}\rangle + \lambda |g_\alpha\rangle, \tag{4.47}$$

where  $\lambda$  is the step length. In practice ONETEP uses conjugate gradients in successive steps to enable more rapid convergence of the NGWFs.

We can obtain the density kernel gradient in a similar manner, by differentiating (4.41) with respect to the elements of the auxiliary density kernel,

$$\begin{aligned}
 G_{\beta\alpha} &= \frac{\partial\Omega}{\partial L^{\alpha\beta}}, \\
 &= [3(\mathbf{SLH} + \mathbf{HLS}) - 2(\mathbf{SLSLH} + \mathbf{SLHLS} + \mathbf{HLSLS})]_{\beta\alpha},
 \end{aligned} \tag{4.48}$$

which as before must be converted to tensorially correct form, using the inverse overlap  $\mathbf{S}^{-1}$ ,

$$\begin{aligned}
 G^{\beta\alpha} &= S^{\beta j} \frac{\partial\Omega}{\partial L^{ij}} S^{i\alpha}, \\
 &= [3(\mathbf{LHS}^{-1} + \mathbf{S}^{-1}\mathbf{HL}) - 2(\mathbf{LSLHS}^{-1} + \mathbf{LHL} + \mathbf{S}^{-1}\mathbf{HLSL})]^{\beta\alpha}.
 \end{aligned} \tag{4.49}$$

### 4.3 Relevant ONETEP functionalities

There are many additional functionalities incorporated within ONETEP beyond simply calculating the ground state energy. In this section we will briefly review the functionalities that will be of use to us in later chapters.

#### 4.3.1 Forces and geometry optimisation

Having developed a format for obtaining the ground state of a quantum system, it is often of interest to calculate the forces of and perform geometry optimisation on molecular structures. The Hellmann-Feynman theorem, as discussed in Chapter 2, can be used to calculate ionic forces within DFT, however this form is only strictly true in the limit of a complete basis [36]. Within the ONETEP formalism, additional contributions will arise from the implicit dependence of the density kernel and NGWFs on ionic positions [37],

$$\mathbf{F}_\gamma = -\frac{dE}{d\mathbf{R}_\gamma} = -\frac{\partial E}{\partial \mathbf{R}_\gamma} - \frac{\partial E}{\partial K^{ij}} \frac{dK^{ji}}{d\mathbf{R}_\gamma} - \int \frac{\delta E}{\delta \phi_i(\mathbf{r})} \frac{d\phi_i(\mathbf{r})}{d\mathbf{R}_\gamma} d\mathbf{r}. \quad (4.50)$$

In principle the derivatives of  $E$  with respect to  $\{K^{\alpha\beta}\}$  and  $\{\phi_\alpha\}$  will be zero at the ground state, such that these terms can be neglected. In practice, while this may be true for the kernel, the NGWF gradient cannot be converged to arbitrarily small accuracy due to the competing effects of the kinetic energy operator, which tends to delocalise the orbitals, and the strict localisation constraint imposed within ONETEP [36, 38]. As a result, a small but significant contribution will be retained from the NGWFs in the form of Pulay forces from the use of a localised atom-centred basis [39]. The total forces calculated in ONETEP are thus the Hellmann-Feynman contribution plus an additional correction for the Pulay forces,

$$\mathbf{F}_\gamma^{\text{total}} = \mathbf{F}_\gamma^{\text{HF}} + \mathbf{F}_\gamma^{\text{Pulay}}. \quad (4.51)$$

#### 4.3.2 Conduction NGWF optimisation

The ONETEP minimal basis approach to finding the ground state energy provides a description of the ground state properties of the system with accuracy comparable to plane-wave methods [8, 32]. The same cannot be said, however, for excited state properties. The NGWF basis  $\{\phi_\alpha\}$  provides overall a poor description of the conduction manifold, without which it is impossible to obtain a valid description of excited state properties such as optical spectra [40]. This is not terribly surprising, as while in principle conduction state properties should be accessible with a complete basis, there is

no reason to believe that this should be true for the minimal NGWF basis used in a ONETEP calculation, as the sole interest in the optimisation is providing an accurate description of the occupied Kohn-Sham states.

A solution to this problem has been implemented within ONETEP in order to accurately calculate conduction state orbitals [40, 41]. First, a ground state calculation is performed, yielding an optimised set of valence NGWFs  $\{\phi_\alpha\}$  and density kernel  $\mathbf{K}_v$ . A new density matrix can be constructed for a set of  $N_c$  virtual orbitals above the Fermi level,

$$\sigma(\mathbf{r}, \mathbf{r}') = \sum_{i=N_{\text{occ}}/2+1}^{(N_{\text{occ}}+N_c)/2} \psi_i(\mathbf{r})\psi_i^*(\mathbf{r}'), \quad (4.52)$$

where we have assumed that the states are doubly occupied due to spin-degeneracy (though this need not be the case in general) and the sum runs over the  $N_c$  lowest unoccupied Kohn-Sham orbitals. We can then expand this quantity in terms of an additional conduction NGWF basis  $\{\chi_\alpha\}$  and density kernel  $\mathbf{K}_c$ ,

$$\sigma(\mathbf{r}, \mathbf{r}') = \chi_i(\mathbf{r})K_c^{ij}\chi_j(\mathbf{r}'). \quad (4.53)$$

The virtual Kohn-Sham orbitals must be orthogonal with respect to the occupied valence orbitals. To achieve this, we construct a projected Hamiltonian of the form,

$$\hat{H}^{\text{proj}} = \hat{H} - \hat{\sigma} \left( \hat{H} - \mu \right) \hat{\sigma}, \quad (4.54)$$

where  $\mu$  is a parameter used to shift the eigenvalues of the valence Hamiltonian above the conduction states we wish to optimise. Thus, following the same procedure as in Section 4.2.2, a minimisation of the functional

$$\Omega_{\text{cond}} = \text{Tr} [\mathbf{H}_c^{\text{proj}} \mathbf{K}_c] \quad (4.55)$$

with respect to  $\mathbf{K}_c$  and the orbitals  $\{\chi_\alpha\}$  will yield an optimised kernel and NGWF basis spanning the  $N_c$  lowest virtual Kohn-Sham states. Subsequently, a one-shot diagonalisation of the Hamiltonian can be performed to obtain optical properties using a joint valence-conduction NGWF basis [40].

The above procedure has been demonstrated to give accurate results for properties that depend on the low-lying states of the conduction manifold [41], however the uppermost, delocalised states are still poorly described, particularly for molecular systems in vacuum. This is to be expected given that our local orbital basis is ill-suited to characterise the high-energy states that can extend into the vacuum. Even for the low-lying states, it is often necessary to use a larger NGWF cutoff radius and even a larger number of NGWFs to capture the conduction state properties. Great care is needed in

the choice of the number of conduction orbitals  $N_c$  and the shift parameter  $\sigma$  in order to avoid becoming trapped in local minima during the optimisation procedure [40], though much of this is now automated by ONETEP.

### 4.3.3 Hartree-Fock exchange

In addition to the standard set of semi-local exchange-correlation functionals, hybrid functionals have been implemented within ONETEP [42]. In the Kohn-Sham formalism, the exact exchange energy is given by

$$E_{\text{HFX}} = - \sum_{i,j} f_i f_j \int \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}) \psi_i(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \quad (4.56)$$

which, expanding the Kohn-Sham orbitals in our NGWF basis as in (4.4), becomes,

$$E_{\text{HFX}} = -K^{\beta\alpha} \int \int \phi_\alpha(\mathbf{r}) \phi_\delta(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_\beta(\mathbf{r}') \phi_\gamma(\mathbf{r}') d\mathbf{r} d\mathbf{r}' K^{\delta\gamma} \quad (4.57)$$

$$= -K^{\beta\alpha} (\phi_\alpha \phi_\delta | \phi_\beta \phi_\gamma) K^{\delta\gamma} = -K^{\beta\alpha} X_{\alpha\beta}, \quad (4.58)$$

where  $(\phi_\alpha \phi_\delta | \phi_\beta \phi_\gamma)$  is a two-electron integral and the last equation defines the exchange matrix  $X_{\alpha\beta}$ . The computational challenge within ONETEP rests in the calculation of the exchange matrix, as the number of two-electron integrals to be computed can scale as  $\sim N_{\text{at}}^2 - N_{\text{at}}^4$  with the number of atoms in the system. Even so, a linear-scaling calculation of the exchange matrix has been implemented in ONETEP [42], though the cost of accurately calculating this term is extremely expensive and can be a severe bottleneck.

### 4.3.4 Excitation energies from ONETEP

ONETEP is fundamentally a ground-state DFT code, and thus all the limitations of standard DFT discussed in Chapter 2 apply here. In particular, even with a highly accurate description of the ground state it is not possible to say anything about the nature of excited states, such as excitation energies, based purely on a ground state calculation. Thus extensions to the ONETEP approach are required for us to be able to obtain excitation energies for the kinds of systems that we wish to study with linear-scaling DFT.

Excited state properties can be obtained using time-dependent density functional theory (TDDFT) [43], an extension of ground-state DFT which comes in two varieties — time-domain TDDFT, which considers the time evolution of the ground state electric density, and the frequency-domain linear-response approach (LR-TDDFT), where

the excitation energies are solutions to an effective eigenvalue equation [44]. The latter has been implemented within ONETEP in both the Tamm-Dancoff Approximation (TDA) [45] and the full linear-response approach [46]. The semi-local adiabatic approximation is used, whereby history dependence is neglected. As things stand, LR-TDDFT in ONETEP is thus restricted to the use of semi-local exchange-correlation functionals, restricting applications of the method due to the severe underestimation of charge-transfer interactions by semi-local functionals.

So far in this chapter we have assumed spin degeneracy throughout and neglected the effects of spin polarisation. Nonetheless, spin-polarised calculations are possible by considering the total density matrix to be the sum of spin up and spin down components,

$$\rho^\uparrow(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r})(K^\uparrow)^{\alpha\beta} \phi_\beta(\mathbf{r}'), \quad (4.59)$$

$$\rho^\downarrow(\mathbf{r}, \mathbf{r}') = \phi_\alpha(\mathbf{r})(K^\downarrow)^{\alpha\beta} \phi_\beta(\mathbf{r}'). \quad (4.60)$$

where now the spin-polarised density kernels are normalised such that

$$\text{Tr} [\mathbf{K}^\uparrow \mathbf{S}] = N_\uparrow, \quad (4.61)$$

$$\text{Tr} [\mathbf{K}^\downarrow \mathbf{S}] = N_\downarrow, \quad (4.62)$$

where  $N_\uparrow$  and  $N_\downarrow$  are the number of electrons in spin up and spin down configurations respectively, such that  $N_\uparrow + N_\downarrow = N$ . Then the total energy functional can be written as

$$\Omega = \text{Tr} [\mathbf{H}^\uparrow \mathbf{K}^\uparrow + \mathbf{H}^\downarrow \mathbf{K}^\downarrow]. \quad (4.63)$$

Note that there will be additional numerical factors of two in the spin-unpolarised formalism, though we have neglected throughout this chapter for clarity. Taking advantage of this functionality, we can in fact extract the ground state electron density for both the spin-polarised (triplet) and spin-unpolarised (singlet) configurations of any system, and thus the ground state energies  $E_T$  and  $E_S$ . Assuming the singlet state is lower in energy, the lowest triplet state excitation energy level can then be calculated as

$$\Delta E_{S_0 \rightarrow T_1} = E_T - E_S. \quad (4.64)$$

This is the  $\Delta$ SCF method and can be used to extract triplet excitation energies, though its applicability is limited to the lowest excitation energy as this is the only excited state which can be rigorously described by a ground-state DFT calculation. On the other hand, unlike LR-TDDFT there is no restriction on the use of hybrid functionals within ONETEP, which enables useful comparisons of the two methods.

## 4.4 Summary

In this chapter we have discussed the motivation and theoretical underpinning for linear-scaling density functional theory in the density matrix formalism. The ONETEP linear-scaling package can perform calculations on systems up to tens of thousands of atoms by virtue of the combinations of the localised NGWF basis, sparse matrices and the FFT box technique [4]. In addition, the expansion of the NGWFs in a basis of psinc functions [33] enables calculations with quality systematically tunable to plane-wave accuracy. While ONETEP is able to perform calculations using implicit solvation models and QM/MM approaches, there is no facilitation for DFT calculations to be performed at different levels of theory for different parts of the system, in spite of the expansive nature of the method. In later chapters we will investigate the possibility of implementing quantum embedding methods within ONETEP.

## References

- [1] Stefan Goedecker. In: *Rev. Mod. Phys.* 71 (4 1999), pp. 1085–1123. DOI: [10.1103/RevModPhys.71.1085](https://doi.org/10.1103/RevModPhys.71.1085).
- [2] D R Bowler and T Miyazaki. In: *Reports on Progress in Physics* 75.3 (2012), p. 036503. ISSN: 0034-4885. DOI: [10.1088/0034-4885/75/3/036503](https://doi.org/10.1088/0034-4885/75/3/036503).
- [3] Yousef Saad. In: *Philadelphia, Pennsylvania: SIAM* (2003), pp. 421–426. ISSN: 1570579X. DOI: [10.2277/0898715342](https://doi.org/10.2277/0898715342).
- [4] N.D.M. Hine et al. In: *Computer Physics Communications* 180.7 (2009), pp. 1041 – 1053. ISSN: 0010-4655. DOI: <http://dx.doi.org/10.1016/j.cpc.2008.12.023>.
- [5] Giulia Galli. In: *Current Opinion in Solid State and Materials Science* 1.6 (1996), pp. 864 –874. ISSN: 1359-0286. DOI: [http://dx.doi.org/10.1016/S1359-0286\(96\)80114-8](http://dx.doi.org/10.1016/S1359-0286(96)80114-8).
- [6] D R Bowler and T Miyazaki. In: *Reports on Progress in Physics* 75.3 (2012), p. 036503. DOI: [10.1088/0034-4885/75/3/036503](https://doi.org/10.1088/0034-4885/75/3/036503).
- [7] R McWeeny. In: *Rev. Mod. Phys.* 32.2 (1960), pp. 335–369. ISSN: 0034-6861. DOI: [10.1103/RevModPhys.32.335](https://doi.org/10.1103/RevModPhys.32.335).
- [8] Chris-Kriton Skylaris, Peter D. Haynes, Arash A. Mostofi, and Mike C. Payne. In: *The Journal of Chemical Physics* 122.8, 084119 (2005). DOI: <http://dx.doi.org/10.1063/1.1839852>.
- [9] E. Prodan. In: *Physical Review B - Condensed Matter and Materials Physics* 73.8 (2006), pp. 0–3. ISSN: 10980121. DOI: [10.1103/PhysRevB.73.085108](https://doi.org/10.1103/PhysRevB.73.085108).
- [10] W. Kohn. In: *Phys. Rev. Lett.* 76 (17 1996), pp. 3168–3171. DOI: [10.1103/PhysRevLett.76.3168](https://doi.org/10.1103/PhysRevLett.76.3168).
- [11] S. Goedecker. In: *Physical Review B - Condensed Matter and Materials Physics* 58.7 (1998), pp. 3501–3502. ISSN: 1550235X. DOI: [10.1103/PhysRevB.58.3501](https://doi.org/10.1103/PhysRevB.58.3501).
- [12] Sohrab Ismail-Beigi and T. A. Arias. In: *Physical Review Letters* 82.10 (1999), pp. 2127–2130. ISSN: 10797114. DOI: [10.1103/PhysRevLett.82.2127](https://doi.org/10.1103/PhysRevLett.82.2127).
- [13] Weitao Yang. In: *Phys. Rev. Lett.* 66 (11 1991), pp. 1438–1441. DOI: [10.1103/PhysRevLett.66.1438](https://doi.org/10.1103/PhysRevLett.66.1438).
- [14] Tai-Sung Lee, James P. Lewis, and Weitao Yang. In: *Computational Materials Science* 12.3 (1998), pp. 259 –277. ISSN: 0927-0256. DOI: [http://dx.doi.org/10.1016/S0927-0256\(98\)00029-9](http://dx.doi.org/10.1016/S0927-0256(98)00029-9).



- [15] D.R. Bowler and M.J. Gillan. In: *Computer Physics Communications* 120.2–3 (1999), pp. 95–108. ISSN: 0010-4655. DOI: [http://dx.doi.org/10.1016/S0010-4655\(99\)00221-0](http://dx.doi.org/10.1016/S0010-4655(99)00221-0).
- [16] Karl Blum. *Density matrix theory and applications; 3rd ed.* Springer Series on Atomic Optical and Plasma Physics. Berlin: Springer, 2012. DOI: [10.1007/978-3-642-20561-3](https://doi.org/10.1007/978-3-642-20561-3).
- [17] E. Hernández and M. J. Gillan. In: *Phys. Rev. B* 51 (15 1995), pp. 10157–10160. DOI: [10.1103/PhysRevB.51.10157](https://doi.org/10.1103/PhysRevB.51.10157).
- [18] Emilio Artacho and Lorenzo Miláns del Bosch. In: *Phys. Rev. A* 43 (11 1991), pp. 5770–5777. DOI: [10.1103/PhysRevA.43.5770](https://doi.org/10.1103/PhysRevA.43.5770).
- [19] Martin Head-Gordon, Paul E. Maslen, and Christopher A. White. In: *The Journal of Chemical Physics* 108.2 (1998), pp. 616–625. DOI: <http://dx.doi.org/10.1063/1.475423>.
- [20] R. McWeeny. In: *Rev. Mod. Phys.* 32 (2 1960), pp. 335–369. DOI: [10.1103/RevModPhys.32.335](https://doi.org/10.1103/RevModPhys.32.335).
- [21] Adam Palser and David Manolopoulos. In: *Physical Review B* 58.19 (1998), pp. 12704–12711. ISSN: 0163-1829. DOI: [10.1103/PhysRevB.58.12704](https://doi.org/10.1103/PhysRevB.58.12704).
- [22] P. D. Haynes, C. K. Skylaris, A. A. Mostofi, and M. C. Payne. In: *Chemical Physics Letters* 422.4–6 (2006), pp. 345–349. ISSN: 00092614. DOI: [10.1016/j.cplett.2006.02.086](https://doi.org/10.1016/j.cplett.2006.02.086).
- [23] Peter D. Haynes and Mike C. Payne. In: *Phys. Rev. B* 59.19 (1999), pp. 12173–12176.
- [24] X.-P. Li, R. W. Nunes, and David Vanderbilt. In: *Phys. Rev. B* 47 (16 1993), pp. 10891–10894. DOI: [10.1103/PhysRevB.47.10891](https://doi.org/10.1103/PhysRevB.47.10891).
- [25] R. W. Nunes and David Vanderbilt. In: *Physical Review B* 50.23 (1994), pp. 17611–17614. ISSN: 01631829. DOI: [10.1103/PhysRevB.50.17611](https://doi.org/10.1103/PhysRevB.50.17611).
- [26] P D Haynes, C-K Skylaris, A A Mostofi, and M C Payne. In: *Journal of Physics: Condensed Matter* 20.29 (2008), p. 294207.
- [27] John M. Millam and Gustavo E. Scuseria. In: *The Journal of Chemical Physics* 106.13 (1997), pp. 5569–5577. ISSN: 0021-9606. DOI: [10.1063/1.473579](https://doi.org/10.1063/1.473579).
- [28] S F Boys and F Bernardi. In: *Molecular Physics* 19.4 (1970), pp. 553–566. DOI: [10.1080/00268977000101561](https://doi.org/10.1080/00268977000101561).
- [29] Pablo Ordejón, David A. Drabold, Matthew P. Grumbach, and Richard M. Martin. In: *Phys. Rev. B* 48 (19 1993), pp. 14646–14649. DOI: [10.1103/PhysRevB.48.14646](https://doi.org/10.1103/PhysRevB.48.14646).

- [30] Per-Olov Lowdin. In: *The Journal of Chemical Physics* 18.3 (1950), pp. 365–375. DOI: <http://dx.doi.org/10.1063/1.1747632>.
- [31] Francesco Mauri, Giulia Galli, and Roberto Car. In: *Phys. Rev. B* 47 (15 1993), pp. 9973–9976. DOI: [10.1103/PhysRevB.47.9973](https://doi.org/10.1103/PhysRevB.47.9973).
- [32] Peter D. Haynes, Chris-Kriton Skylaris, Arash A. Mostofi, and Mike C. Payne. In: *physica status solidi (b)* 243.11 (2006), pp. 2489–2499. ISSN: 1521-3951. DOI: [10.1002/pssb.200541457](https://doi.org/10.1002/pssb.200541457).
- [33] Chris-Kriton Skylaris et al. In: *Phys. Rev. B* 66 (3 2002), p. 035119. DOI: [10.1103/PhysRevB.66.035119](https://doi.org/10.1103/PhysRevB.66.035119).
- [34] Chris-Kriton Skylaris et al. In: *Computer Physics Communications* 140.3 (2001), pp. 315–322. ISSN: 0010-4655. DOI: [http://dx.doi.org/10.1016/S0010-4655\(01\)00248-X](http://dx.doi.org/10.1016/S0010-4655(01)00248-X).
- [35] Arash A. Mostofi, Chris-Kriton Skylaris, Peter D. Haynes, and Mike C. Payne. In: *Computer Physics Communications* 147.3 (2002), pp. 788–802. ISSN: 0010-4655. DOI: [https://doi.org/10.1016/S0010-4655\(02\)00461-7](https://doi.org/10.1016/S0010-4655(02)00461-7).
- [36] N. D. M. Hine A. Ruiz-Serrano and C.-K. Skylaris. In: *J. Chem. Phys.* 136 (2012), p. 234101.
- [37] N. D. M. Hine et al. In: *Phys. Rev. B* 83 (2011), p. 195102.
- [38] Chris Kriton Skylaris, Oswaldo Diéguez, Peter D. Haynes, and Mike C. Payne. In: *Physical Review B - Condensed Matter and Materials Physics* 66.7 (2002), pp. 731031–731034. ISSN: 1550235X. DOI: [10.1103/PhysRevB.66.073103](https://doi.org/10.1103/PhysRevB.66.073103).
- [39] P Pulay. In: *Molecular Physics* 17.2 (1969), pp. 197–204. ISSN: 0026-8976. DOI: [10.1080/00268976900100941](https://doi.org/10.1080/00268976900100941).
- [40] Laura E. Ratcliff, Nicholas D. M. Hine, and Peter D. Haynes. In: *Phys. Rev. B* 84 (2011), p. 165131.
- [41] L. Ratcliff. *Optical Absorption Spectra Calculated Using Linear-Scaling Density-Functional Theory*. Imperial College London: Springer Theses, 2013.
- [42] J. Dziedzic, Q. Hill, and C.-K. Skylaris. In: *The Journal of Chemical Physics* 139.21, 214103 (2013). DOI: <http://dx.doi.org/10.1063/1.4832338>.
- [43] Erich Runge and E. K. U. Gross. In: *Phys. Rev. Lett.* 52 (12 1984), pp. 997–1000. DOI: [10.1103/PhysRevLett.52.997](https://doi.org/10.1103/PhysRevLett.52.997).
- [44] Giovanni Onida, Lucia Reining, and Angel Rubio. In: *Rev. Mod. Phys.* 74 (2 2002), pp. 601–659. DOI: [10.1103/RevModPhys.74.601](https://doi.org/10.1103/RevModPhys.74.601).
- [45] T. J. Zuehlsdorff et al. In: *The Journal of Chemical Physics* 139.6, 064104 (2013). DOI: <http://dx.doi.org/10.1063/1.4817330>.

- [46] T. J. Zuehlsdorff, N. D. M. Hine, M. C. Payne, and P. D. Haynes. In: *The Journal of Chemical Physics* 143.20 (2015), p. 204107. DOI: [10.1063/1.4936280](https://doi.org/10.1063/1.4936280).

## 5 Influence of the crystal host on excitons in pentacene derivatives

The key motivation behind the work in this chapter is the desire to obtain a better understanding of the pentacene in *p*-terphenyl room temperature maser (RTM), in particular regarding the interaction of the active pentacene molecule with its host structure. In this chapter, we perform an *ab initio* study of the ground state and excited state properties of this system using pre-existing electronic structure techniques. This work has previously been published in [1] and is reproduced here with the permission of AIP Publishing.

### 5.1 Background and Motivation

Pentacene ( $C_{22}H_{14}$ ) is an organic molecule in the class of linear acenes, consisting of five benzene rings fused together. As a family, the acene molecules have attracted a great deal of interest in the scientific world, particularly in the last decade due to their intriguing electronic properties which make them suitable for use in organic electronic devices [2]. Amongst their general properties that make acenes amenable for electronic applications is their high degree of mechanical flexibility and high charge-carrier mobility [3], which has enabled them and their derivatives to be used as thin-film transistors [2, 4–6]. Pentacene in particular has found use in organic photovoltaic systems [7], not least due to the observation of singlet fission within the pentacene excitation spectrum which enables extremely high efficiencies [8]. While these cases tend to deal with bulk crystalline and thin-film pentacene, for our purposes we are mainly concerned with the use of pentacene as a dopant molecule in a crystalline structure composed of *p*-terphenyl, which, by exploiting the excitation spectrum of pentacene, can be used as the gain medium for a room-temperature maser [9].

The masing mechanism is illustrated in Figure 5.1. Initially, the pentacene molecule is excited from the singlet ground state  $S_0$  to the lowest singlet excited state  $S_1$  using a yellow light laser. This state coincides in energy with the triplet excited state  $T_2$ , such that there is a strong mixing of the two states via a mixture of spin-orbit coupling and electron-phonon interactions. The result is a pair of states consisting of mixed singlet and triplet character,  $S_1$  being dominated by the former and  $T_2$  by the latter, allowing

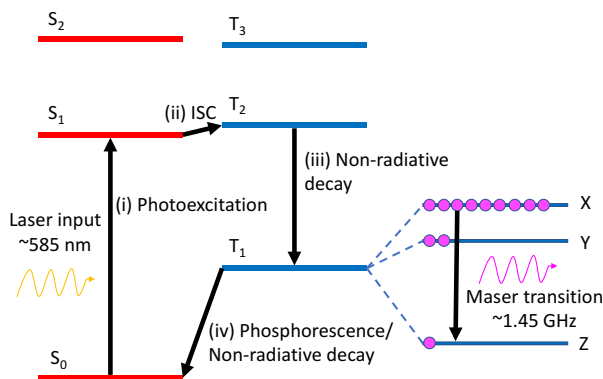


FIGURE 5.1: Pentacene excitation spectrum as part of the room-temperature maser mechanism cycle [9]. Starting from the singlet ground state  $S_0$ , (i) the pentacene is photoexcited by a yellow light laser into the  $S_1$  state. From there, (ii) an intersystem crossing (ISC) transfers the molecule to the triplet  $T_2$  state (see text for more details). The molecule then (iii) decays to the triplet ground state  $T_1$ , wherein a population inversion exists between the X, Y and Z states in the ratio 0.76:0.16:0.08. The maser transition then occurs within the  $T_1$  state. Finally, the molecule decays back to the  $S_0$  state. This figure is reproduced from [1].

an intersystem crossing (ISC) from the singlet to the triplet manifold [10] (note that, strictly speaking, spin is not a good quantum number due to spin-orbit coupling, such that the  $S_1 \rightarrow T_2$  ISC is a vertical transition between the mixed states). The sublevels of the  $T_2$  state are populated such that the uppermost level is preferred, giving rise to a population inversion that is preserved in the decay to the  $T_1$  triplet ground state. The maser transition takes place within the  $T_1$  state, before decaying back to  $S_1$ . This system is the first successful demonstration of masing at room-temperature, although further improvements are necessary before this can become a practical device [9, 10]. For example, the pentacene in *p*-terphenyl maser is a pulsed device and is not suitable for continuous operation. Another major constraint is the slow decay rate from  $T_1$  back to  $S_0$ , causing the accumulation of excitons in the lowest triplet sublevel and breaking the population inversion in  $T_1$ .

Experimental studies of pentacene have shown significant variation in its excitonic behaviour depending on the choice of host environment [11–13]. The change in molecular excitation energies due to exposure to solvent effects is known as a solvatochromic shift, and indeed for a *p*-terphenyl host even the choice of lattice site can be significant — Patterson *et al.* observed an increase in the ISC rate by two orders of magnitude due to the pentacene being placed in one of the four crystalline *p*-terphenyl lattice sites, corresponding to a shift of 20 meV in the excitation spectrum [14]. Thus in order to describe the behaviour of this system from first principles, it is clear that we must develop an understanding of not just pentacene in vacuum but also the impact

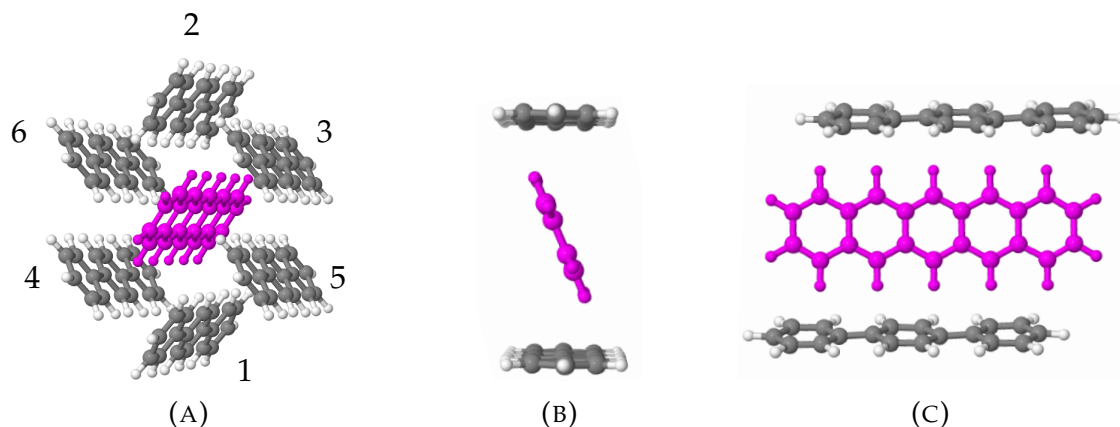


FIGURE 5.2: Various views of the pentacene (pink) in *p*-terphenyl (grey and white) cluster used in this work. (A) shows the complete cluster configuration with a pentacene molecule surrounded by its six nearest neighbour *p*-terphenyls, each labelled by a number. (B) shows the orientation of the molecules within the *p*-terphenyl herringbone crystal structure, with a pentacene swapped for one *p*-terphenyl molecule. (C) displays the alignment of the acene rings within the neighbouring molecules. This figure is reproduced from [1].

of the host structure. This would also enable us to consider other molecules that could replace pentacene as the active molecule, potentially yielding better performance for the maser. Past studies of the linear acene anthracene and its nitrogen-doped derivative phenazine, in which two  $sp^2$  nitrogen groups are contained within the central aromatic ring, showed a comparable zero-field splitting in both molecules but with a significantly more pronounced population inversion for phenazine, producing strong phosphorescence [15]. An equivalent effect for pentacene could potentially reduce the impact of the bottleneck in the return to  $S_0$ .

## 5.2 Chromophores and the crystal host

As discussed in Chapter 3, there are a number of ways to embed an active subsystem in a solvent environment [16]. Past studies on the importance of explicit solvent effects have focused on the inclusion of explicit solvent regions when performing QM/MM type calculations within an implicit solvent environment. For instance, solvation effects in water can be understood by including shells of water molecules located within a given radius of the solute in the explicit quantum mechanical region, while treating the remainder as a polarisable continuum medium (PCM). In these works, solvatochromic shifts were observed in the excitation spectra that varied by inclusion of additional solvent molecules in the active region, suggesting that at least

some environment is needed explicitly to accurately describe the solute-solvent interaction [17–19]. Our simulation of the pentacene in *p*-terphenyl system will be based on two models for the host — an implicit solvent using a conductor-like polarisable continuum medium (CPCM, see Chapter 3) and an explicit molecular host consisting of the six nearest neighbour molecules of pentacene in the *p*-terphenyl herringbone crystal structure as shown in Figure 5.2. Crystalline *p*-terphenyl has a dielectric constant of  $\epsilon_{\text{microwave}} = 2.98$  in the microwave spectrum, compared to an optical light value of  $\epsilon_{\text{optical}} = 2.81$  [20, 21]. Preliminary tests showed similar results with both models, thus a single value of  $\epsilon_{\text{CPCM}} = 3$  was selected for the CPCM to represent the host for implicit solvent calculations.

### 5.2.1 Choosing the cluster

The explicit crystal host was selected using a  $3 \times 4 \times 3$  periodic cell of 72 *p*-terphenyl molecules, the atomic positions and lattice parameters of which were taken from experiment [22]. At low temperatures, crystalline *p*-terphenyl has four distinct lattice positions, though in room temperature conditions as would be applicable to the maser, these four positions are averaged down to two in a herringbone configuration. A *p*-terphenyl molecule was replaced with a single pentacene, which has a similar length and width as *p*-terphenyl (see Figure 5.2) and so can maintain the same orientation and centre of mass. The geometry was subsequently reoptimised at the PBE level of theory in ONETEP, the linear-scaling nature of which makes geometry optimisation on structures containing thousands of atoms feasible [23, 24]. Since PBE lacks dispersion and therefore cannot accurately capture long-range effects, the *p*-terphenyl carbon atom positions were fixed in their correct experimental locations, with only the carbon atoms in pentacene and the hydrogen atoms across the crystal being allowed to reoptimise. The lattice vectors were kept fixed for all geometry optimisations.

The full periodic crystal structure is too large to be of use for excited state calculations with hybrid functionals, so a cluster consisting of the pentacene and its six nearest neighbour *p*-terphenyls was extracted to vacuum as the representation of the explicit host. The size of this structure was chosen via an analysis of the HOMO and LUMO states in the crystal. Due to the semi-local nature of the PBE functional, the HOMO and LUMO states both spill out somewhat onto the neighbouring *p*-terphenyl molecules in the crystal host. Since the  $S_1$  and  $T_1$  excited states in vacuum are primarily composed of the HOMO→LUMO transition, it is important that any effects that the explicit host may have on these transitions is accurately captured by the extracted cluster structure. Thus the proportion of the HOMO/LUMO states localised on the pentacene and its neighbouring molecules was determined by partitioning the Kohn-Sham orbitals



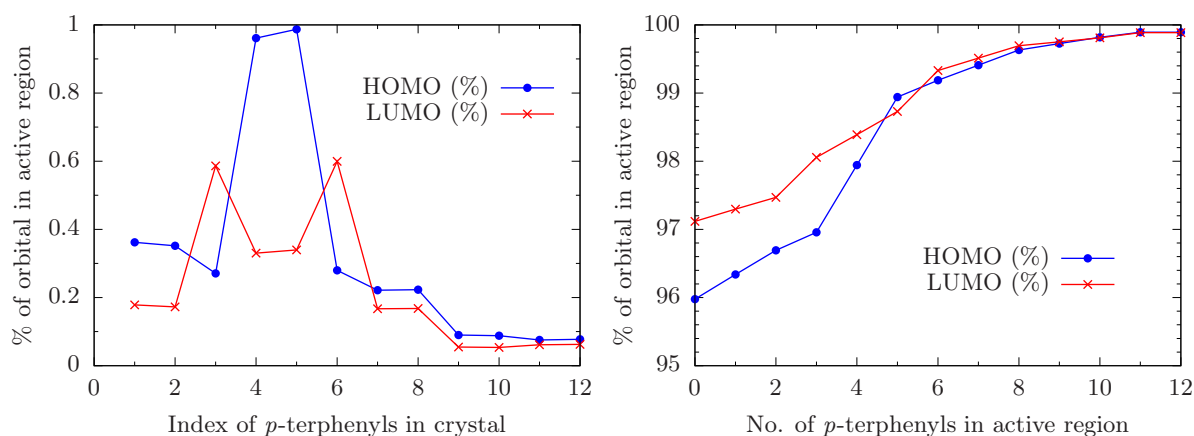


FIGURE 5.3: Left: contribution to HOMO and LUMO states of pentacene by each of its 12 nearest neighbour *p*-terphenyls in the periodic crystal. The first six data points correspond to the six molecules in Figure 5.2 and the remainder are the next nearest neighbours. Right: proportion of HOMO/LUMO in active region. This figure is reproduced from the supplementary material of [1].

based on contributions from the local orbitals assigned to each molecule. Indeed, in the periodic *p*-terphenyl crystal structure, only 96.0% of the HOMO and 97.1% of the LUMO are localised on the pentacene chromophore, with the remainder being spread primarily across the 12 nearest *p*-terphenyls. Figure 5.3 shows the contributions towards the HOMO and LUMO states from each of these molecules, with the first six indices corresponding to the molecules indicated in Figure 5.2. We see that including the six nearest neighbour *p*-terphenyls in the active region enables the cluster to reproduce over 99% of the crystalline HOMO/LUMO states. Thus the cluster of pentacene surrounded by six *p*-terphenyls was selected for explicit host structure calculations.

## 5.2.2 Pentacene-based derivatives

As discussed earlier, nitrogen doping of acene molecules in past experimental studies has resulted in an increased population inversion compared to undoped molecules, enabling stronger phosphorescence and thus potentially a faster ISC [15]. This could potentially improve the efficiency of the pentacene in *p*-terphenyl maser. For this reason, four additional chromophores were considered, all of which include the doping of pentacene with nitrogen. The four shown in Figure 5.4 are known as diazapentacene (dap) molecules, labelled 6,13-dap, 5,12-dap, 1,8-dap and 2,9-dap according to the location of the two nitrogen atoms. These were chosen based on a previous *ab initio* study by Bogatko *et al.* [25], whereby a variety of alternative maser active molecules



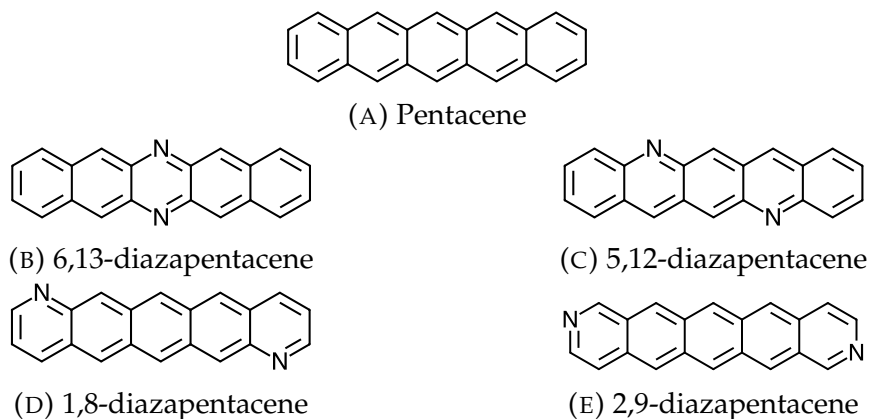


FIGURE 5.4: Dopant molecules considered as part of this study. For brevity, each diazapentacene molecule will be referred to by the shorthand “m,n-dap” throughout this work. This figure is reproduced from [1].

	Chromophore optimised in crystal (mHa/ $a_0$ )	Chromophore extracted to vacuum (mHa/ $a_0$ )
Pentacene	1.25	5.35
1-8 dap	1.90	4.48
2-9 dap	1.59	3.40
5-12 dap	1.48	3.90
6-13 dap	1.68	3.79

TABLE 5.1: Maximum root mean square (RMS) forces on each chromophore for crystal-optimised geometry in crystal host and vacuum. This table is reproduced from the supplementary material of [1].

were screened using time-dependent density functional theory (TDDFT) to find a selection of molecules that bear qualitative excitonic similarity to pentacene. The diazapentacene structures were selected for this study due to the similar size and shape of these molecules to pentacene, such that they can still fit within the *p*-terphenyl herringbone lattice structure. We used the same procedure for optimisation of the pentacene structure in crystal with all of the nitrogen-doped analogues studied here. The maximum root mean square (RMS) forces for each molecule are displayed in Table 5.1 both in the crystal environment and extracted to vacuum. We see that the forces for all five molecules are of comparable magnitude for both configurations, indicating that the diazapentacene molecules are a good fit for the *p*-terphenyl crystal and are not substantially distorted compared to the pure vacuum arrangement.

### 5.3 Excited state methods

Although density functional theory (DFT) is often thought of as a ground-state method, and thus not suitable for the study of excited states, it is possible to extract information on excited states from standard DFT calculations. One possibility is known as the  $\Delta$ SCF method, whereby the lowest triplet excitation energy is obtained from two separate SCF calculations. In the case of pentacene, since the lowest-lying triplet state  $T_1$  is also the triplet ground state, the triplet excitation energy  $\Delta E_{S_0 \rightarrow T_1}$  can be obtained from the energy difference between a spin-polarised and a spin-unpolarised calculation,

$$\Delta E_{S_0 \rightarrow T_1} = E_{\text{pol}} - E_{\text{unp}}. \quad (5.1)$$

$T_1$  is a special case, however, and for other excited states it is necessary to utilise dedicated excitonic methods such as time-dependent DFT (TDDFT) [26]. Excitation energies can be calculated in the linear-response form of TDDFT [27–29], whereby the TDDFT equations are recast as effective eigenvalue equations. TDDFT retains a theoretical similarity to regular ground-state DFT, along with a scalability that enables the application of TDDFT to large systems such as the pentacene in *p*-terphenyl cluster used as part of this study. Linear-response TDDFT, in particular, can be made linear-scaling [30–32]. Nonetheless, TDDFT does have its drawbacks, such as the well-documented failure of semi-local exchange-correlation functionals to capture the excited states of linear acenes [33, 34]. In particular, singlet energies are severely underestimated due to the failure of semi-local TDDFT to describe the  $\pi$ -conjugated states of pentacene. This limits the applicability of functionals such as PBE, and indeed has detrimental effects for hybrid functionals such as B3LYP which also struggle to accurately determine such states [35]. Instead, our focus here will be on the solvatochromic shifts between vacuum and host calculations for comparisons of semi-local and hybrid functionals. This raises the question of whether the failure of semi-local functionals to capture acene excitation energies is systematic and thus if such comparisons are valid. Long-range interactions in particular are not well represented by such functionals, causing charge-transfer (CT) states to be unphysically low in energy.

Self-interaction errors can only be truly eradicated using exact exchange (see Section 2.4), enabling hybrid functionals to provide a more realistic description of the CT states and avoid the difficulties associated with semi-local functionals in explicit host. Nonetheless, standard hybrid functionals also struggle to produce accurate excitation energies for linear acenes [35], with results severely underestimated relative to more accurate quantum chemical methods [11]. Optimally-tuned range-separated hybrid functionals (OT-RSH, see Section 2.4.3) have in past studies produced highly accurate results for systems that are difficult for semi-local TDDFT, such as acenes and

CT excitations [36–38]. For this reason, in this study we make use of the OT-LC $\omega$ PBE functional, which has a 100% semi-local exchange form at short range and 100% exact exchange at long range [39]. As these functionals are not currently available in ONETEP, all excited state calculations presented in this chapter were performed using Gaussian 09 [40] with the cc-pVTZ basis set [41]. For each chromophore, the range separation parameter  $\lambda$  was tuned in vacuum such that the ionisation energy matched the negative of the HOMO eigenvalue energy. These vacuum-tuned functionals were also used for all calculations in implicit and explicit host.

## 5.4 Vacuum results

We will now look more closely at our choice of dopant molecules by examining their ground and excited state properties in vacuum. Ground state calculations were done using the linear-scaling ONETEP code, using the PBE exchange-correlation functional with a kinetic energy cutoff of 750 eV. A minimal basis set of non-orthogonal generalised Wannier functions (NGWFs) was used, with 1 per H, 4 per C and 4 per N. All NGWFs have a radius of 10.0 bohr. Excited state calculations were performed using Gaussian 09 [40] with the cc-pVTZ basis set [41], chosen such that the excited state energies using the PBE functional for all chromophores in vacuum matched the ONETEP excitation energies for  $\Delta E_{S_0 \rightarrow S_1}$  and  $\Delta E_{S_0 \rightarrow T_1}$  using the NGWF basis mentioned above to within  $\pm 5$  meV. This ensured that, as much as possible, the ONETEP ground state and Gaussian excited state calculations used equivalent basis sets.

### 5.4.1 Natural bond orbital (NBO) analysis

Nitrogen is a more electronegative element than carbon, so one would expect that the substitution of N in place of C-H for the diazapentacene molecules would cause measurable distortions in the electronic structure compared to pentacene. This can be evaluated via a natural population analysis (NPA) of the molecules to determine the electron population associated with each atom in a natural bond orbital (NBO) basis set [42]. Figure 5.5 illustrates the population difference associated with each atom for the diazapentacenes relative to pentacene. As expected, electric charge accumulates on the nitrogens, such that there is a net reduction in the overall charge of the N sites relative to the corresponding C-H sites of pentacene. These results are consistent with previous studies of the impact of nitrogen substitution on pentacene [43], along with the formation of electric dipoles in pyridine (nitrogen-doped benzene) due to charge redistribution across the acene ring, reducing the spatial extent of the  $\pi$ -electron density.

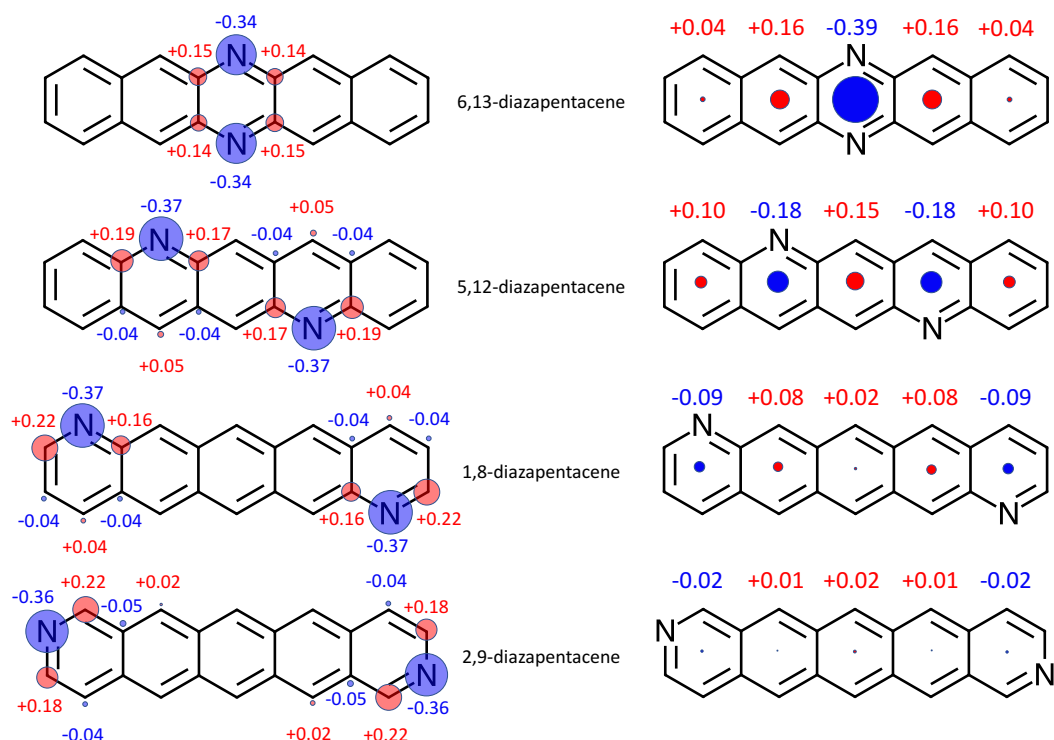


FIGURE 5.5: Changes in the electron populations associated with each atom and acene ring for the diazapentacene molecules compared to pentacene using natural population analysis (NPA). Circle radii are proportional to the net electron population indicated. Nitrogen population differences are compared to the combined C-H value for the corresponding position on the pentacene molecule. This figure is reproduced from [1].

The acene rings of pentacene are not entirely equivalent in terms of their charge distributions, which can cause the inner rings to be more chemically reactive than the outer rings. Using our NPA, we do indeed see a small but measurable charge redistribution across the molecule, such that the three inner rings each possess a net negative charge of  $-0.004e$  and the two outer rings possess a  $+0.006e$  net positive charge. Given the accumulation of electric charge in nitrogen-doped sites seen above, the next question is whether this has any overall impact on the charge associated with each ring. Figure 5.5 also shows the change in the net population for each ring relative to pentacene. The most significant change is seen for 6,13-dap, which accumulates  $-0.39e$  in the central ring where the two nitrogens are located, plus a positive charge of  $+0.16e$  on each of the neighbouring rings. A similar though less striking effect can be seen for 5,12-dap, where there is a substantial charge variation on each ring though it is more even than for 6,13-dap. 1,8- and 2,9-dap show comparatively little variation, seemingly due to much of the charge redistribution being isolated within the outermost rings where the nitrogens are located, as can be seen by examining the atomic population changes. Nonetheless, for all chromophores the charge variation is significantly

	$S_1$ (eV)	$T_1$ (eV)
Experiment [13]	2.3	0.86
Multi-reference Møller-Plesset perturbation theory (MRMP) [44]	2.31	0.87
CCSD(T) (vertical excitation) [45]	-	1.37
PBE TDDFT	1.62	0.81
OT-LC $\omega$ PBE TDDFT	2.15	0.58
PBE $\Delta$ SCF	-	0.91
OT-LC $\omega$ PBE $\Delta$ SCF	-	1.08

TABLE 5.2: Singlet and triplet excitation energies for pentacene in vacuum from experimental and *ab initio* studies. TDDFT and  $\Delta$ SCF results were performed as part of this work, while MRMP and CCSD(T) are taken from the references indicated. This table is reproduced from [1].

greater than between the individual acene rings of pentacene, indicating that there is indeed a substantial change in the electronic structure for each diazapentacene structure. Based on these results, we would expect the 6,13-dap molecule to show the greatest divergence in behaviour compared to pentacene, with 1,8- and 2,9-dap remaining comparatively similar.

### 5.4.2 Vacuum excitations

In this section we will look more carefully at the excited state properties of pentacene and its nitrogen-based derivatives. For this purpose we will use two methods to calculate the excitation energies  $\Delta E_{S_0 \rightarrow S_1}$  and  $\Delta E_{S_0 \rightarrow T_1}$ , representing transitions from the  $S_0$  ground state to the singlet excited state  $S_1$  and the  $T_1$  triplet ground state respectively. For both states we make use of LR-TDDFT for determining both vacuum and host excitation energies, while for the triplet state only we also use  $\Delta$ SCF. Both the semi-local PBE functional and the OT-LC $\omega$ PBE functional were used in all calculations for comparison.

Table 5.2 shows the vacuum energy levels of the  $S_1$  and  $T_1$  states for pentacene, along with experimental and reference *ab initio* results. As expected, TDDFT with PBE significantly underestimates the singlet excitation energy  $\Delta E_{S_0 \rightarrow S_1}$  for pentacene, while OT-LC $\omega$ PBE yields much more accurate results for the  $S_1$  state. By contrast, for the triplet state  $T_1$  PBE generally yields results closer to both experiment and wavefunction based quantum chemical methods than OT-LC $\omega$ PBE, regardless of method used.  $\Delta$ SCF overestimates  $\Delta E_{S_0 \rightarrow T_1}$  with range-separated hybrids relative to experiment, while LR-TDDFT underestimates it significantly. The former is in line with the tendency of quantum chemical methods to overestimate the difference between the singlet and triplet ground states of pentacene [45, 46], while the latter is reminiscent of

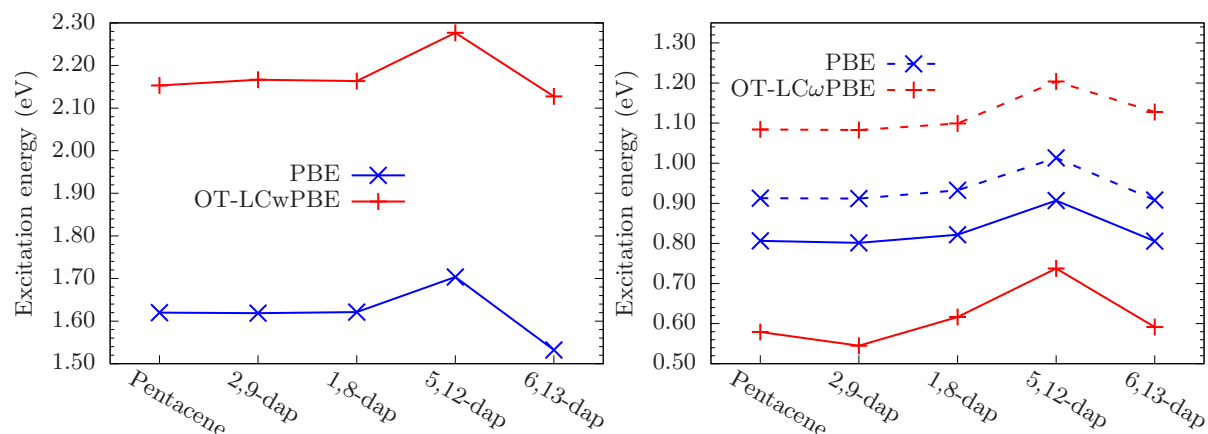


FIGURE 5.6: Singlet and triplet excitation energies  $\Delta E_{S_0 \rightarrow S_1}$  and  $\Delta E_{S_0 \rightarrow T_1}$  for each chromophore in vacuum with PBE and OT-LC $\omega$ PBE exchange-correlation functionals. TDDFT calculations (solid lines) were performed for both singlet and triplet states, and  $\Delta$ SCF calculations (broken lines) were used for triplet states only. This figure is reproduced from [1].

the triplet instability associated with linear-response TDDFT [47–49]. Thus the most accurate singlet results are obtained using LR-TDDFT with OT-LC $\omega$ PBE, while for triplet energies the closest results to experiment for pentacene can be found with  $\Delta$ SCF and PBE.

Turning our attention to the other chromophores, Figure 5.6 displays the results for the  $\Delta E_{S_0 \rightarrow S_1}$  and  $\Delta E_{S_0 \rightarrow T_1}$  excitation energies for all five of our molecules in vacuum. Generally, 1,8- and 2,9-dap closely match pentacene for all levels of theory, while 5,12-dap produces higher excitation energies. 6,13-dap yields lower singlet excitations but the vacuum triplet energy is similar to that of pentacene. The similarity of 1,8- and 2,9-dap confirms the predictions made using our NBO analysis previously, whereby these two molecules show the least disturbance in their ground state electronic structures compared to pentacene. 5,12-dap does show significant divergence in its excitation spectrum, moreso than 6,13-dap, though this should be interpreted with caution due to the difficulties in calculating accurate excitation energies for acene systems outlined above. From the perspective of the maser, it would be more useful to consider the relative behaviours of these molecules in host structures, as this presents a more meaningful comparison of the molecular candidates and computational effects as opposed to interpreting absolute energies.

## 5.5 Host effects

In this section we focus our efforts on the impact of the host on the various chromophores discussed so far using the implicit CPCM model and the explicit *p*-terphenyl

cluster. Based on the analysis of the vacuum pentacene results from earlier, we will only work with the LR-TDDFT method for singlet calculations and the  $\Delta$ SCF method for triplet calculations.

Figure 5.7 displays the solvatochromic shift in  $\Delta E_{S_0 \rightarrow S_1}$  and  $\Delta E_{S_0 \rightarrow T_1}$  for each chromophore both in implicit and explicit host. The singlet results for CPCM show a PBE redshift of 30–40 meV less than the OT-LC $\omega$ PBE for each chromophore, with 2,9- and 6,13-dap producing greater redshifts than pentacene while 1,8- and 5,12-dap show a reduced shift. The singlet results for CPCM show a PBE redshift of 30–40 meV less than the OT-LC $\omega$ PBE for each chromophore, showing the same trend in both cases of 2,9- and 6,13-dap producing greater redshifts than pentacene while 1,8- and 5,12-dap show a reduced shift. The explicit host cluster indicates a far greater divergence between the two functionals for singlet excitations. While there is a greater redshift for pentacene with OT-LC $\omega$ PBE than for PBE, the nitrogen-doped analogues display the opposite behaviour, with the PBE redshift consistently increasing as the nitrogen atoms are brought closer together. The scale of this disparity can be seen clearly from a comparison of the  $S_1$  excitation energy for pentacene and 6,13-dap in cluster— with OT-LC $\omega$ PBE, the former shows a redshift of 65 meV, compared to 92 meV for the latter. However, with PBE, the pentacene redshift amounts to 58 meV compared to the 6,13-dap result of 189 meV.

For  $T_1$  in CPCM, an insignificant redshift of just 2 meV is observed for pentacene with both OT-LC $\omega$ PBE and PBE. Similarly, all the diazapentacene molecules apart from 6,13-dap also display small energy shifts in CPCM, with a minor redshift for 2,9-dap compared to a blueshift of 6 meV and 11 meV for 1,8- and 5,12-dap for OT-LC $\omega$ PBE, a trend also followed using semi-local functionals. 6,13-dap produces the greatest overall solvatochromic shift of all the chromophores, amounting to a 15 – 18 meV redshift. Turning our attention to the explicit cluster host, both functionals again predict the same trends, with PBE suggesting a slightly greater redshift in all cases. Once again, the standout result is for 6,13-dap, which yields a redshift of 30–35 meV relative to vacuum, far greater than any other molecule for  $T_1$  in either host.

### 5.5.1 Discussion

In order to understand the impact of the host on the chromophores, we must disentangle two effects — the importance of the choice of exchange-correlation functional and the difference between the explicit and implicit hosts. The key result which requires explanation is the large redshift in  $\Delta E_{S_0 \rightarrow S_1}$  for 5,12- and 6,13-dap in explicit host. An examination of the composition of these excitations in Figure 5.8 in terms of the contribution of the HOMO  $\rightarrow$  LUMO transition to the excitation reveals the cause of



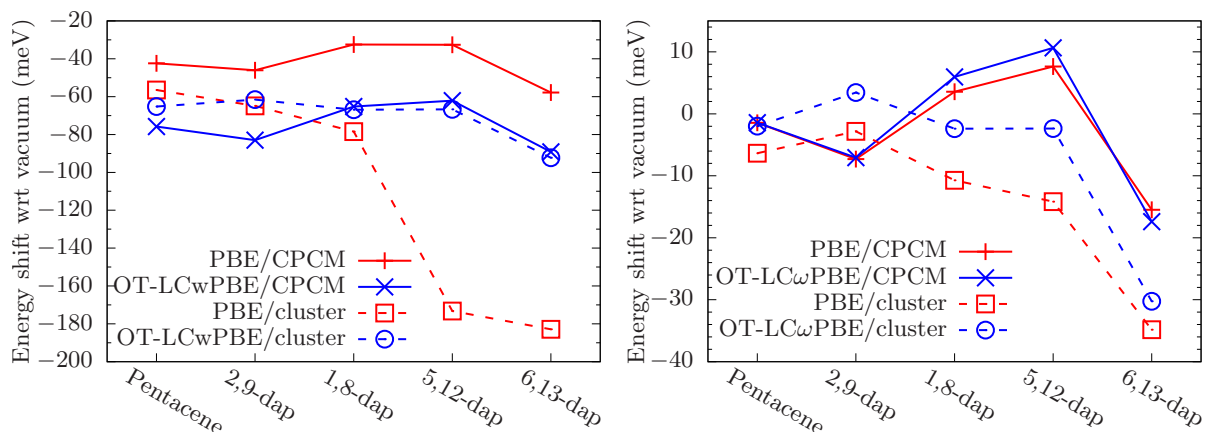


FIGURE 5.7: Solvatochromic shifts in  $\Delta E_{S_0 \rightarrow S_1}$  (left) and  $\Delta E_{S_0 \rightarrow T_1}$  (right) by placing each chromophore in a conductor-like polarisable continuum model (CPCM, solid) and an explicit *p*-terphenyl cluster (dashed lines). Singlet calculations were performed using TDDFT and triplet calculations using  $\Delta$ SCF at either the PBE or OT-LC $\omega$ PBE level of theory. This figure is reproduced from [1].

this effect. The normally dominant HOMO $\rightarrow$ LUMO in vacuum (and CPCM) becomes slightly diluted in explicit host due to the availability of additional molecular orbitals from the neighbouring *p*-terphenyl molecules. While pentacene, 2,9- and 1,8-dap retain their  $> 99\%$  HOMO $\rightarrow$ LUMO makeup in explicit host with OT-LC $\omega$ PBE, this is not the case for 5,12- and 6,13-dap, which now contain significant contributions from other molecules in their singlet excited state. This effect is even more dramatic with PBE, for which only 56% and 66% of the  $S_1$  excitation are explicitly HOMO $\rightarrow$ LUMO for 5,12- and 6,13-dap respectively, due to the inclusion of portions of spurious charge-transfer (CT) states in the excitation. In this context, it is easy to see why the PBE cluster redshift is so large — the CT states are unphysically low in energy for 5,12- and 6,13-dap, becoming mixed into their excitation spectra and lowering their overall energy relative to vacuum. This suggests that the  $\Delta E_{S_0 \rightarrow S_1}$  PBE redshifts for 5,12- and 6,13-dap are not physically meaningful and are simply a consequence of the failure of semi-local TDDFT to describe long-range CT interactions. This indicates that for singlet states in explicit host PBE is not a useful functional as it predicts unphysical behaviour.

Nonetheless, the fact that a qualitatively similar effect regarding the dilution of the  $S_1$  state is observable with OT-LC $\omega$ PBE, which should be free of such distortions, implies that there is a significant change to the excitonic properties of these two molecules in the presence of the explicit host. This is not represented by the CPCM data in Figure 5.8. To better understand this effect, Figure 5.9 shows the local density of states (LDOS) plots for each molecule in explicit host with OT-LC $\omega$ PBE. Energetically, the



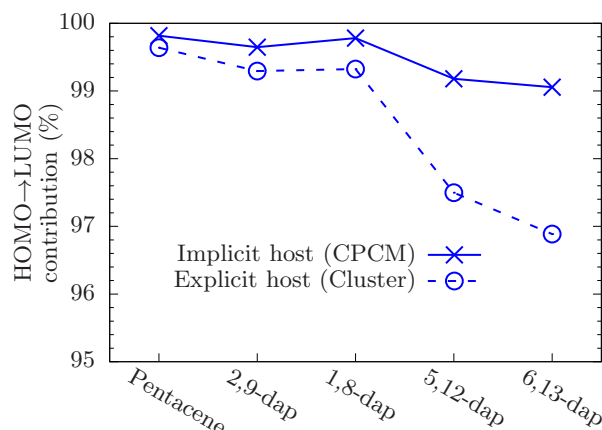


FIGURE 5.8: Proportion of the  $S_1$  excited state that consists of the HOMO→LUMO Kohn-Sham transition vector for each chromophore in both implicit (CPCM) and explicit (cluster) host. All calculations were performed using TDDFT with the OT-LC $\omega$ PBE functional. This figure is reproduced from [1].

HOMO and LUMO of vacuum pentacene are isolated from the rest of the molecular orbitals, but placing the pentacene in explicit host causes additional *p*-terphenyl molecular orbitals to be located above the HOMO-1 and below the LUMO+1 states. Nitrogen-doping brings these host states closer in energy to the HOMO of the chromophore, an effect that becomes more noticeable as the two nitrogens of the diazapentacene structure are brought closer together. In this way, the nitrogen lone pairs provide a bridging mechanism between the pentacene HOMO/LUMO and the orbitals of *p*-terphenyl, enabling the CT mixing observed in Figure 5.8.

For the  $S_1$  state, the two solvent models produce very similar behaviour for 1,8-, 5,12- and 6,13-dap with OT-LC $\omega$ PBE, though they diverge for the other two molecules. This does not capture the full picture, however, as the CPCM predicts a smaller redshift in  $\Delta E_{S_0 \rightarrow S_1}$  for 1,8- and 5,12-dap than pentacene, in contrast to the cluster structure results, as well as underestimating the shift for 6,13-dap relative to pentacene. By contrast, for  $\Delta E_{S_0 \rightarrow T_1}$  the two solvent models yield different trends despite predicting the same redshift with OT-LC $\omega$ PBE for pentacene. These results suggest that the CPCM does not provide a full description of the impact of the *p*-terphenyl host structure on the diazapentacenes, despite giving a good description of the electrostatic interaction of the pentacene in *p*-terphenyl solute-solvent system, as expected by our earlier analysis of the chromophore excited states.

From a structural perspective, the CT coupling between the chromophore and its host appears to depend significantly on the location of the nitrogen atoms in the diazapentacene structure, as only 5,12- and 6,13-dap show significant divergence from pentacene in their excitation spectrum in explicit host. This is in line with the NBO

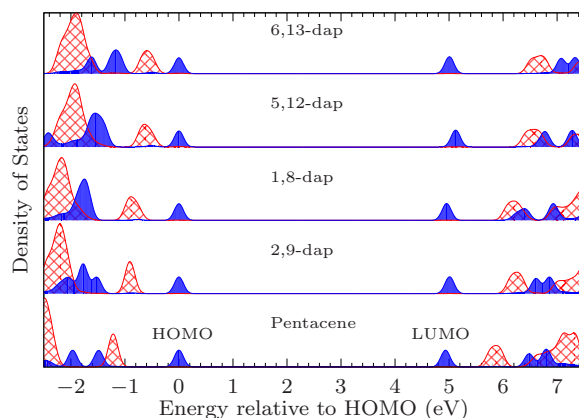


FIGURE 5.9: Local density of states (LDOS) analysis of the various chromophores in explicit host using OT-LC $\omega$ PBE. Chromophore states are shown in blue, *p*-terphenyl states in red. This figure is reproduced from [1].

analysis of Section 5.4.1, whereby significant charge redistribution across the acene structure are only seen for the latter two molecules, while the former two retain an electronic structure similar to pentacene. The location of the nitrogens themselves relative to the acene rings of the *p*-terphenyls (see Figure 5.2) is another possible cause of this effect, since only 6,13-dap retains mirror symmetry through the major axis plane while the other three molecules have two possible configurations within the cluster corresponding to the alternative lattice positions in the herringbone structure. However, rotating these three molecules within the cluster yielded no appreciable change in the excitation energies or molecular orbital composition, so we must rule this out.

## 5.6 Maser candidates

Given the results presented thus far, we are finally in a position to evaluate the suitability of each diazapentacene molecule as maser candidates. The similarity of 1,8- and 2,9-dap to pentacene in their ground and excited state properties both in vacuum and host suggests that they have the potential to function as maser active molecules. 5,12-dap behaves in a similar manner to pentacene in the presence of the *p*-terphenyl host but has the drawback that a significantly higher  $\Delta E_{S_0 \rightarrow T_1}$  energy would slow the decay back to the singlet state at the end of the masing process. The electronic structure of 6,13-dap is significantly distorted by the presence of the nitrogens on the central acene ring, promoting a stronger interaction with the host environment and yielding a greater redshift than any of the other chromophore candidates. In particular, the decrease in  $\Delta E_{S_0 \rightarrow T_1}$  in host could produce a faster decay rate from  $T_1$  back to  $S_0$  reducing the risk of bottlenecking in the maser process that impacts the pentacene in *p*-terphenyl

device. While an analysis of the impact of nitrogen-doping on the intersystem crossing rates and populations is beyond this study, based on previous studies of phenazine and anthracene it is possible that the introduction of the nitrogen lone pairs would facilitate a stronger population inversion and thus produce a faster ISC rate. Combined with the results discussed above, this would enable the tuning of the maser device via selective doping of pentacene.

## 5.7 Summary

In this chapter we have presented the results of a ground and excited state study of the excitonic properties of pentacene and its nitrogen-based analogues in explicit and implicit *p*-terphenyl hosts. We have demonstrated the impact that nitrogen-doping can have on the electronic structure of pentacene both in vacuum and in host structures, which enabled us to consider the viability of each of the test molecules as potential replacements for pentacene in the room-temperature maser. From the perspective of *ab initio* methods for dopant molecules in large host structures, we saw that the use of implicit hosts is insufficient to describe the full range of effects associated with such molecular systems. With explicit host, the semi-local PBE exchange-correlation functional fails to accurately describe long-range charge-transfer interactions, producing physically meaningless results and requiring the use of methods that include portions of exact exchange. Such methods can be computationally very expensive, which restricts their use to systems of a few hundred atoms such as the cluster structure used in this study. The combination of linear-scaling DFT to study ground state properties and to produce the explicit host structure, and advanced hybrid functionals for excited states has thus been demonstrated to be a useful combination, though it still leaves room for improvement.

## References

- [1] R. J. Charlton et al. In: *The Journal of Chemical Physics* 148.10 (2018), p. 104108. ISSN: 0021-9606. DOI: [10.1063/1.5017285](https://doi.org/10.1063/1.5017285).
- [2] John E. Anthony. In: *Chemical Reviews* 106.12 (2006). PMID: 17165682, pp. 5028–5048. DOI: [10.1021/cr050966z](https://doi.org/10.1021/cr050966z).
- [3] Ingrid M. Graz and Stéphanie P. Lacour. In: *Applied Physics Letters* 95.24 (2009). ISSN: 00036951. DOI: [10.1063/1.3265737](https://doi.org/10.1063/1.3265737).
- [4] Marcia M. Payne et al. In: *Journal of the American Chemical Society* 127.14 (2005), pp. 4986–4987. ISSN: 00027863. DOI: [10.1021/ja042353u](https://doi.org/10.1021/ja042353u).
- [5] Norbert Koch. In: *ChemPhysChem* 8.10 (2007), pp. 1438–1455. ISSN: 14394235. DOI: [10.1002/cphc.200700177](https://doi.org/10.1002/cphc.200700177).
- [6] Ming L Tang et al. In: *Journal of the American Chemical Society* 130.19 (2008), pp. 6064–6065. ISSN: 0002-7863. DOI: [10.1021/ja8005918](https://doi.org/10.1021/ja8005918).
- [7] S. Yoo, B. Domercq, and B. Kippelen. In: *Applied Physics Letters* 85.22 (2004), pp. 5427–5429. ISSN: 00036951. DOI: [10.1063/1.1829777](https://doi.org/10.1063/1.1829777).
- [8] Millicent B. Smith and Josef Michl. In: *Annual Review of Physical Chemistry* 64.1 (2013), pp. 361–386. ISSN: 0066-426X. DOI: [10.1146/annurev-physchem-040412-110130](https://doi.org/10.1146/annurev-physchem-040412-110130).
- [9] Mark Oxborrow, Jonathan D. Breeze, and Neil M. Alford. In: *Nature* 488.7411 (2012), pp. 353–356. ISSN: 0028-0836. DOI: [10.1038/nature11339](https://doi.org/10.1038/nature11339).
- [10] Jonathan Breeze et al. In: *Nat. Commun.* 6.6215 (2015). DOI: [10.1038/ncomms7215](https://doi.org/10.1038/ncomms7215).
- [11] Paul M. Zimmerman, Zhiyong Zhang, and Charles B. Musgrave. In: *Nat Chem* 2.8 (2010), pp. 648–652.
- [12] J. Köhler, A.C.J. Brouwer, E.J.J. Groenen, and J. Schmidt. In: *Chemical Physics Letters* 250.1 (1996), pp. 137–144. ISSN: 0009-2614. DOI: [http://dx.doi.org/10.1016/0009-2614\(95\)01393-8](http://dx.doi.org/10.1016/0009-2614(95)01393-8).
- [13] E. Heinecke, D. Hartmann, R. Müller, and A. Hese. In: *The Journal of Chemical Physics* 109.3 (1998), pp. 906–911. DOI: <http://dx.doi.org/10.1063/1.476631>.
- [14] F.G. Patterson, H.W.H. Lee, William L. Wilson, and M.D. Fayer. In: *Chemical Physics* 84.1 (1984), pp. 51–60. ISSN: 0301-0104. DOI: [http://dx.doi.org/10.1016/0301-0104\(84\)80005-1](http://dx.doi.org/10.1016/0301-0104(84)80005-1).
- [15] Bernhard Nickel and A. Andreas Ruth. In: *The Journal of Physical Chemistry* 95.5 (1991), pp. 2027–2036. DOI: [10.1021/j100158a026](https://doi.org/10.1021/j100158a026).

- [16] Andre Severo Pereira Gomes and Christoph R. Jacob. In: *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* 108 (1 2012), pp. 222–277. DOI: [10.1039/C2PC90007F](https://doi.org/10.1039/C2PC90007F).
- [17] T. J. Zuehlsdorff et al. In: *Journal of Chemical Theory and Computation* 12.4 (2016). PMID: 26967019, pp. 1853–1861. DOI: [10.1021/acs.jctc.5b01014](https://doi.org/10.1021/acs.jctc.5b01014).
- [18] T. J. Zuehlsdorff, P. D. Haynes, M. C. Payne, and N. D. M. Hine. In: *The Journal of Chemical Physics* 146.12 (2017), p. 124504. DOI: [10.1063/1.4979196](https://doi.org/10.1063/1.4979196).
- [19] Makenzie R. Provorse, Thomas Peev, Chou Xiong, and Christine M. Isborn. In: *The Journal of Physical Chemistry B* 120.47 (2016). PMID: 27797196, pp. 12148–12159. DOI: [10.1021/acs.jpcc.6b09176](https://doi.org/10.1021/acs.jpcc.6b09176).
- [20] Vincenzo Barone and Maurizio Cossi. In: *The Journal of Physical Chemistry A* 102.11 (1998), pp. 1995–2001. ISSN: 1089-5639. DOI: [10.1021/jp9716997](https://doi.org/10.1021/jp9716997).
- [21] Ishii Kikujiro, Kinoshita Minoru, and Kuroda Haruo. In: *Bulletin of the Chemical Society of Japan* 46.11 (1973), pp. 3385–3391. DOI: [10.1246/bcsj.46.3385](https://doi.org/10.1246/bcsj.46.3385).
- [22] H. M. Rietveld, E. N. Maslen, and C. J. B. Clews. In: *Acta Crystallographica Section B* 26.6 (1970), pp. 693–706. DOI: [10.1107/S0567740870003023](https://doi.org/10.1107/S0567740870003023).
- [23] N. D. M. Hine et al. In: *Phys. Rev. B* 83 (2011), p. 195102.
- [24] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. In: *Phys. Rev. Lett.* 77 (18 1996), pp. 3865–3868. DOI: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [25] Stuart Bogatko et al. In: *The Journal of Physical Chemistry C* 120.15 (2016), pp. 8251–8260. DOI: [10.1021/acs.jpcc.6b00150](https://doi.org/10.1021/acs.jpcc.6b00150).
- [26] Erich Runge and E. K. U. Gross. In: *Phys. Rev. Lett.* 52 (12 1984), pp. 997–1000. DOI: [10.1103/PhysRevLett.52.997](https://doi.org/10.1103/PhysRevLett.52.997).
- [27] G Onida, L Reining, and A Rubio. In: *Rev. Mod. Phys.* 74.April (2002), p. 601.
- [28] M. A. L. Marques et al. *Time-Dependent Density Functional Theory*. Vol. 706. 11. 2006, pp. 427–455. ISBN: 978-3-540-35422-2. DOI: [10.1007/b11767107](https://doi.org/10.1007/b11767107).
- [29] M.E. Casida and M. Huix-Rotllant. In: *Annual Review of Physical Chemistry* 63.1 (2012), pp. 287–323. ISSN: 0066-426X. DOI: [10.1146/annurev-physchem-032511-143803](https://doi.org/10.1146/annurev-physchem-032511-143803).
- [30] Matt Challacombe. In: *arXiv* 1001.2586 (2010).
- [31] Sergei Tretiak, Christine M. Isborn, Anders M.N. Niklasson, and Matt Challacombe. In: *Journal of Chemical Physics* 130.5 (2009). ISSN: 00219606. DOI: [10.1063/1.3068658](https://doi.org/10.1063/1.3068658).
- [32] T. J. Zuehlsdorff et al. In: *The Journal of Chemical Physics* 139.6, 064104 (2013). DOI: <http://dx.doi.org/10.1063/1.4817330>.

- [33] Eugene S. Kadantsev, M. J. Stott, and Angel Rubio. In: *The Journal of Chemical Physics* 124.13, 134901 (2006). DOI: <http://dx.doi.org/10.1063/1.2186999>.
- [34] Antonio Prlj et al. In: *Journal of Chemical Theory and Computation* 12.6 (2016), pp. 2652–2660. ISSN: 15499626. DOI: [10.1021/acs.jctc.6b00245](https://doi.org/10.1021/acs.jctc.6b00245).
- [35] Stefan Grimme and Maja Parac. In: *ChemPhysChem* 4.3 (2003), pp. 292–295. ISSN: 1439-7641. DOI: [10.1002/cphc.200390047](https://doi.org/10.1002/cphc.200390047).
- [36] Leeor Kronik, Tamar Stein, Sivan Refaely-Abramson, and Roi Baer. *Excitation gaps of finite-sized systems from optimally tuned range-separated hybrid functionals*. 2012. DOI: [10.1021/ct2009363](https://doi.org/10.1021/ct2009363).
- [37] Natalia Kuritz, Tamar Stein, Roi Baer, and Leeor Kronik. In: *Journal of Chemical Theory and Computation* 7.8 (2011), pp. 2408–2415. ISSN: 1549-9618. DOI: [10.1021/ct2002804](https://doi.org/10.1021/ct2002804).
- [38] Tamar Stein, Leeor Kronik, and Roi Baer. In: *Journal of the American Chemical Society* 131.8 (2009), pp. 2818–2820. ISSN: 0002-7863. DOI: [10.1021/ja8087482](https://doi.org/10.1021/ja8087482).
- [39] Oleg A. Vydrov and Gustavo E. Scuseria. In: *Journal of Chemical Physics* 125.23 (2006). ISSN: 00219606. DOI: [10.1063/1.2409292](https://doi.org/10.1063/1.2409292).
- [40] M J Frisch et al. *Gaussian 09, Revision B.01*. 2009. DOI: [111](https://doi.org/10.1021/111).
- [41] Rick A. Kendall, Thom H. Dunning, and Robert J. Harrison. In: *The Journal of Chemical Physics* 96.9 (1992), pp. 6796–6806. ISSN: 0021-9606. DOI: [10.1063/1.462569](https://doi.org/10.1063/1.462569).
- [42] Tai-Sung Lee, James P. Lewis, and Weitao Yang. In: *Computational Materials Science* 12.3 (1998), pp. 259–277. ISSN: 0927-0256. DOI: [http://dx.doi.org/10.1016/S0927-0256\(98\)00029-9](http://dx.doi.org/10.1016/S0927-0256(98)00029-9).
- [43] Christopher Sutton, Chad Risko, and Jean-Luc Brédas. In: *Chemistry of Materials* 28.1 (2016), pp. 3–16. ISSN: 0897-4756. DOI: [10.1021/acs.chemmater.5b03266](https://doi.org/10.1021/acs.chemmater.5b03266).
- [44] Tao Zeng, Roald Hoffmann, and Nandini Ananth. In: *Journal of the American Chemical Society* 136.15 (2014), pp. 5755–5764. ISSN: 0002-7863. DOI: [10.1021/ja500887a](https://doi.org/10.1021/ja500887a).
- [45] B. Hajgató et al. In: *Journal of Chemical Physics* 131.22 (2009). ISSN: 00219606. DOI: [10.1063/1.3270190](https://doi.org/10.1063/1.3270190).
- [46] Johannes Hachmann, Jonathan J. Dorando, Michael Avilés, and Garnet Kin Lic Chan. In: *Journal of Chemical Physics* 127.13 (2007). ISSN: 00219606. DOI: [10.1063/1.2768362](https://doi.org/10.1063/1.2768362).

- [47] Michael J. G. Peach and David J. Tozer. In: *The Journal of Physical Chemistry A* 116.39 (2012), pp. 9783–9789. ISSN: 1089-5639. DOI: [10.1021/jp308662x](https://doi.org/10.1021/jp308662x).
- [48] Yi-Lei Wang and Guo-Shi Wu. In: *International Journal of Quantum Chemistry* 108.3 (2008), pp. 430–439. ISSN: 00207608. DOI: [10.1002/qua.21510](https://doi.org/10.1002/qua.21510).
- [49] Tonatiuh Rangel, Samia M Hamed, Fabien Bruneval, and Jeffrey B Neaton. In: *The Journal of Chemical Physics* 146.19 (2017), p. 194108. ISSN: 0021-9606. DOI: [10.1063/1.4983126](https://doi.org/10.1063/1.4983126).

## 6 Freeze-and-thaw embedding in ONETEP

In Chapter 5 we discussed the impact of a *p*-terphenyl environment on the excited states of pentacene using ground state DFT and TDDFT as implemented in Gaussian and ONETEP. We saw that in order to accurately capture the vacuum and solvent excited state energies with DFT, it is important to make use of optimally-tuned range-separated hybrid (OT-RSH) functionals as semi-local functionals can fail drastically to describe the absolute energies and the solvatochromic shifts for acenes in explicit hosts. Implicit solvent models do not capture the full picture of the crystal host effects on the excitonic properties of the acene guest molecules. However, the use of hybrid functionals on molecules in explicit host is limited by the computational cost of performing the exact exchange calculation, which can become prohibitive. Here we can see the potential for the embedding methods discussed in Chapter 3 which raise the possibility of performing a high-level calculation (e.g. with hybrids) on the subsystem of interest while treating the environment at a lower level of theory (e.g. with semi-local functionals). Combining embedding approaches with ONETEP’s linear-scaling infrastructure requires a considerable adaptation to the ONETEP algorithm, which conventionally involves a full optimisation of all non-orthogonalised generalised Wannier functions (NGWFs) and density kernel elements as part of the self-consistent optimisation procedure discussed in Chapter 4. In this chapter we will outline a toy model for performing freeze-and-thaw (F+T) and frozen density embedding (FDE) calculations with ONETEP, plus the potential ramifications for the implementation of embedding.

### 6.1 Embedding and linear-scaling DFT

We saw in Chapter 3 that quantum embedding can be represented as a partitioning of the electron density for various subsystems which, via the Wesolowski and Warshel frozen density embedding (FDE) method, provides a complete description of the quantum system [1, 2]. From the discussion of density-matrix theory in Chapter 4, we have also seen that the density matrix approach to electronic structure provides a complete description of the system since all information can be recovered from the density matrix. The inherent nearsightedness of the density matrix [3, 4] facilitates the use of



linear-scaling methods as it implies the localised nature of electronic structure. This locality indicates that, from a physical standpoint, linear-scaling methods can be straightforwardly adapted for embedding [5]. For example, the divide and conquer technique involves calculating segments of the density matrix for separate parts of the system separately, which are then pieced together to form the full density matrix [6]. This bears a distinct similarity to the FDE self-consistent method for obtaining the ground state density and highlights the similarity between the two approaches. In this section we will explore in more detail the potential for combining embedding with ONETEP.

### 6.1.1 Embedding in ONETEP

We begin by outlining the ONETEP formalism outlined in Chapter 4 with the molecular system partitioned into two regions, an active subsystem A and an environment B. The partitioning of the full system is done on an atom-by-atom basis, such that the non-orthogonal generalised Wannier functions (NGWFs) related to a particular atom are all assigned to one subsystem. The active and environment NGWFs are denoted  $\{\phi^A\}$  and  $\{\phi^B\}$ . Consequently, the elements of the density kernel  $\mathbf{K}$  are also divided into diagonal region blocks and off-diagonal, cross-overlap blocks,

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{AA} & \mathbf{K}_{AB} \\ \mathbf{K}_{BA} & \mathbf{K}_{BB} \end{pmatrix}. \quad (6.1)$$

The density matrix can then be reconstructed in terms of the subsystem NGWFs and kernels,

$$\rho_{IJ}(\mathbf{r}, \mathbf{r}') = \sum_{\alpha \in I} \sum_{\beta \in J} \phi_{\alpha}^I(\mathbf{r}) K_{IJ}^{\alpha\beta} \phi_{\beta}^J(\mathbf{r}'), \quad I, J \in \{A, B\}, \quad (6.2)$$

from which one can obtain the subsystem and total electronic densities,

$$n_{\text{tot}}(\mathbf{r}) = \sum_{I,J} n_{IJ}(\mathbf{r}) = \sum_{I,J} \left( \sum_{\alpha \in I} \sum_{\beta \in J} K_{IJ}^{\alpha\beta} [\phi_{\beta}^J(\mathbf{r}) \phi_{\alpha}^I(\mathbf{r})] \right). \quad (6.3)$$

The overlap and Hamiltonian matrices for the complete system are then defined as

$$\mathbf{S} = \begin{pmatrix} \mathbf{S}_{AA} & \mathbf{S}_{AB} \\ \mathbf{S}_{BA} & \mathbf{S}_{BB} \end{pmatrix}, \quad \mathbf{H} = \begin{pmatrix} \mathbf{H}_{AA} & \mathbf{H}_{AB} \\ \mathbf{H}_{BA} & \mathbf{H}_{BB} \end{pmatrix}, \quad (6.4)$$

where

$$(\mathbf{S}_{IJ})_{\alpha\beta} = \langle \phi_{\alpha}^I | \phi_{\beta}^J \rangle, \quad (6.5)$$

$$(\mathbf{H}_{IJ})_{\alpha\beta} = \langle \phi_{\alpha}^I | \hat{H} | \phi_{\beta}^J \rangle. \quad (6.6)$$

At this stage we assume that there is a single Hamiltonian operator  $\hat{H}$  that is applicable to all subsystems, composed of the kinetic energy  $\hat{T}$ , the Hartree potential  $\hat{V}_H$ , the external potential  $\hat{V}_{\text{ext}}$  and the exchange-correlation potential  $\hat{V}_{\text{XC}}$ ,

$$\hat{H} = \hat{T} + \hat{V}_H + \hat{V}_{\text{ext}} + \hat{V}_{\text{XC}}. \quad (6.7)$$

The usual conditions of normalisation and idempotency apply for the full matrices rather than the subsystem structures,

$$\text{Tr} [\mathbf{KS}] = N_A + N_B, \quad (6.8)$$

$$\mathbf{KSK} = \mathbf{K}, \quad (6.9)$$

allowing for the transfer of density between the subsystems. In keeping with the standard NGWF optimisation procedure in ONETEP, we then seek to minimise the quantity

$$\Omega = \text{Tr} [\mathbf{KH}], \quad (6.10)$$

subject to the usual constraints (see Chapter 4).

At this stage it is worth noting that, aside from a change of notation, we have not actually altered anything about the ONETEP algorithm. An optimisation of the elements of the full density kernel  $\mathbf{K}$  and the NGWFs  $\{\phi^A\}$  and  $\{\phi^B\}$  would yield the exact same results as the standard algorithm laid out in Chapter 4. Nonetheless, the format presented above highlights the two potential approaches for combining embedding techniques with the ONETEP method. Firstly, partitioning the NGWFs and density kernel elements into subsystem structures enables the possibility of selectively optimising the terms that are explicitly associated with a particular region of interest. Secondly, the Hamiltonian submatrices (6.4) can be generalised to use different Hamiltonian operators in each subsystem, enabling the use of different levels of theory within a single linear-scaling DFT calculation. We will return to this latter approach in Chapter 7, in this chapter we will solely focus on the subsystem optimisation concept using a single Hamiltonian.

The selective optimisation technique bears a distinct similarity to the frozen density embedding (FDE) method of Wesolowski and Warshel [1], whereby a subsystem electron density is optimised self-consistently in the presence of a frozen environment. However, as previously discussed in Chapter 3.5, the Wesolowski-Warshel FDE method relies on a direct partitioning of the Kohn-Sham orbitals between the various subsystems. The electron densities are then constructed for the subsystems, with normalisation and orthogonality of the Kohn-Sham orbitals enforced for each subsystem separately. The loss of explicit orthogonality between the subsystem Kohn-Sham

orbitals then gives rise to the non-additive kinetic potential (3.27) [7, 8], the primary computational difficulty associated with such embedding methods. By contrast, in the ONETEP embedding approach we perform the partitioning on an atom-by-atom basis, assigning the NGWFs on each atom to the specified region. The density kernel is then divided into parts related to the NGWFs in subsystems A and B only, plus the cross-overlap between the subsystems. Normalisation and idempotency are thus enforced for the full system in ONETEP embedding, rather than the subsystem structures in the standard Wesolowski-Warshel embedding method, and is in fact closer in nature to the Bowler-Gillan view of embedding [5]. This averts the need to construct the non-additive kinetic potential, since idempotency of the full density matrix is equivalent to orthonormality of all the Kohn-Sham orbitals. Optimising the NGWFs and density kernel elements selectively while imposing the full system constraints should thus be possible in ONETEP without the need to construct elaborate kinetic potentials, or enforce explicit orthogonality via the projector methods set out in Section 3.4.

For the remainder of this chapter, we will focus our efforts on the selective optimisation of the minimal NGWF basis set utilised in ONETEP for linear-scaling DFT calculations. In general, during a self-consistent optimisation in ONETEP all NGWFs are periodically updated along with the density kernel elements until the convergence criteria are met. One aspect that is not well understood from an embedding perspective, however, is the validity of this approach when only a subset of the orbitals are ever updated, with the remainder fixed either as pseudoatomic orbitals (PAOs) or another type of basis function. In the next section we will examine the theory behind this by examining the NGWF gradient for subsystem optimisation.

### 6.1.2 NGWF gradient for subsystem optimisation

Starting from the Kohn-Sham band-structure energy (see 4.2.2),

$$\Omega = \text{Tr} [\mathbf{KH}], \quad (6.11)$$

our task is to calculate the NGWF gradient by optimising only a subset of the NGWFs available, namely  $\{\phi_i^A\}$ , but using the full density kernel and Hamiltonian matrices. The matrix product can be written as

$$\mathbf{KH} = \begin{pmatrix} \mathbf{K}_{AA}\mathbf{H}_{AA} + \mathbf{K}_{AB}\mathbf{H}_{BA} & \mathbf{K}_{AA}\mathbf{H}_{AB} + \mathbf{K}_{AB}\mathbf{H}_{BB} \\ \mathbf{K}_{BB}\mathbf{H}_{AA} + \mathbf{K}_{BB}\mathbf{H}_{BA} & \mathbf{K}_{BA}\mathbf{H}_{AB} + \mathbf{K}_{BB}\mathbf{H}_{BB} \end{pmatrix} \quad (6.12)$$

which gives,

$$\Omega = \text{Tr} [\mathbf{K}_{AA} \mathbf{H}_{AA} + \mathbf{K}_{AB} \mathbf{H}_{BA}] + \text{Tr} [\mathbf{K}_{BA} \mathbf{H}_{AB} + \mathbf{K}_{BB} \mathbf{H}_{BB}] \quad (6.13)$$

$$= \Omega_A + \Omega_B + \text{Tr} [\mathbf{K}_{AB} \mathbf{H}_{BA} + \mathbf{K}_{BA} \mathbf{H}_{AB}]. \quad (6.14)$$

We thus see that the total energy functional can be split into the contributions from each subsystem plus a non-additive contribution. However, we must also impose the idempotency of the density kernel via the LNV method (Section 4.1.2),

$$\mathbf{K} = 3\mathbf{L}\mathbf{S}\mathbf{L} - 2\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{S}\mathbf{L}, \quad (6.15)$$

where  $\mathbf{L}$  is an auxiliary density kernel, which results in a ‘convolution’ of the subsystem terms in (6.14).

For the moment we consider only the structure of the matrices in (6.4) i.e. we effectively treat these as  $2 \times 2$  matrices and ignore their tensorial properties. This will clarify which matrix blocks must be considered without having to write the individual orbitals explicitly.

$$\begin{aligned} (\mathbf{K}\mathbf{H})_{\alpha\beta} &= \mathbf{K}_{\alpha I} \mathbf{H}_{I\beta}, \\ &= (3\mathbf{L}_{\alpha J} \mathbf{S}_{JK} \mathbf{L}_{KI} - 2\mathbf{L}_{\alpha J} \mathbf{S}_{JK} \mathbf{L}_{KM} \mathbf{S}_{MN} \mathbf{L}_{NI}) \mathbf{H}_{I\beta}, \\ &= 3\mathbf{L}_{\alpha J} \mathbf{S}_{JK} \mathbf{L}_{KI} \mathbf{H}_{I\beta} - 2\mathbf{L}_{\alpha J} \mathbf{S}_{JK} \mathbf{L}_{KM} \mathbf{S}_{MN} \mathbf{L}_{NI} \mathbf{H}_{I\beta}, \end{aligned}$$

where we make use of the Einstein summation convention. Thus

$$\Omega = \text{Tr}[\mathbf{K}\mathbf{H}] = 3\mathbf{L}_{IJ} \mathbf{S}_{JK} \mathbf{L}_{KM} \mathbf{H}_{MI} - 2\mathbf{L}_{IJ} \mathbf{S}_{JK} \mathbf{L}_{KM} \mathbf{S}_{MN} \mathbf{L}_{NP} \mathbf{H}_{PI}. \quad (6.16)$$

Even with our  $2 \times 2$  picture, there are 80 terms added together in this trace. Comparing with (6.14), two terms are  $\Omega_A$ , another two are  $\Omega_B$  and the other 76 are the trace  $\text{Tr} [\mathbf{K}_{AB} \mathbf{H}_{BA} + \mathbf{K}_{BA} \mathbf{H}_{AB}]$ ! This can be rewritten as

$$\begin{aligned} \Omega &= 3(\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{H})_{AA} - 2(\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{H})_{AA} \\ &\quad + 3(\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{H})_{BB} - 2(\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{S}\mathbf{L}\mathbf{H})_{BB}. \end{aligned} \quad (6.17)$$

Taking the derivative of our total energy functional with respect to an NGWF in subsystem A, we obtain the contravariant gradient,

$$\begin{aligned}
 |g_A^\alpha\rangle &= \frac{\partial \Omega}{\partial \langle \phi_\alpha^A |} \\
 &= |\phi_i^A\rangle (3\text{LHL} - 2\text{LSLHL} - 2\text{LHLSL})_{AA}^{i\alpha} + \hat{H} |\phi_i^A\rangle (3\text{LSL} - 2\text{LSLSL})_{AA}^{i\alpha} \\
 &\quad + |\phi_j^B\rangle (3\text{LHL} - 2\text{LSLHL} - 2\text{LHLSL})_{BA}^{j\alpha} + \hat{H} |\phi_j^B\rangle (3\text{LSL} - 2\text{LSLSL})_{BA}^{j\alpha}, \\
 &= |\phi_i^A\rangle \mathcal{A}^{i\alpha} + \hat{H} |\phi_i^A\rangle \mathcal{B}^{i\alpha} + |\phi_j^B\rangle \mathcal{C}^{j\alpha} + \hat{H} |\phi_j^B\rangle \mathcal{D}^{j\alpha},
 \end{aligned} \tag{6.18}$$

where

$$\mathcal{A} = [3\text{LHL} - 2\text{LHLSL} - 2\text{LSLHL}]_{AA} = \mathbf{Q}_{AA}, \tag{6.19}$$

$$\mathcal{B} = [3\text{LSL} - 2\text{LSLSL}]_{AA} = \mathbf{K}_{AA}, \tag{6.20}$$

$$\mathcal{C} = [3\text{LHL} - 2\text{LHLSL} - 2\text{LSLHL}]_{BA} = \mathbf{Q}_{BA}, \tag{6.21}$$

$$\mathcal{D} = [3\text{LSL} - 2\text{LSLSL}]_{BA} = \mathbf{K}_{BA}. \tag{6.22}$$

As usual, we must lower the index of the NGWF search direction vector such that the resulting term is covariant:

$$|g_{\alpha,A}\rangle = |\phi_i^A\rangle \tilde{\mathcal{A}}_{,\alpha}^i + \hat{H} |\phi_i^A\rangle \tilde{\mathcal{B}}_{,\alpha}^i + |\phi_j^B\rangle \tilde{\mathcal{C}}_{,\alpha}^j + \hat{H} |\phi_j^B\rangle \tilde{\mathcal{D}}_{,\alpha}^j, \tag{6.23}$$

where

$$\tilde{\mathcal{A}} = \mathbf{Q}_{AA}\mathbf{S}_{AA} + \mathbf{Q}_{AB}\mathbf{S}_{BA}, \tag{6.24}$$

$$\tilde{\mathcal{B}} = \mathbf{K}_{AA}\mathbf{S}_{AA} + \mathbf{K}_{AB}\mathbf{S}_{BA}, \tag{6.25}$$

$$\tilde{\mathcal{C}} = \mathbf{Q}_{BA}\mathbf{S}_{AA} + \mathbf{Q}_{BB}\mathbf{S}_{BA}, \tag{6.26}$$

$$\tilde{\mathcal{D}} = \mathbf{K}_{BA}\mathbf{S}_{AA} + \mathbf{K}_{BB}\mathbf{S}_{BA}. \tag{6.27}$$

This result is simply the gradient for a single NGWF in the presence of all the other orbitals in a standard ONETEP calculation. We can thus construct a frozen-density embedding (FDE) style optimisation algorithm by enforcing a zero NGWF gradient for the (frozen) B orbitals while using the standard NGWF gradient for the A orbitals,

$$|g_{\alpha,A}\rangle = |\phi_i^A\rangle \tilde{\mathcal{A}}_{,\alpha}^i + \hat{H} |\phi_i^A\rangle \tilde{\mathcal{B}}_{,\alpha}^i + |\phi_j^B\rangle \tilde{\mathcal{C}}_{,\alpha}^j + \hat{H} |\phi_j^B\rangle \tilde{\mathcal{D}}_{,\alpha}^j, \tag{6.28}$$

$$|g_{\alpha,B}\rangle = |\mathbf{0}\rangle. \tag{6.29}$$

Alternatively, an iterative freeze-thaw (F+T) NGWF optimisation cycle can be devised by swapping the orbitals to be optimised at any given step i.e. improving the A orbitals for one step in the presence of the frozen B orbitals, before freezing  $\{\phi_i^A\}$  and ‘thawing’

$\{\phi_i^B\}$  for one step and so on until self-consistency is obtained. For initial tests, the zero environment gradient was implemented by explicitly zeroing the components of the NGWF gradient vector that correspond to the B subsystem orbitals after the full gradient was calculated. Later this feature was integrated into ONETEP, alongside the work explored in Chapter 7, by only calculating the relevant coefficient matrices for thawed NGWFs and skipping the environment parts — both approaches yielded equivalent results.

## 6.2 Results — iterative subsystem NGWF optimisation

Before we can examine the viability of FDE as an optimisation strategy for NGWFs, we must check that an iterative F+T NGWF optimisation routine is capable of reproducing the results that one would expect for a regular total energy optimisation in ONETEP. For this purpose, we will consider the study of a water dimer in the configuration shown in Figure 6.1a. The molecules are bound by a weak hydrogen bond ( $O \cdots H$ ) between the oxygen of one molecule (acceptor) and an adjacent hydrogen of the other molecule (donor). The water dimer is a system that has been previously studied to determine the validity of NGWFs in accurately reproducing ground state properties to plane-wave accuracy and removing the spurious basis set superposition error (BSSE) associated with fixed localised basis functions [9]. This structure allows for a straightforward partitioning of the NGWF regions between the two weakly-bound molecules while still enabling the different components of the hydrogen bond to be treated at different levels of theory via the iterative freezing and thawing of the NGWFs. For all calculations in this section we will use the PBE exchange-correlation functional [10], a generalised gradient approximation (GGA, see Chapter 2.4) to the exchange-correlation problem. The standard norm-conserving pseudopotentials (see Chapter 3.1) for oxygen and hydrogen distributed with ONETEP were used, along with a kinetic energy cutoff of 1000 eV and an NGWF radius of 7.0 bohr for all atoms. All density kernel elements are updated as per usual for a ONETEP calculation.

Figure 6.1 shows the behaviour of the NGWF gradient for the water dimer by performing an iterative optimisation of the NGWFs for the acceptor and donor molecules compared with a full NGWF optimisation of the entire system. We see that the NGWF gradient gradually decreases throughout the optimisation, eventually dropping below the specified threshold after 36 steps, albeit somewhat slower than the full system time of 10 steps. Given that only half of the NGWFs are optimised at any stage in the F+T cycle, the time required to perform this calculation is equivalent to performing 18 iterations for the whole system. This confirms that the F+T approach can successfully

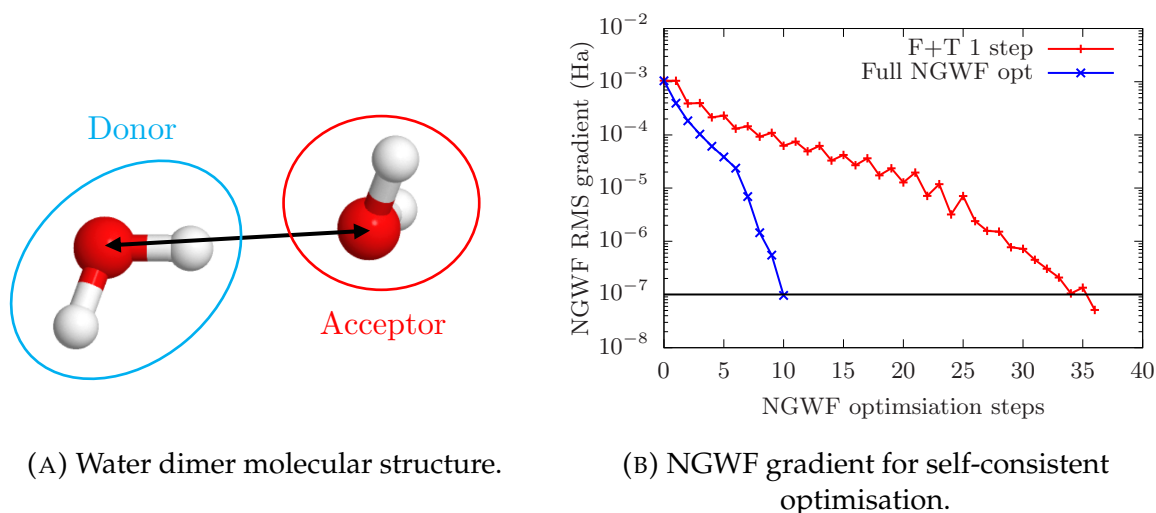


FIGURE 6.1: Left: Water dimer molecule used for embedding tests. The distance between the oxygen atoms is used as a measure of the separation between the molecules. Figure generated with Jmol [11]. Right: NGWF gradient for water dimer for optimisation of donor and acceptor molecules in freeze-and-thaw (F+T) cycles at a separation of  $2.95 \text{ \AA}$ . The behaviour of the gradient for a regular optimisation of all NGWFs is also shown for comparison. Note that for F+T each step only optimises half of the NGWFs in the system, and the gradient plotted at each data point is the gradient for the NGWFs optimised at that step only.

optimise the NGWFs and density kernel for an interacting molecular system, thus illustrating that the disparity between the NGWFs caused by the freeze-thaw algorithm is no barrier to the use of NGWFs in total energy calculations. This is a fairly simplistic test, since at any given stage the discrepancy in the NGWFs between the acceptor and donor molecules remains small since once one molecule's NGWFs are updated, the other's are updated immediately afterwards, with only an intermediate kernel recalculation distinguishing this from a standard ONETEP optimisation process.

To show that this is a general result, Figure 6.2 displays the error in total energy between the full NGWF total energy and a series of F+T calculations for the same system with up to 10 successive steps in the NGWF optimisation for each region. For example, with two steps the acceptor NGWFs are updated for two successive iterations, before switching back to the donor for the next two, then the acceptor and so on until self-consistency is obtained. We see that the error in the total energy relative to full NGWF optimisation never exceeds 1.5 meV, or  $\sim 0.0002\%$  of the Kohn-Sham energy for this system. All F+T arrangements do eventually converge, taking as little as 35 NGWF optimisation steps with four F+T cycles per region and a maximum of 59 NGWF optimisation steps for nine cycles per region. While the error is negligible for all F+T calculations, we notice that the error drops to a minimum for 3–4 steps,

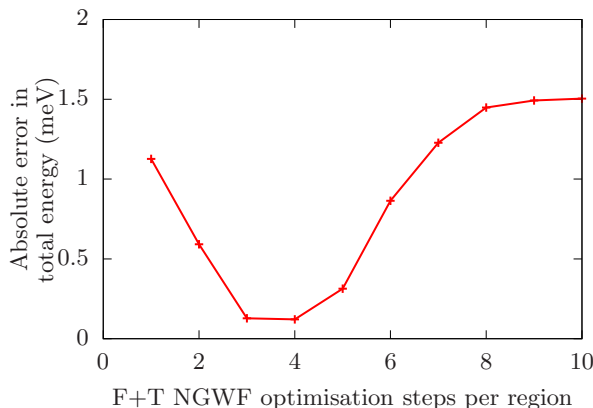


FIGURE 6.2: Absolute difference in the final energies obtained from iterative optimisation of NGWFs in freeze-thaw (F+T) cycles compared to full optimisation of all NGWFs simultaneously for a water dimer. Each F+T calculation was performed by updating the NGWFs for a fixed number of steps in each region in turn. The acceptor and donor molecules were treated as separate subsystems. Calculations were terminated once the NGWF RMS gradient dropped below a preset threshold of  $10^{-7}$  Ha.

before gradually increasing for greater numbers of steps. This suggests that it is in fact optimal to allow some degree of flexibility in the NGWFs by optimising one region in the presence of frozen NGWFs for a few optimisation steps, though too many F+T cycles can reduce accuracy due to the growing disparity between NGWFs in different regions.

Overall, these results confirm the validity of the F+T approach and provides some confidence that a similar method can be deployed to the optimisation of only a subset of NGWFs in the presence of a frozen environment, which we shall consider in the next section.

### 6.2.1 Frozen density embedding (FDE)

Having established that it is possible to obtain the ground state density via an iterative NGWF optimisation, we now consider the possibility of optimising a subset of the NGWFs in our system in a quasi-FDE scheme. We focus our attention on the interaction potential for the water dimer structure, using the same calculation parameters as set out in Section 6.2. Even without using the selective NGWF optimisation scheme outlined in Section 6.1.2, we can make direct comparisons of the quality of different types of basis functions that can be used as starting points for further NGWF improvements. Figure 6.3 shows the interaction potentials as a function of the separation of the oxygen atoms in the dimer with three different sets of basis functions. The first are pseudoatomic orbitals (PAOs) for which only the density kernel is optimised. The



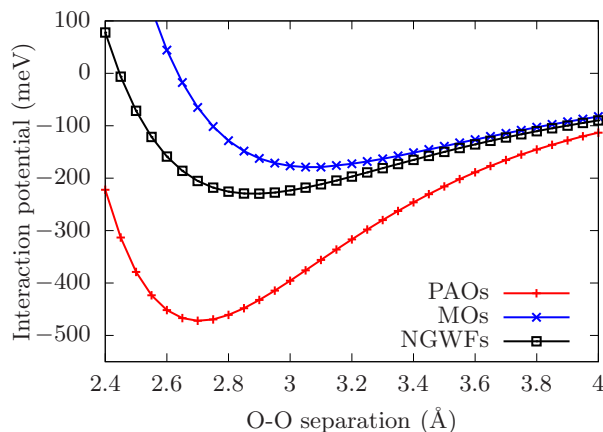


FIGURE 6.3: Interaction potential for water dimer using pseudoatomic orbitals (PAOs), molecular orbitals (MOs) and non-orthogonal generalised Wannier functions (NGWFs). All density kernel elements are optimised self-consistently at all stages of the calculations. No counterpoise correction is applied to account for basis set superposition error (BSSE) with PAOs or MOs.

second are molecular orbitals (MOs) determined from an NGWF optimisation of each water molecule in isolation, for which again there is no subsequent optimisation of the basis and only the kernel is updated. Finally we show the results of a full NGWF optimisation of all basis functions. We see that the PAOs vastly overbind the dimer relative to the NGWFs, as well as yielding a much shorter optimal molecular separation. The MO basis by comparison yields a bond length that is too long with a shallow energy minimum. Both the MO and PAO results suffer from BSSE and can be improved by utilising a counterpoise correction for the dimer calculation, however plane-wave accuracy is only recovered with the BSSE-free NGWF basis [9].

Our FDE model enables us to perform selective NGWF optimisation for our system, with which we can evaluate the accuracy of a partially optimised basis set. Figure 6.4a shows the previous NGWF and PAO curves along with FDE calculations for the donor and acceptor molecule. All orbitals are initialised with the same PAOs as were used to perform the density kernel-only optimisation calculation. Then only the NGWFs of one water molecule are updated during the calculation, the other's being fixed at the PAO level. In each case the named molecule is the one whose NGWFs are optimised. We see that only updating the donor molecule orbitals produces a binding energy that is even more inaccurate than the pure PAO calculation. While there is a lengthening of the optimal O-O separation to 2.8, this is still underestimated compared to the NGWF results. By contrast, an FDE optimisation of the acceptor water yields results that are more in alignment with the NGWF basis, with a significant reduction in the overbinding effect and a bond length that matches the expected NGWF result. Though the FDE

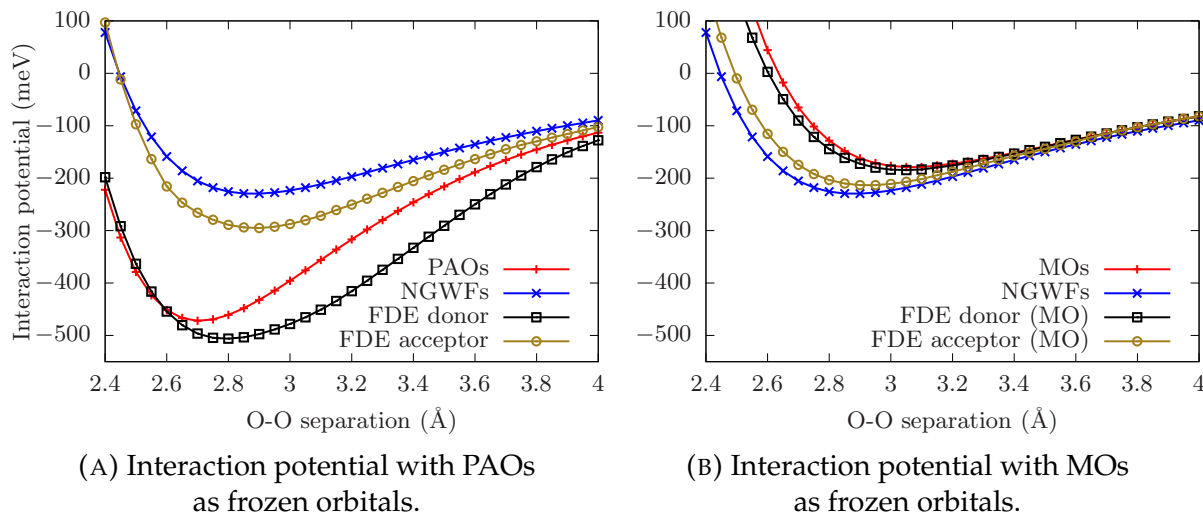


FIGURE 6.4: Interaction potential for water dimer via selective optimisation of subsystem NGWFs, using either pseudoatomic orbitals (PAOs, left) or molecular orbitals (MOs, right) as the starting point for the dimer calculations. The donor and acceptor molecules were treated as separate subsystems.

optimisation in this case does not correctly reproduce the NGWF interaction potential, it does highlight the fact that a selective orbital optimisation can significantly improve results compared to a PAO basis, in this case by updating the acceptor rather than the donor orbitals.

Next, we consider the impact of using the MO basis as the starting point for an FDE optimisation. Figure 6.4b shows the interaction potentials for FDE optimisation of the donor and acceptor molecules compared to the MO functions discussed earlier. Notice that a donor optimisation yields little improvement compared to the pure MO calculation. The acceptor curve, however, shows significant improvement, with an energy minimum that is only 20 meV higher than the full NGWF optimisation. We see a divergence between the two curves as the separation is reduced, possibly due to the increasing importance of the unoptimised orbitals at short range due to greater overlap.

In each of the FDE graphs, optimisation of the acceptor orbitals yields significantly better results than the donor. In our minimal basis set, each hydrogen atom has only one orbital, while each oxygen atom possesses four, such that improving the oxygen orbitals allows for a greater degree of flexibility in the density matrix. It is clear, then, from the above results that an accurate description of the orbitals constituting the  $\text{O} \cdots \text{H}$  bond is essential to obtain reliable results for the interaction potential. Figure 6.5 shows a comparison of the interaction potentials obtained by optimising only the orbitals in the  $\text{O} \cdots \text{H}$  bond while holding the rest fixed and vice versa. The result is

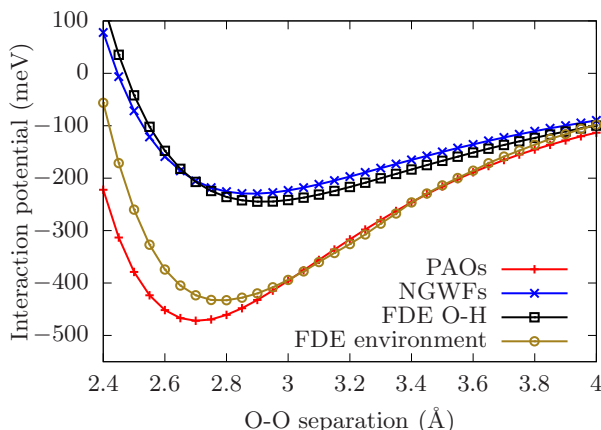


FIGURE 6.5: Interaction potential for water dimer via selective optimisation of subsystem NGWFs. Here the  $\text{O} \cdots \text{H}$  hydrogen bond between the molecules is treated as a separate subsystem from the environment consisting of all other atoms in the system. Pseudoatomic orbitals (PAOs) obtained from monomer calculations were used as the starting point for the dimer runs.

a significantly better fit with the  $\text{O} \cdots \text{H}$  optimisation than for any other schemes presented so far. The energy minimum is only 14 meV higher than that obtained from a full NGWF optimisation of the dimer, while the minimum position is correctly reproduced at 2.90 Å. Comparison can be made to the optimisation of the environment orbitals only, such that the  $\text{O} \cdots \text{H}$  bond is described using only PAOs. There is scarcely any improvement in the binding energy which remains vastly overestimated as for the case of using PAOs for the whole system, however there is a noticeable improvement in the optimal bond distance, which much more closely matches the all-NGWF result. It would appear that, while all orbitals are significant for describing the full effects of the water dimer interaction curve, ensuring that the  $\text{O} \cdots \text{H}$  bond is part of the active system and that both of these atoms are treated at the same level of theory is important for an accurate FDE description of this system.

In summary, an NGWF-in-PAO FDE optimisation of the  $\text{O} \cdots \text{H}$  bond of a water dimer produces results that are in close agreement with an all-NGWF optimisation of the system, an encouraging result for the correct description of structures with subsystem orbital optimisation.

## 6.2.2 Excitation energies with FDE

We now consider the application of the subsystem NGWF method to the study of excited state properties. For this purpose we will focus on the familiar pentacene in *p*-terphenyl cluster that we used in Chapter 5. As we saw in our study of the lower excited states of this system, the excitation formed by the transition from the singlet

Basis functions	$\Delta E_{S_0 \rightarrow T_1}$ (eV)
PAOs	0.913
NGWFs	0.885

TABLE 6.1: Vacuum energy differences between the  $S_0$  singlet ground state and the  $T_1$  triplet ground state for pentacene using non-orthogonal generalised Wannier functions (NGWFs) and pseudoatomic orbitals (PAOs). All calculations were performed using the PBE exchange-correlation functional.

ground state  $S_0$  to the triplet ground state  $T_1$  is heavily localised on the pentacene molecule, with a solvatochromic shift of only a few meV resulting from placing the pentacene within an implicit or explicit  $p$ -terphenyl host. Thus this is a system that would appear to be well suited to treating the pentacene at a higher level of theory than the neighbouring molecules. In addition, using the  $\Delta$ SCF method it is possible to obtain the excitation energy  $\Delta E_{S_0 \rightarrow T_1}$  by calculating the difference between two ground state DFT calculations, one in the singlet spin-unpolarised configuration for  $S_0$  and the other in the spin-polarised triplet arrangement for  $T_1$ . This can be done in both vacuum and explicit solvent, the latter we represent using the six  $p$ -terphenyl cluster from Chapter 5.

Vacuum calculations were performed using both a set of unoptimised pseudoatomic orbitals (PAOs) and self-consistently updated NGWFs. For cluster calculations, a mixture of optimised and fixed basis functions were used. Molecular orbitals (MOs) were generated for the vacuum pentacene and the six neighbouring  $p$ -terphenyls separately, which were then used for kernel-only optimisation and as starting points for FDE optimisation of the pentacene orbitals. Thus, five different combinations of basis sets were used in total energy optimisation of the pentacene in  $p$ -terphenyl cluster — PAOs, MOs, NGWFs, NGWFs-in-PAOs and NGWFs-in-MOs. The first two involve kernel-only optimisation with fixed basis functions, the third a standard kernel and NGWF optimisation process, and the final two are selective optimisation of the orbitals on the pentacene only in the presence of either a PAO or MO  $p$ -terphenyl environment. In all cases, both singlet and triplet calculations were performed. All results presented here use the PBE exchange-correlation functional with a kinetic energy cutoff of 1000 eV and a basis function radius of 10.0 bohr.

Table 6.1 displays the vacuum excitation energy  $\Delta E_{S_0 \rightarrow T_1}$  for pentacene using both PAOs and NGWFs as benchmark results. We see that the PAO basis overestimates the energy of the transition by 28 meV compared to NGWFs. This is not a surprising result as the minimal PAO basis is by no means an optimal basis choice — normally one would perform a kernel-only optimisation with a larger number of basis functions, but for illustration purposes the minimal basis highlights the difference in accuracy

Basis functions	$\Delta E_{S_0 \rightarrow T_1}$ (eV)
PAOs	0.928
MOs	0.881
NGWFs	0.879
FDE PAOs	0.101
FDE MOs	0.882

TABLE 6.2: Energy differences between the  $S_0$  singlet ground state and the  $T_1$  triplet ground state for the pentacene in  $p$ -terphenyl cluster host for various basis functions. All calculations were performed using the PBE exchange-correlation functional. Frozen density embedding (FDE) calculations were performed via selective optimisation of orbitals associated with the pentacene molecule only — ‘FDE PAOs’ uses pseudoatomic orbitals as the initial basis functions, ‘FDE MOs’ uses molecular orbitals (see text for more details). All calculations were performed using  $\Delta$ SCF.

between NGWFs and PAOs.

Moving on to the cluster calculations, Table 6.2 shows the results for  $\Delta E_{S_0 \rightarrow T_1}$  for each of the five basis types mentioned above. The stand-out result is the dreadful performance of the FDE PAO basis, whereby only the pentacene molecule orbitals are reoptimised relative to the PAO starting point. Most likely this reflects a breaking of the variational principle for the spin-polarised run, since the PAO basis for the  $p$ -terphenyl basis is not optimal and thus there remains room for variation in the density matrix that is otherwise not possible with PAOs. The MO basis gives good results without orbital reoptimisation, predicting an excitation energy only 2 meV above the NGWF result. Allowing the pentacene orbitals to improve in the presence of  $p$ -terphenyl MOs yields a result of 0.882 meV, just 3 meV higher than the optimal NGWF number. From the perspective of solvatochromic shifts, there is no practical difference between the NGWF, MO or FDE-MO results, which all predict a negligible redshift of under 6 meV compared to vacuum. By contrast, the PAO-only basis predicts a blueshift of 15 meV, thus illustrating that the minimal PAO basis gives qualitatively poor results for understanding the nature of this excitation. Combined with the failure of the NGWF-in-PAO calculation to reproduce the  $\Delta E_{S_0 \rightarrow T_1}$  excitation energy, it appears that the use of MOs for pentacene and its environment is essential to obtain accurate results for this system.

### 6.3 Summary

In this chapter we have outlined a preliminary implementation of subsystem NGWF embedding in ONETEP. This has enabled the use of freeze-and-thaw style (F+T)

NGWF optimisation as an alternative optimisation technique, which we have demonstrated can reproduce the ground state total energy for a water dimer. Subsequently, the use of frozen density embedding (FDE) NGWF optimisation, whereby only a subset of orbitals are improved during the energy minimisation routine, was demonstrated for the interaction potential of a water dimer and the  $T_1$  excited states of pentacene in *p*-terphenyl. The former highlights the importance of carefully choosing the orbitals for the optimisation is to be performed, as we saw that the choice of the O...H bond in a water dimer yielded the most accurate results compared to a full NGWF optimisation. The latter reveals the importance of the choice of basis functions for the environment orbitals, even in a system with a straightforward partitioning such as the pentacene in *p*-terphenyl cluster.

While the use of FDE NGWF optimisation does present theoretical benefits from a computational perspective, two caveats must be noted. Firstly, all the tests in this chapter involved full optimisation of all density kernel matrix elements, the calculation of which is the main bottleneck during the self-consistent optimisation procedure, as opposed to improving the NGWFs. Secondly, the approach discussed here is only a toy model which has not been designed to truly take advantage of the potential benefits of selective NGWF optimisation — the only difference compared to a standard ONETEP optimisation being that the covariant NGWF gradient vector is zero for the frozen orbitals. Nevertheless, the tests presented here provide a useful foundation for the use of embedding techniques in ONETEP as we have seen that selective NGWF optimisation and the use of different levels of theory in subsystems is a feasible strategy within the ONETEP linear-scaling architecture.

## References

- [1] Tomasz Adam Wesolowski and Arieh Warshel. In: *The Journal of Physical Chemistry* 97.30 (1993), pp. 8050–8053. DOI: [10.1021/j100132a040](https://doi.org/10.1021/j100132a040).
- [2] Andre Severo Pereira Gomes and Christoph R. Jacob. In: *Annu. Rep. Prog. Chem., Sect. C: Phys. Chem.* 108 (1 2012), pp. 222–277. DOI: [10.1039/C2PC90007F](https://doi.org/10.1039/C2PC90007F).
- [3] W. Kohn. In: *Phys. Rev. Lett.* 76 (17 1996), pp. 3168–3171. DOI: [10.1103/PhysRevLett.76.3168](https://doi.org/10.1103/PhysRevLett.76.3168).
- [4] E. Prodan. In: *Physical Review B - Condensed Matter and Materials Physics* 73.8 (2006), pp. 0–3. ISSN: 10980121. DOI: [10.1103/PhysRevB.73.085108](https://doi.org/10.1103/PhysRevB.73.085108).
- [5] D. R. Bowler and M. J. Gillan. In: *Chemical Physics Letters* 355.3-4 (2002), pp. 306–310. ISSN: 00092614. DOI: [10.1016/S0009-2614\(02\)00273-7](https://doi.org/10.1016/S0009-2614(02)00273-7).
- [6] Weitao Yang. In: *Phys. Rev. Lett.* 66 (11 1991), pp. 1438–1441. DOI: [10.1103/PhysRevLett.66.1438](https://doi.org/10.1103/PhysRevLett.66.1438).
- [7] Tomasz Adam Wesolowski and Arieh Warshel. In: *Journal of Physical Chemistry* 97.30 (1993), pp. 8050–8053. ISSN: 0022-3654. DOI: [10.1021/j100132a040](https://doi.org/10.1021/j100132a040).
- [8] J. D. Goodpaster, N. Ananth, F. R. Manby, and T. F. Miller III. In: *J. Chem Phys.* 133.084103 (2010).
- [9] P. D. Haynes, C. K. Skylaris, A. A. Mostofi, and M. C. Payne. In: *Chemical Physics Letters* 422.4-6 (2006), pp. 345–349. ISSN: 00092614. DOI: [10.1016/j.cplett.2006.02.086](https://doi.org/10.1016/j.cplett.2006.02.086).
- [10] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. In: *Phys. Rev. Lett.* 77 (18 1996), pp. 3865–3868. DOI: [10.1103/PhysRevLett.77.3865](https://doi.org/10.1103/PhysRevLett.77.3865).
- [11] Robert M. Hanson. In: *Journal of Applied Crystallography* 43.5 Part 2 (2010), pp. 1250–1260. DOI: [10.1107/S0021889810030256](https://doi.org/10.1107/S0021889810030256).



## 7 EMFT in ONETEP

In Chapter 5 we studied the excited state properties of pentacene and its nitrogen-doped analogues in both implicit and explicit solvent. We saw that the environment can have significant impact on the excitonic properties of the solute molecule, which manifests itself in the solvatochromic shift of the excited state energy levels relative to vacuum. The inclusion of an explicit description for the host structure can result in qualitatively different behaviour compared to an implicit solvent, plus it can yield more substantial information regarding the nature of the interaction between the solute and solvent molecules, as we saw with the local density of state (LDOS) analysis of the chromophores in explicit host in Figure 5.9. This can also be important in other chemical systems such as defects in crystals or active sites in large proteins, for which there is a large environment that must be considered explicitly using quantum mechanical methods. Ideally one would like to be able to study such systems using the most accurate electronic structure methods available, for example using hybrid exchange-correlation functionals such as discussed in Chapter 2, but such calculations may span hundreds or thousands of atoms, thus being computationally prohibitive. The quantum embedding techniques outlined in Chapter 3 offer a potentially attractive solution to this dilemma, by enabling a high-level calculation to be performed on an ‘active’ subsystem within a cheaper, less accurate ‘environment’ system. In particular, the embedded mean-field theory (EMFT) method discussed in Section 3.5 offers a very simple approach to quantum embedding grounded within the realms of the single-particle density matrix, which as we saw in Chapter 4 is also a central quantity for achieving linear-scaling with density-functional theory (DFT) calculations. This makes the EMFT scheme an ideal candidate for combining quantum embedding with linear-scaling DFT.

In this chapter, we will outline the implementation and testing of EMFT within ONETEP. We begin with a theoretical formulation of embedding techniques in the ONETEP framework, compared to the frozen density embedding (FDE) toy-model discussed in Chapter 6. From there, a brief derivation and discussion of the central quantities required for EMFT will be outlined, including the use of the block-orthogonalisation (BO) method discussed earlier. Testing of the new approach will follow for ground and excited state properties of a variety of molecular systems, demonstrating the accuracy and viability of the linear-scaling EMFT and the challenges that



remain with this approach. Reproduced in part with permission from [1]. Copyright 2019 American Chemical Society.

## 7.1 Combining EMFT and ONETEP

The starting point for the implementation of EMFT (see Section 3.5) within ONETEP is the same as our subsystem approach outlined in Chapter 6. We partition the atoms of the system into two subsystems, an active region A and an environment B, with corresponding NGWFs  $\{\phi^A\}$  and  $\{\phi^B\}$ . As before, the density kernel and overlap matrices are divided into diagonal blocks related to a particular subsystem and off-diagonal cross-overlap terms,

$$\mathbf{K} = \begin{pmatrix} \mathbf{K}_{AA} & \mathbf{K}_{AB} \\ \mathbf{K}_{BA} & \mathbf{K}_{BB} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} \mathbf{S}_{AA} & \mathbf{S}_{AB} \\ \mathbf{S}_{BA} & \mathbf{S}_{BB} \end{pmatrix}, \quad (7.1)$$

such that the density matrix can be defined in terms of its subsystem components,

$$\rho_{IJ}(\mathbf{r}, \mathbf{r}') = \sum_{\alpha \in I} \sum_{\beta \in J} \phi_{\alpha}^I(\mathbf{r}) K_{IJ}^{\alpha\beta} \phi_{\beta}^J(\mathbf{r}'), \quad I, J \in \{A, B\}. \quad (7.2)$$

The key difference with EMFT is that we now allow for the use of different Hamiltonian operators in the active and environment subsystems. It is assumed that the kinetic, Hartree and external potentials are the same for all subsystems, with the only divergence between the low-level and high-level Hamiltonians being contained within the exchange-correlation potential,

$$\hat{H}^{\text{high}} = \hat{T} + \hat{V}_H + \hat{V}_{\text{ext}} + \hat{V}_{\text{XC}}^{\text{high}}, \quad (7.3)$$

$$\hat{H}^{\text{low}} = \hat{T} + \hat{V}_H + \hat{V}_{\text{ext}} + \hat{V}_{\text{XC}}^{\text{low}}. \quad (7.4)$$

From this the Hamiltonian matrix is constructed such that the off-diagonal terms are treated at the low-level of theory using  $\hat{H}^{\text{low}}$ ,

$$\mathbf{H}^{\text{EMFT}} = \begin{pmatrix} \mathbf{H}_{AA}^{\text{high}} & \mathbf{H}_{AB}^{\text{low}} \\ \mathbf{H}_{BA}^{\text{low}} & \mathbf{H}_{BB}^{\text{low}} \end{pmatrix}. \quad (7.5)$$

The core Hamiltonian from Equation (3.51) consists of the kinetic and external potential terms, while the two-electron interaction energy is made up of the Hartree and exchange-correlation potentials. The optimal density matrix can thus be found by minimising the quantity

$$E^{\text{EMFT}} = \min_{\{K^{\alpha\beta}\}, \{\phi_{\alpha}^A\}} \text{Tr} [\mathbf{K} \mathbf{H}^{\text{EMFT}}]. \quad (7.6)$$

Following the same procedure as in Section 6.1.2, we can obtain the covariant EMFT NGWF gradient,

$$|g_{\alpha, A}\rangle = |\phi_i^A\rangle \tilde{\mathcal{A}}_{\alpha}^i + \hat{H}^{\text{high}} |\phi_i^A\rangle \tilde{\mathcal{B}}_{\alpha}^i + |\phi_j^B\rangle \tilde{\mathcal{C}}_{\alpha}^j + \hat{H}^{\text{low}} |\phi_j^B\rangle \tilde{\mathcal{D}}_{\alpha}^j, \quad (7.7)$$

where the coefficient matrices have the same form as Equations (6.24)–(6.27) but now include the adjusted EMFT Hamiltonian matrix in  $\mathbf{Q}$ . The key difference is that the contribution from the environment NGWFs is operated on by the low-level Hamiltonian  $\hat{H}^{\text{low}}$ .

We now have the tools at our disposal to utilise EMFT within ONETEP. To begin, the NGWFs and density kernel are initialised for the full system at the lower level of theory as in a conventional ONETEP calculation. The total and subsystem electronic densities  $n_{\text{tot}}$  and  $n_{\text{AA}}$  are computed from the respective NGWFs and density kernel elements. The Hamiltonian operator is built using the kinetic, Hartree and external potentials i.e. the first three terms in (6.7). At this stage, normally the density  $n_{\text{tot}}$  would be used to construct the exchange-correlation potential  $V_{\text{XC}}(\mathbf{r})$  using whatever level of approximation is chosen for the calculation, which is then added on to the Hamiltonian  $\hat{H}$ . Instead, with EMFT we also determine the regional potentials  $V_{\text{XC}}^{\text{low}, A}(\mathbf{r})$  and  $V_{\text{XC}}^{\text{high}, A}(\mathbf{r})$  using the subsystem density  $n_{\text{AA}}$  at the low and high levels of theory respectively. The final EMFT XC potential is then given by

$$V_{\text{XC}}^{\text{EMFT}}[n_{\text{tot}}, n_{\text{AA}}](\mathbf{r}) = V_{\text{XC}}^{\text{low}}[n_{\text{tot}}](\mathbf{r}) + \left( V_{\text{XC}}^{\text{high}, A}[n_{\text{AA}}](\mathbf{r}) - V_{\text{XC}}^{\text{low}, A}[n_{\text{AA}}](\mathbf{r}) \right). \quad (7.8)$$

This corrected potential is used to construct the high-level Hamiltonian  $\hat{H}^{\text{high}}$ , while the uncorrected  $V_{\text{XC}}^{\text{low}}(\mathbf{r})$  forms part of  $\hat{H}^{\text{low}}$ . Finally, the Hamiltonian matrix (7.5) is built, with which we can proceed with a standard ONETEP optimisation for the DFT-in-DFT system. The XC potentials and Hamiltonian are recalculated after each iteration of the density kernel and NGWF optimisation

### 7.1.1 Block-orthogonalised EMFT

In addition to regular EMFT, we have implemented the block-orthogonalised version (BO-EMFT) outlined in Section 3.5.1 for comparison. This proceeds in much the same manner as regular EMFT, except for an additional step in constructing the block-orthogonalised overlap matrix. The sub-blocks of the overlap matrix  $\mathbf{S}$  are built using the initial set of NGWFs  $\{\phi_{\alpha}^A\}$  and  $\{\phi_{\beta}^B\}$ . The inverse overlap matrix for the active system  $(\mathbf{S}_{\text{AA}})^{-1}$  is calculated in linear-scaling time using the Hotelling algorithm [2],

from which the projector matrix can be constructed,

$$\mathbf{P}^A_B = (\mathbf{S}_{AA})^{-1} \mathbf{S}_{AB}. \quad (7.9)$$

Note that the inverse matrix in (7.9) is a contravariant tensor, and so  $\mathbf{P}^A_B$  is a mixed contravariant-covariant quantity. This enables us to project out the environment NGWFs,

$$|\tilde{\phi}^B_\alpha\rangle = |\phi^B_\alpha\rangle - |\phi^A_i\rangle (\mathbf{P}^A_B)^i_\alpha, \quad (7.10)$$

such that, by construction, the overlap matrix in the block-orthogonalised basis becomes

$$\tilde{\mathbf{S}} = \begin{pmatrix} \mathbf{S}_{AA} & 0 \\ 0 & \tilde{\mathbf{S}}_{BB} \end{pmatrix}. \quad (7.11)$$

We can then carry out the optimisation of the density kernel and NGWFs with EMFT as discussed in Section 7.1. In the BO-EMFT formalism, the covariant NGWF gradient takes on the same form as (7.7) but with simplified coefficient matrices,

$$|g^{\text{BO}}_{\alpha,A}\rangle = |\phi^A_i\rangle (\tilde{\mathcal{A}}_{\text{BO}})^i_{.\alpha} + \hat{H}^{\text{high}} |\phi^A_i\rangle (\tilde{\mathcal{B}}_{\text{BO}})^i_{.\alpha} + |\phi^B_j\rangle (\tilde{\mathcal{C}}_{\text{BO}})^j_{.\alpha} + \hat{H}^{\text{low}} |\phi^B_j\rangle (\tilde{\mathcal{D}}_{\text{BO}})^j_{.\alpha}, \quad (7.12)$$

where

$$\tilde{\mathcal{A}}_{\text{BO}} = \mathbf{Q}_{AA} \mathbf{S}_{AA}, \quad (7.13)$$

$$\tilde{\mathcal{B}}_{\text{BO}} = \mathbf{K}_{AA} \mathbf{S}_{AA}, \quad (7.14)$$

$$\tilde{\mathcal{C}}_{\text{BO}} = \mathbf{Q}_{BA} \mathbf{S}_{AA}, \quad (7.15)$$

$$\tilde{\mathcal{D}}_{\text{BO}} = \mathbf{K}_{BA} \mathbf{S}_{AA}. \quad (7.16)$$

with similar equations for the environment orbital gradient. Interestingly, the subsystem NGWF gradients in the block-orthogonalised format is equivalent to using the subsystem matrix block  $\mathbf{S}_{AA}$  as the matrix for transforming from contravariant to covariant quantities, rather than the full overlap matrix  $\mathbf{S}$ .

## 7.2 Results — initial tests of EMFT in ONETEP

Following on from the frozen density embedding (FDE) toy model discussed in Chapter 6, we begin our testing of the EMFT implementation in ONETEP by studying the water dimer interaction potential. For now we restrict ourselves to using the semi-local PBE and LDA functionals for the active and environment exchange-correlation functionals. All calculations were performed using a kinetic energy cut-off of 1000 eV and an NGWF radius of  $7.0 a_0$ .

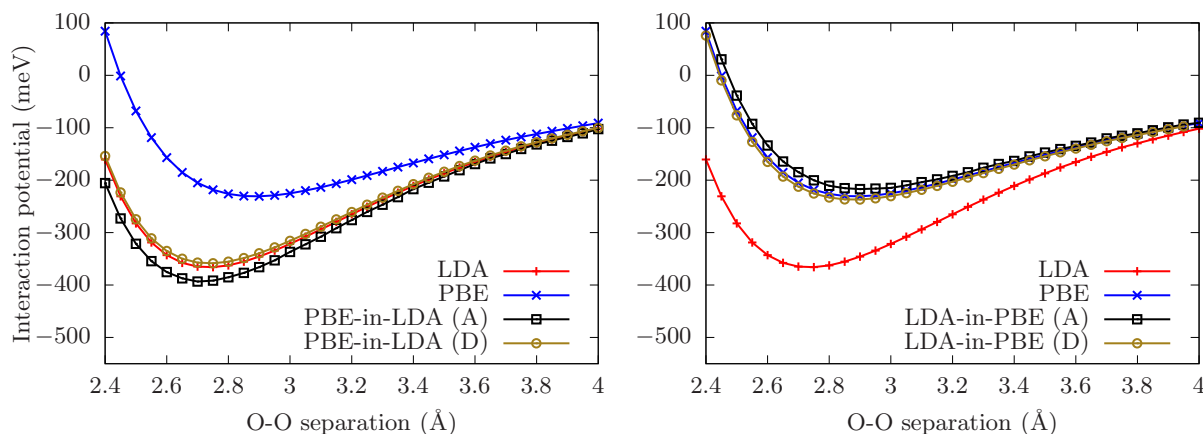


FIGURE 7.1: Interaction potential for a water dimer using EMFT, treating either the acceptor (A) or donor (D) as the active subsystem. ‘PBE-in-LDA’ refers to calculations where the active system is treated using the PBE exchange-correlation functional while the environment is treated using LDA, and vice versa for ‘LDA-in-PBE’. All NGWFs and density kernel elements are optimised self-consistently at all stages of the calculations.

### 7.2.1 NGWF optimisation with EMFT

Figure 7.1 shows the interaction potentials for PBE-in-LDA and LDA-in-PBE, treating either the donor or the acceptor molecule as the active region, compared to the pure PBE and LDA energy curves. We observe that, in both cases, the EMFT interaction potential closely follows the behaviour of the environment functional, with only a small energy shift differentiating PBE-in-LDA from pure LDA or LDA-in-PBE from pure PBE, and no significant change in the O-O separation minimum.

A possible reason for these poor results can be inferred from the form of the EMFT potential in (7.8). Since the interaction between the active subsystem and environment is described at the low-level of theory, with the EMFT correction only applied strictly to the active NGWFs and kernel, it appears that the interaction between the donor and acceptor water molecules is indeed dominated by the low-level functional, with only a perturbative change in the interaction potential arising from the use of the more accurate potential. This suggests that, as in Chapter 6 with our FDE model, we must treat the  $\text{O} \cdots \text{H}$  bond as the active system in order to have a sensible description of the bond interaction, since this is the region that determines the dimer interaction. However, an NGWF optimisation using EMFT in this configuration fails to converge, with the NGWF gradient stalling while the total energy continues to decrease well below the PBE or LDA total energies. Figure 7.2a shows the behaviour of the NGWF gradient during an optimisation of the water dimer structure treating the hydrogen bond as the active system. We see that for both PBE-in-LDA and LDA-in-PBE, the NGWF gradient does not converge as we would expect with regular PBE or LDA, or indeed when

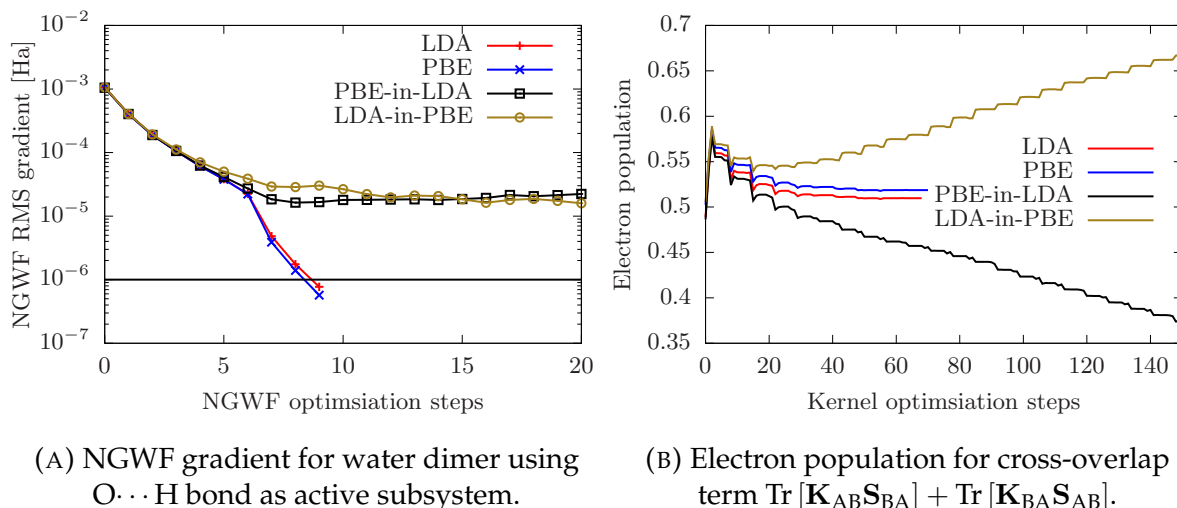


FIGURE 7.2: Behaviour of the NGWF gradient and electron populations during a self-consistent optimisation of the NGWFs and density kernel with EMFT, treating the O...H hydrogen bond as the active system.

the partitioning was done on the molecular level. Examining the electron population of the cross-overlap matrix blocks  $\text{Tr}[\mathbf{K}_{AB}\mathbf{S}_{BA}] + \text{Tr}[\mathbf{K}_{BA}\mathbf{S}_{AB}]$  in Figure 7.2b, we notice that in contrast to the single functional calculations, EMFT causes a significant change in the proportions of electric charge contained within each subregion. While the total electron number  $\text{Tr}[\mathbf{KS}]$  is conserved during the NGWF optimisation, the subsystem populations can be altered as the density is redistributed. Interestingly, the PBE-in-LDA and LDA-in-PBE calculations behave in an opposite manner to each other, with the former seeing a decrease in the cross-overlap population and the latter an increase. There are most likely two factors behind this phenomenon. First, the PBE functional generally gives lower energies for this system than LDA, so dumping charge in the PBE part of the system should, naively, lower the total energy. Secondly, the cross-overlap term is treated at the ‘lower’ (environment) level of theory, such that the interaction between the subsystems is computed with the low level functional. Thus, placing additional charge in the region that uses the functional with a lower energy minimum (i.e. PBE in this example) has the net effect of reducing the energy of the overall system to physically meaningless values. The lack of constraints on the electron density beyond normalisation of the total electron number thus facilitates the breaking of the self-consistent cycle as there is no longer a well-defined energy minimum to converge with respect to. This is reminiscent of the observation by Ding *et al.*[3] that the use of very different exchange-correlation functionals can cause the collapse of the self-consistent EMFT solution. The optimisation of the NGWF basis here introduces an additional degree of variation compared to past studies with EMFT, which may explain the collapse of the self-consistent cycle with the relatively similar PBE and LDA

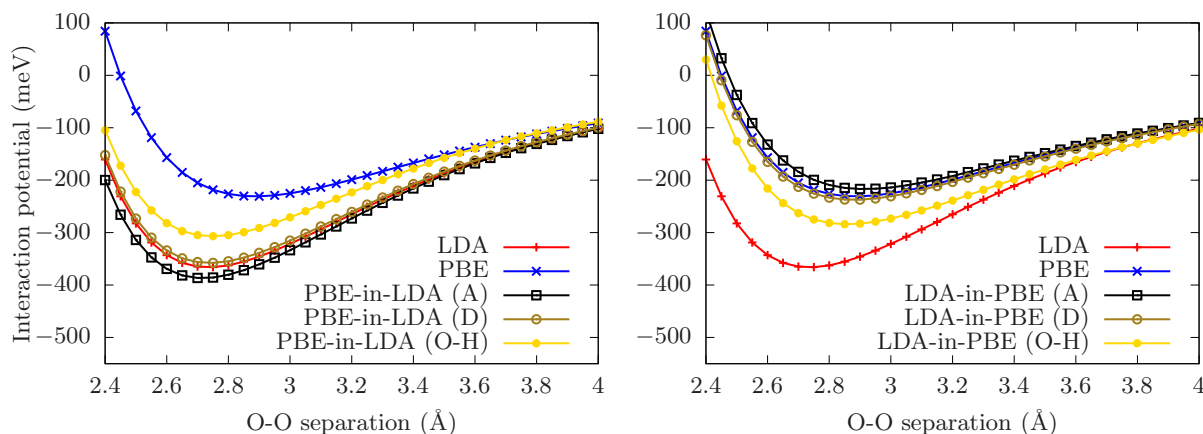


FIGURE 7.3: Interaction potential for water dimer. For EMFT calculations, only the density kernel is optimised at the high-level of theory, with the NGWFs fixed at the lower level — see text for more details.  
Active regions: acceptor (A), donor (D), hydrogen bond ( $\text{O} \cdots \text{H}$ ).

functionals, compared to the hybrid functionals used by Ding *et al.*

## 7.2.2 Kernel-only optimisation

The difficulties presented in the preceding section with NGWF optimisation using EMFT motivates an alteration to the self-consistent algorithm laid out in Section 7.1. Instead of optimising the density kernel and NGWFs using the EMFT Hamiltonian, we perform an initial optimisation of these quantities at the low-level of theory, then proceed with a kernel-only optimisation with EMFT using the NGWFs optimised with the low-level functional. This averts the issue observed previously with charge spillover causing the breakdown of the self-consistent cycle, while preserving the theoretical computational benefits of treating only a small part of the calculation at the higher level of theory, in this case as a post-processing step. For the moment we focus solely on our PBE/LDA example.

Figure 7.3 shows the results for the kernel-only EMFT calculation for various combinations of functional and active subsystems. As with the NGWF optimisation, there is scarcely any difference between the full low-level calculation and treating either the acceptor or the donor as the active system with EMFT. With the  $\text{O} \cdots \text{H}$  active system, we finally see an appreciable difference between the low-level and EMFT calculations, whereby the energy minimum lying between the high- and low-level energies, producing a PBE-in-LDA minimum of  $-307$  meV compared to the PBE and LDA results of  $-231$  meV and  $-366$  meV respectively. The minimum O-O separation with EMFT coincides with the low-level minimum, which amounts to  $2.75$  meV for both LDA and PBE-in-LDA compared to  $2.90$  meV for PBE only. It thus appears that selecting the

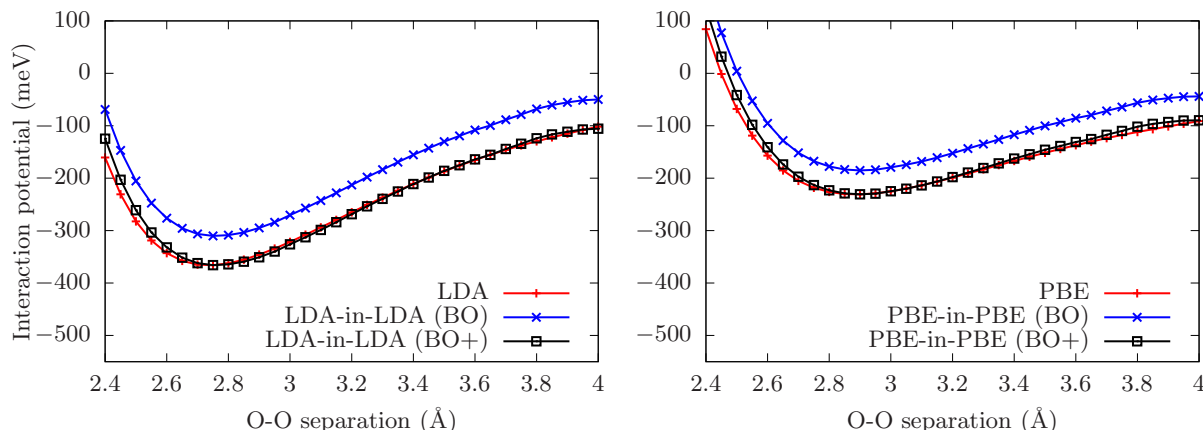


FIGURE 7.4: Interaction potential for water dimer using block orthogonalised (BO) and corrected (BO+) NGWFs, treating the  $\text{O} \cdots \text{H}$  bond as the active subsystem.

$\text{O} \cdots \text{H}$  bond as the active system with the kernel-only EMFT calculation produces the best results of the systems tested so far, though there is clearly room to improve in terms of both the depth of the potential and the molecule separation.

### 7.2.3 Block orthogonalisation

Given the resemblance between the difficulties discussed above regarding the use of EMFT with NGWFs and the breakdown of the self-consistent method observed in past studies, we now look at the possibility of using block-orthogonalised EMFT (BO-EMFT) in ONETEP. As before, the NGWFs and density kernel are initially evaluated at the low-level of theory, but now we block-orthogonalise the NGWFs in the environment with respect to the active subsystem orbitals as discussed in Section 7.1.1. The density kernel is then reoptimised using BO-EMFT, holding the (now orthogonalised) NGWFs fixed. This follows the formal structure of the initial BO-EMFT implementation of Ding *et al.* while also maintaining the use of a minimal set of basis functions. From now on, all embedding calculations will treat the  $\text{O} \cdots \text{H}$  hydrogen bond as the active system in the water dimer.

One potential difficulty with this procedure is that the block-orthogonalised orbitals are not an optimal choice of basis functions for the system, which may cause an error in the interaction potential similar to those observed in Chapter 6 for molecular orbitals (MOs) and selective NGWF optimisation. As a test of the validity of this approach, Figure 7.4 shows the results for block-orthogonalised LDA-in-LDA and PBE-in-PBE, compared to normal SCF calculations. We see that in both cases the location of the minimal bond length is preserved, though there is also an increase in the energies leading to a reduction in the depth of the potential. This effect amounts to 46 meV for PBE and



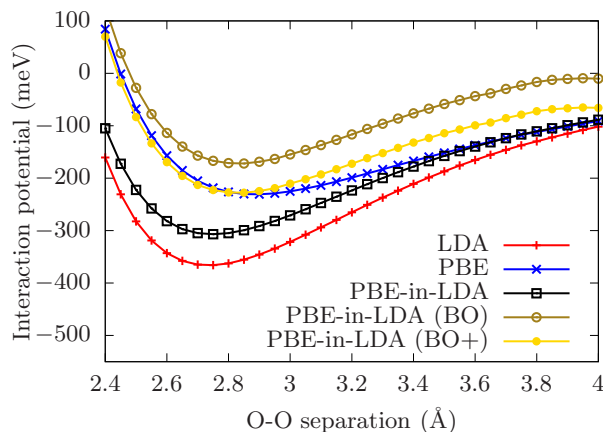


FIGURE 7.5: PBE-in-LDA interaction potential for water dimer using block orthogonalised (BO) NGWFs, treating the  $\text{O} \cdots \text{H}$  bond as the active subsystem.

56 meV for LDA at the potential minima. Figure 7.4 also shows a corrected version of the block-orthogonalised calculations (BO+), where we have shifted the energies such that the energy minima coincide with the non-orthogonalised NGWF results. In both cases, the BO+ curve shows excellent agreement with the normal NGWF results, suggesting that the discrepancy between the BO and non-BO results for this system can be accounted for by a simple energetic shift in the potential. Thus, by evaluating the difference in the depth of the energy minima for PBE with NGWFs and PBE-in-PBE with BO-NGWFs, we can avoid the need to perform expensive additional counterpoise corrections and maintain the use of a minimal set of basis functions to obtain accurate interaction energies for this system.

Switching our attention to a PBE-in-LDA system, Figure 7.5 illustrates the use of both standard and BO NGWFs evaluated at the LDA level. The energy minimum has a value of  $-172$  meV, compared to the actual PBE minimum of  $-231$  meV and the non-orthogonalised value of  $-307$  meV. Significantly, the optimal O-O distance is located at a separation of  $2.85$  Å, which is much closer to the expected PBE result of  $2.90$  Å than the non-orthogonalised version of  $2.75$  Å. Applying the BO+ correction of  $56$  meV for LDA, the energy minimum is found at  $-228$  meV, in very close agreement with the full PBE result and a great improvement over all our previous attempts at PBE-in-LDA embedding for this system. While visually the two curves diverge at greater distances, the agreement for the bond minimum verifies the accuracy of the BO+ method for an embedded system with different functionals.



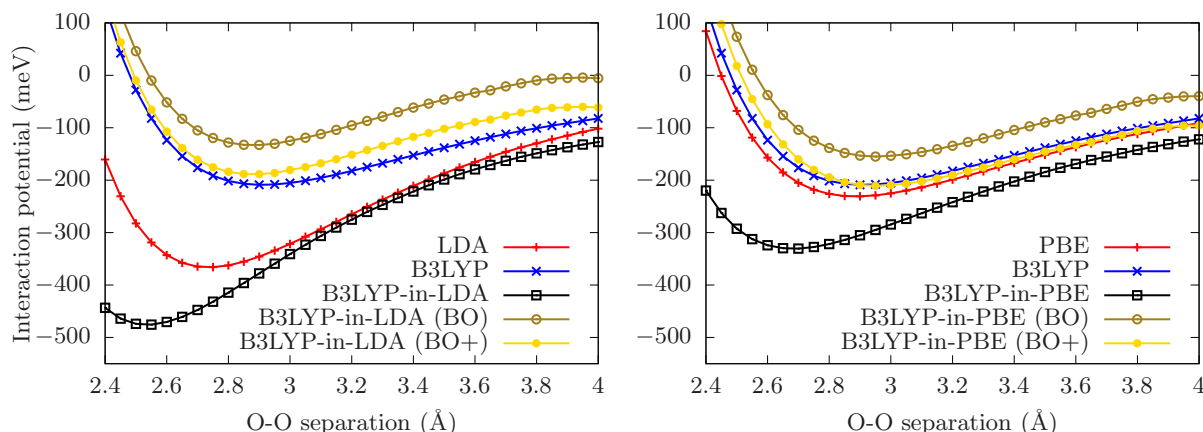


FIGURE 7.6: Interaction potential for a water dimer. For EMFT calculations, B3LYP is used for the active subsystem, with the environment treated at either the LDA (left) or PBE (right) levels of theory. All EMFT were performed by optimising the NGWFs at the low-level of theory, before optimising the density kernel with EMFT. Also shown is the effect of block-orthogonalising (BO) the environment NGWFs, with an additional correction for the block-orthogonalisation error (BO+) also included.

## 7.2.4 Hybrid functionals

Having established that it is possible to reproduce the behaviour of a water dimer with the PBE functional by treating only a subsystem (the hydrogen bond) with PBE while the rest is described at the LDA level of theory, we now turn our attention to the use of hybrid functionals. Computationally the exchange matrix is the bottleneck during a ONETEP exact-exchange calculation, so reducing the size of the high-level region by embedding the Hartree-Fock exchange region within a semi-local exchange-correlation environment would be of significant computational value. We will initially focus on B3LYP-in-LDA since the behaviour of these two functionals for this system is very different and is thus a more challenging test of the algorithm. Initially the NGWFs will be optimised at the low-level of theory (LDA), before performing a kernel-only LNV optimisation of the density kernel using B3LYP-in-LDA.

Figure 7.6 illustrates the behaviour of an EMFT B3LYP-in-LDA calculation using both non-orthogonalised and block-orthogonalised NGWFs. Without BO, EMFT yields dreadful results, with the potential well being several hundred meV too deep and the bond length too short. The BO+ approach reproduces the correct O-O separation of 2.90 Å, while the energy minimum is −189 meV, compared to the pure B3LYP result of −209 meV. While this is not as accurate as the aforementioned PBE-in-LDA result, it is clear that BO-EMFT using LDA NGWFs can provide a good description of the interaction between two water dimers.

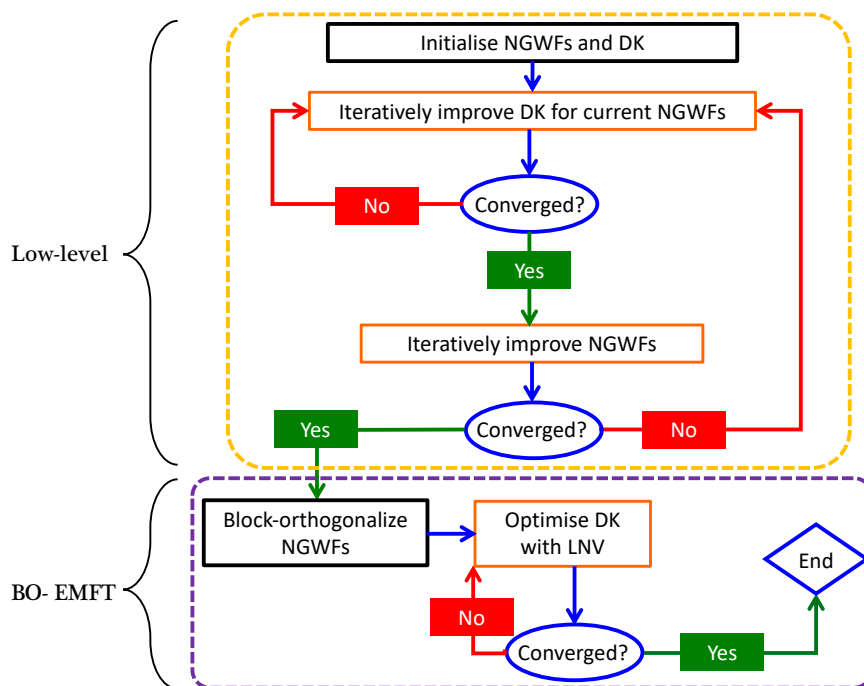


FIGURE 7.7: Block-orthogonalised embedded mean-field theory (BO-EMFT) algorithm as implemented in ONETEP. Compare with Figure 4.2

For a system such as this, it would be more realistic to make use of PBE as the low-level functional, although the results here are less interesting given the close agreement between the PBE and B3LYP energy curves as shown in Figure 7.6. Nonetheless, we once again see that regular EMFT with kernel optimisation produces poor results that significantly overbind the dimer. The BO+ method again gives results that are in agreement with full B3LYP, with a potential depth of  $-201$  meV. The B3LYP-in-PBE and B3LYP results do diverge at short separation distances, but in spite of this the energy minimum is reconstructed to a high level of accuracy. Thus we can conclude that our implementation of BO-EMFT is working correctly and is capable of accurately reproducing the behaviour of the high-level of theory with exact exchange.

These findings suggest that an alteration to the standard ONETEP algorithm is required for performing calculations with BO-EMFT for embedded systems. Figure 7.7 illustrates an adjusted form of the flow diagram shown in Figure 4.2, whereby the first stage is performed entirely at the low-level of theory, before block-orthogonalising the NGWFs and reoptimising the density kernel with the EMFT Hamiltonian. This is the form of BO-EMFT that we will utilise throughout the remainder of this chapter.

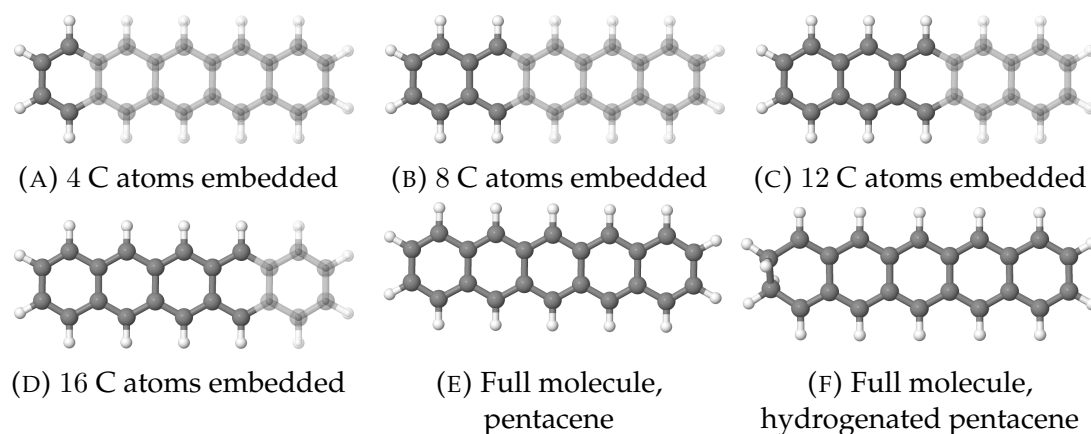


FIGURE 7.8: Embedding regions used for the various calculations on pentacene and terminally-hydrogenated pentacene. Active region atoms are displayed as opaque, environment atoms are translucent. (A)-(D) show the embedding regions used for pentacene — the regions used for hydrogenated pentacene are identical, but with the 2 additional hydrogen atoms bonded to the two left-most carbon atoms. (E) and (F) show the full structure of molecular pentacene and terminally-hydrogenated pentacene, respectively. Each configuration is identified by the number of carbon atoms included in the active region. Note that all hydrogen atoms are included in the same subsystem as their neighbouring carbon. This figure is reproduced from[1].

### 7.3 Results — Hydrogenation of pentacene

Pentacene is a molecule of widespread interest within the organic chemistry community, due to its utilisation for applications such as organic photovoltaics [4] and, as explored in Chapter 5, room-temperature masers [5]. Past studies of EMFT have used pentacene as a test system [3, 6], making it an ideal candidate to benchmark the use of EMFT on a more complex system than the water dimer used earlier. The structure of pentacene consists of five fused benzene rings, giving rise to molecular orbitals that are delocalised across the system. Since relative energies are more meaningful and generally of greater interest than total energies in *ab initio* contexts, we will focus our attention on determining the reaction energy for the terminal hydrogenation of pentacene, whereby two hydrogen atoms are bonded with two carbon atoms at one end of the molecule as shown in Figure 7.8.

Our tests will focus on partitioning the rings between the low- and high-level regions shown in Figure 7.8. We use PBE as the low-level of theory and B3LYP as the high-level functional for all calculations here, with an NGWF radius of 9.0 bohr (4.76 Å) and a cutoff energy of 800 eV. For EMFT calculations, we make use of the block-orthogonalised method (BO-EMFT) outlined in Figure 7.7, whereby the density kernel and NGWFs are optimised at the PBE level of theory, before the NGWFs are

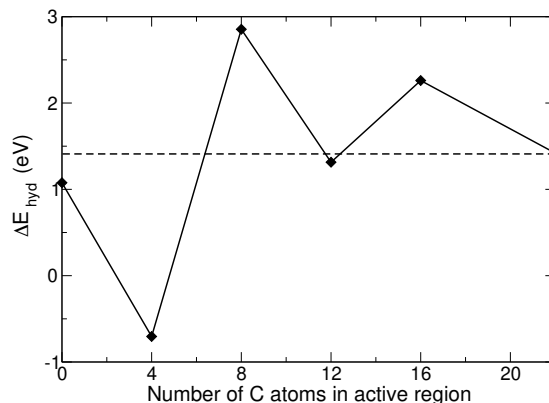


FIGURE 7.9: Terminal hydrogenation energy,  $\Delta E_{\text{hyd}}$ , of a pentacene molecule as a function of the number of C atoms treated with B3LYP, with all NGWFs optimised at the PBE level of theory. The energy obtained when both the kernel *and* the NGWFs are optimised using B3LYP is marked with a dashed line.  $\Delta E_{\text{hyd}}$  is calculated from Eq. (7.17). The precise atoms included within the embedded active region for each calculation are shown in Figure 7.8. This figure is reproduced from[1].

block-orthogonalised and the kernel reoptimised using BO-EMFT for B3LYP-in-PBE calculations. Two full molecule calculations are performed for each molecule at the B3LYP level of theory, the first using NGWFs optimised at the PBE level of theory, and the second with NGWFs optimised at the B3LYP level. The latter acts as the benchmark for all other calculations.

For each embedding configuration, the terminal hydrogenation energy is defined as

$$\Delta E_{\text{hyd}} = E_{\text{HP}} - E_{\text{P}} - E_{\text{H}_2}, \quad (7.17)$$

where  $E_{\text{HP}}$ ,  $E_{\text{P}}$ , and  $E_{\text{H}_2}$  are the energy of terminally-hydrogenated pentacene, pentacene, and an isolated  $\text{H}_2$  hydrogen molecule respectively. While  $E_{\text{HP}}$  and  $E_{\text{P}}$  are calculated with EMFT using the various subsystem partitions illustrated in Figure 7.8,  $E_{\text{H}_2}$  is always evaluated at the same level of theory as the extra hydrogen atoms in the active part of the terminally hydrogenated structure.

The resulting hydrogenation energies are displayed in Figure 7.9. Starting on the left of the graph, the first data point, corresponding to zero carbon atoms in the active region, is a full PBE calculation. The final data point has all C atoms in the active region and is thus a B3LYP-level calculation using PBE NGWFs. All intermediary data points are BO-EMFT calculations whereby the number of C atoms specified in both the normal and hydrogenated pentacene molecules (the partitions shown in Figure 7.8) are treated at the B3LYP level with PBE NGWFs, while the remaining parts of the molecules are treated at the PBE level with NGWFs block-orthogonalised relative to

the active subsystem orbitals. The dashed line represents the benchmark B3LYP calculation whereby all NGWFs and density kernel elements are evaluated at the B3LYP level in a non-embedding run.

Examining these results, we first notice that the B3LYP-level calculation using PBE NGWFs is in excellent agreement with the benchmark B3LYP calculation. The difference amounts to 0.01 eV, an order of magnitude smaller than the difference between the benchmark and all other results presented here. This result is well within the bounds of chemical accuracy, and thus demonstrates that the use of low-level NGWFs for the active subsystem at the high-level of theory is a valid approximation. Significantly, this minimises the discrepancies between the basis functions used in each calculation, thus making direct comparison of results more meaningful.

Turning to the EMFT results, we see that increasing the size of the active region causes the hydrogenation energy  $\Delta E_{\text{hyd}}$  to tend towards the full B3LYP result in an oscillatory manner. Increasing the number of active C atoms from four to eight reduces the absolute error by 32%, from 2.12 eV to 1.44 eV. There is a further 59% reduction in absolute error by increasing the active region size to 16 C atoms. In addition, we see very close agreement between the 12 C atom result and the full B3LYP calculation, though this is most likely just a fortuitous coincidence. We have thus demonstrated that it is possible to systematically reduce the error in the hydrogenation energy from using EMFT by increasing the size of the active region, such that the accuracy of our embedding calculation is readily controllable. From a practical perspective, this suggests that we can indeed use the expensive hybrid functionals in a small part of a complex structure in order to obtain results with reasonable accuracy at a significantly reduced computation cost. We will take advantage of this in the next section by applying our BO-EMFT approach to systems that would otherwise not be accessible without a joint linear-scaling and embedding approach.

## 7.4 Excitation energies with pentacene derivatives

With the validity of our EMFT implementation confirmed for ground state calculations, we now turn our attention to excited state energies. For this, we will focus on pentacene and its nitrogen-doped derivatives in a *p*-terphenyl cluster host that we used earlier in Chapter 5 for our study of solvation effects. At this stage, we possess the tools to perform a true embedding calculation by treating the dopant molecule as the active subsystem and the *p*-terphenyl environment as the low-level environment. As before, a BO-EMFT calculation proceeds by optimising the NGWFs and density kernel at the

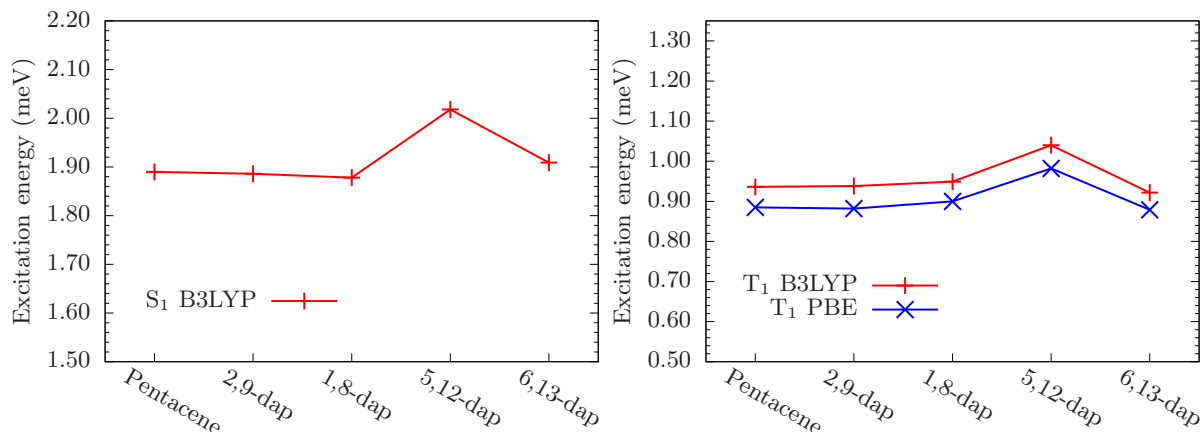


FIGURE 7.10: Singlet and triplet excitation energies of chromophores in vacuum. Both PBE and B3LYP calculations make use of NGWFs optimised at the PBE level of theory.

low-level of theory, before block-orthogonalising the NGWFs and reoptimising the kernel with the LNV algorithm. The low-level functional here is PBE and the high-level is the exact exchange B3LYP functional.

Although EMFT is only implemented for ground state optimisation of the NGWFs and density kernel at present, we can still obtain excitation energies via the use of spin-polarised density matrices as outlined in Section 4.3.4. The EMFT formulation within a triplet configuration involves a trivial change to our preceding discussion by making the density kernel the sum of an up and down kernel, such that the triplet excitation energy  $\Delta E_{S_0 \rightarrow T_1}$  can be determined using the  $\Delta$ SCF method as the difference in total energy between two BO-EMFT calculations. For pentacene, the singlet excited state  $S_1$  is dominated by the HOMO  $\rightarrow$  LUMO transition. Consequently, the singlet-triplet splitting  $\Delta E_{T_1 \rightarrow S_1}$  can be determined by the virial exciton method [7], whereby the splitting is defined as the electron-repulsion integral (ERI) constructed from the frontier orbitals,

$$\Delta E_{T_1 \rightarrow S_1} = (\psi_H \psi_H | \psi_L \psi_L) = \iint \frac{\psi_H^*(\mathbf{r}) \psi_H(\mathbf{r}) \psi_L(\mathbf{r}') \psi_L^*(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r} d^3\mathbf{r}', \quad (7.18)$$

where  $\psi_H$  and  $\psi_L$  are the HOMO and LUMO wavefunctions, respectively. This method has previously been applied by Becke and coworkers to the evaluation of  $S_1$  for single molecules, producing promising results for a multitude of systems, including aromatic systems with significant charge-transfer effects [7, 8]. The quantity in Eq. (7.18) is readily available from an exact exchange calculation in ONETEP, which enables us to calculate this in a post-processing step after a triplet ground state energy calculation. Thus the singlet excitation energy  $\Delta E_{S_0 \rightarrow S_1}$  can be easily found as the sum of the triplet excitation energy and the singlet-triplet splitting.

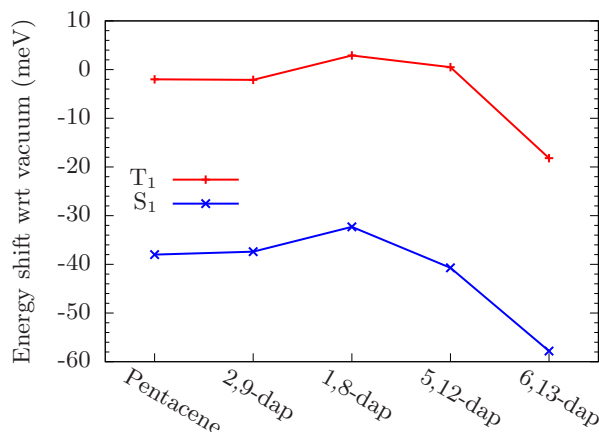


FIGURE 7.11: Solvatochromic shift of  $T_1$  and  $S_1$  states for various dopant molecules in  $p$ -terphenyl cluster relative to vacuum.

Figure 7.10 shows the absolute  $T_1$  and  $S_1$  excitation energies for the various chromophores in vacuum. For  $T_1$ , the excitation energies with B3LYP (using PBE NGWFs) follow the same trend as the PBE only results and, compared to the equivalent vacuum calculations with OT-LC $\omega$ PBE, the pentacene energy is much closer to the expected experimental value of 0.86 eV. The energies for the  $S_1$  state lie in an intermediate range between the PBE and OT-LC $\omega$ PBE results discussed in Chapter 5. While somewhat underestimated compared to experiment, the pentacene B3LYP excitation energy is in much closer agreement than PBE for the vacuum state. From the perspective of embedding, the more interesting phenomenon to consider is whether the B3LYP-in-PBE calculation can capture the solvatochromic shift for  $\Delta E_{S_0 \rightarrow S_1}$ , since we expect this excitation to be much more significantly affected than  $\Delta E_{S_0 \rightarrow T_1}$ .

Figure 7.11 displays the solvatochromic shift for the singlet and triplet excitations using B3LYP for the chromophore and PBE for the environment. Focusing on the  $T_1$  state, we see that there is a negligible change (less than 3 meV) in  $\Delta E_{S_0 \rightarrow T_1}$  for every chromophore except 6,13-diazapentacene, which undergoes a redshift of 18 meV. This is a reduction compared to the predicted redshift of 30 meV for OT-LC $\omega$ PBE in explicit solvent from earlier, however we do see the same qualitative trend using hybrid BO-EMFT as we did with the range-separated hybrid functional for all chromophores, namely that 6,13-dap is the only molecule to witness a significant change in its  $\Delta E_{S_0 \rightarrow T_1}$  energy.

Turning our attention to the  $S_1$  state, as expected there is a much more appreciable redshift for all molecules in the  $\Delta E_{S_0 \rightarrow S_1}$  energy. The singlet calculation yields a redshift of 38 meV for pentacene, compared to the 65 meV explicit host shift we saw for OT-LC $\omega$ PBE in Figure 5.7. For the nitrogen-based analogues, the smallest redshift observed is 32 meV for 1,8-dap, while 6,13-dap sees the greatest redshift of 58 meV,



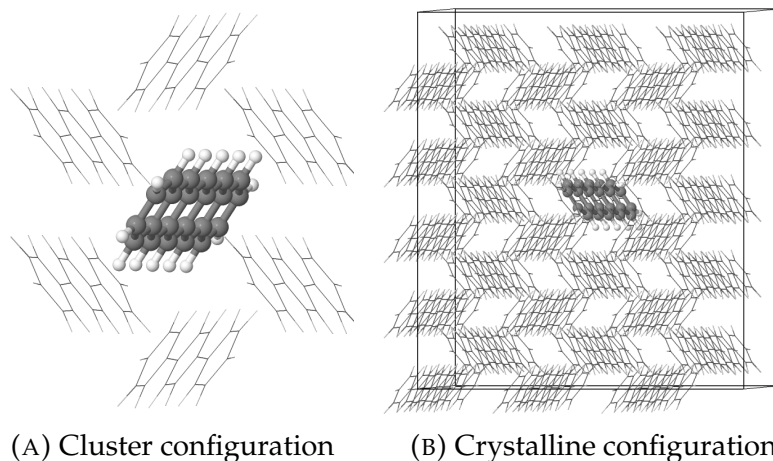


FIGURE 7.12: The two configurations used in this study for the pentacene in *p*-terphenyl system. (A) shows the cluster configuration, consisting of a pentacene molecule surrounded by six nearest neighbour *p*-terphenyl molecules, the same structure as used previously in Chapters 5 and 6. (B) shows the crystalline configuration, built using a  $3 \times 5 \times 3$  supercell of crystalline *p*-terphenyl with the central *p*-terphenyl replaced by a pentacene molecule. For both images, *p*-terphenyl molecules are shown as wireframes, and the pentacene using a ball-and-stick model for ease of viewing. H and C atoms are white and grey respectively. This figure is reproduced from [1].

compared to 67 meV and 92 meV for OT-LC $\omega$ PBE respectively. It thus appears that the B3LYP-in-PBE calculation is underestimating the OT-LC $\omega$ PBE results in explicit solvent by 27–35 meV for all molecules. Significantly, we do not witness anything comparable to the spurious charge-transfer (CT) mixing that was observed using PBE in explicit host for these molecules in Chapter 5, with the B3LYP-in-PBE results instead following the general trend of the pure OT-LC $\omega$ PBE results. Indeed, we see the same qualitative trends for the singlet excitation energy in explicit host between Figures 5.7 and 7.11, with pentacene, 2,9-, 1,8- and 5,12-dap all yielding comparable redshifts for  $\Delta E_{S_0 \rightarrow S_1}$ , while 6,13-dap yields a redshift that is greater by  $\sim 20$ –25 meV.

## 7.5 EMFT for large systems

So far we have focused on the application of EMFT to the treatment of systems of up to a few hundred atoms. For verification purposes this is useful as it enables direct comparison to past studies with hybrid functionals, which can treat systems of similar size. However, this does not push the boundaries of what is possible with EMFT, since the use of a cheaper semi-local functional for the environment in ONETEP allows the simulation of molecular structures containing thousands of atoms. We can therefore go beyond our cluster calculations by considering a periodic crystal of *p*-terphenyl. This



Configuration	$\Delta E_{S_0 \rightarrow T_1}$ (eV)		
	PBE	B3LYP-in-PBE	Exp.[11]
Vacuum	0.885	0.942	$0.86 \pm 0.03$
Cluster	0.879	0.934	
Crystal	0.891	0.934	

TABLE 7.1: Energy differences between the  $S_0$  singlet ground state and the  $T_1$  triplet ground state for pentacene, both in vacuum and in the explicit  $p$ -terphenyl hosts shown in Figure 7.12. For B3LYP-in-PBE calculations in vacuum, PBE NGWFs were used, with only the density kernel being re-optimised at the B3LYP level. In host, B3LYP-in-PBE refers to an EMFT calculation, whereby the pentacene is treated as the active subsystem and the  $p$ -terphenyl as the low-level environment. All calculations were performed using  $\Delta$ SCF. This figure is reproduced from[1].

structure [9] was obtained from the Cambridge Structural Database [10]. A  $3 \times 5 \times 3$  supercell with 90  $p$ -terphenyl molecules was produced, and a pentacene substituted for the central  $p$ -terphenyl molecule, similar to the cluster geometry used earlier. The resulting structure, shown in Figure 7.12, consists of 2884 atoms, an order of magnitude greater than the 228 atoms in our pentacene-doped  $p$ -terphenyl cluster. Such a calculation would only be viable using linear-scaling DFT as implemented in ONETEP, and the use of quantum embedding via EMFT enables us to consider the behaviour of a pentacene molecule at the B3LYP level surrounded by a several thousand atom, semi-local  $p$ -terphenyl explicit solvent.

Table 7.1 shows the  $\Delta E_{S_0 \rightarrow T_1}$  excitation energies for PBE and B3LYP-in-PBE calculations of pentacene in both vacuum and explicit host compared to experiment. For B3LYP-in-PBE, we see that the cluster and periodic crystal structures yield excitation energies that match to the nearest meV, indicating a solvatochromic shift of only 8 meV relative to vacuum. This is in line with the experimental observation that the  $T_1$  state remains largely unaffected in a host structure, retaining its highly localised Frenkel exciton nature [11]. In addition, the agreement between the cluster and crystal show that the cluster is in fact producing a valid description of the excitation in the  $p$ -terphenyl host for  $T_1$ , in line with our earlier discussions in Chapter 5.

Looking at the  $S_1$  state, Table 7.2 shows a comparison of the singlet energies  $\Delta E_{S_0 \rightarrow S_1}$  for each functional in host compared to experiment. Both PBE and B3LYP-in-PBE significantly underestimate the vacuum and host excitation energies compared to experiment. The PBE vacuum energy of 1.835 eV is much higher than the results we saw earlier in Chapter 5 using time-dependent density functional theory (TDDFT), and indeed compared to TDDFT calculations in the literature [15, 16]. This would suggest that the virial exciton method can reduce somewhat the error that arises from the use of semi-local TDDFT in linear acenes. For B3LYP-in-PBE, the  $\Delta E_{S_0 \rightarrow S_1}$  energy of 1.895 eV is in

Configuration	$\Delta E_{S_0 \rightarrow S_1}$ (eV)		
	PBE[12]	B3LYP-in-PBE	Exp.
Vacuum	1.835	1.895	2.31[13]
Cluster	1.801	1.853	$2.1 \pm 0.1$ [14]
Crystal	1.804	1.756	

TABLE 7.2: Energy differences between the  $S_0$  singlet ground state and the  $S_1$  singlet excited state for pentacene, both in vacuum and in the explicit  $p$ -terphenyl hosts shown in Figure 7.12. The comments for Table 7.1 regarding B3LYP-in-PBE also apply here. All calculations were performed using the virial exciton method [7]. This figure is reproduced from[1].

line with past first principles calculations of the excitation energy using TDDFT with the B3LYP functional [17].

Focusing on the solvatochromic shift from vacuum to explicit host, we see a much greater effect than for the  $T_1$  state, as we would expect from our previous discussion of pentacene in Chapter 5. The PBE functional yields a redshift of 34 meV and 31 meV in  $\Delta E_{S_0 \rightarrow S_1}$  for the cluster and crystal structures, respectively. The former is in fact lower than the TDDFT result presented in Chapter 5. This could be explained by the underestimation of long-range charge-transfer effects with semi-local TDDFT, such that the  $S_1$  state is not a pure HOMO→LUMO transition in the  $p$ -terphenyl host. The lack of a significant redshift in the crystal structure using the Becke virial exciton method would appear to confirm that the PBE excitation is largely unaffected by the presence of the wider crystal, and that the larger redshift with TDDFT is simply an artifact of the generally dismal behaviour of semi-local TDDFT in describing the excitonic properties of linear acenes like pentacene.

B3LYP-in-PBE produces significantly different results to PBE in the explicit host structure. The cluster calculation gives a redshift of 42 meV, while the periodic crystal produces a redshift of 139 meV. The crystal result represents a significant improvement in agreement with the experimental finding of  $\sim 200$  meV for pentacene in  $p$ -terphenyl [14]. This suggests that the six  $p$ -terphenyl cluster is too small to include all the effects of the environment that one would expect from a fully periodic crystal host. For instance, the lack of end-to-end molecules in the same orientation as the pentacene may have significant effects upon the interaction between the dopant and its environment, or in this case there are additional effects from including an additional layer of  $p$ -terphenyls surrounding the six molecule cluster. An alternative explanation may be that this is a spurious effect caused by the dilution of the pentacene  $S_0 \rightarrow S_1$  excitation in the host structure — if this is not a strictly HOMO→LUMO transition, then the virial exciton method may not be valid for this system, which would undermine the reliability of the calculation of  $\Delta E_{T_1 \rightarrow S_1}$ . In Chapter 5 we saw that this excitation is primarily HOMO→LUMO for pentacene in cluster, so if the cluster is sufficient to describe host

effects then it is not clear why the difference between these calculations is so large. It would therefore seem reasonable to conclude that the increased solvatochromic shift for the periodic crystal represents a real effect that can be described using B3LYP but is omitted from a pure PBE analysis of the system. Further studies of the virial exciton method are required to verify the validity of this approach for embedded systems.

## 7.6 Summary

In this chapter we have outlined the implementation and testing of the embedded mean-field theory (EMFT) approach to quantum embedding within the ONETEP programme. We have seen that the block-orthogonalised formalism (BO-EMFT) produces results that are in good agreement with non-embedding calculations for ground and excited properties of a variety of systems, including single molecules and molecular crystals. This approach is the most transferable method for using EMFT in *ab initio* calculations as we have seen its usefulness when applied to subsystem partitions that standard EMFT can struggle to describe, such as cutting covalent bonds. Based on BO-EMFT, we have extended the ONETEP linear-scaling DFT package to enable multi-level DFT-in-DFT embedding with different exchange-correlation functionals, including hybrid exact-exchange functionals such as B3LYP in a semi-local host such as PBE. This approach enables calculations to be performed on systems containing thousands of atoms with an active subsystem described using exact-exchange. In terms of the pentacene in *p*-terphenyl system that forms the basis for the room-temperature maser, BO-EMFT can be used to qualitatively capture the importance of host environment effects, both in terms of the selection of dopant molecule and the size of the low-level explicit host, indicating behaviour that cannot be understood from the use of purely semi-local functionals.

## References

- [1] Joseph C A Prentice, Robert J Charlton, Arash A. Mostofi, and Peter D Haynes. In: *Journal of Chemical Theory and Computation* (2019), acs.jctc.9b00956. ISSN: 1549-9618. DOI: [10.1021/acs.jctc.9b00956](https://doi.org/10.1021/acs.jctc.9b00956).
- [2] T. Ozaki. In: *Physical Review B* 64.19 (2001), p. 195110. ISSN: 0163-1829. DOI: [10.1103/PhysRevB.64.195110](https://doi.org/10.1103/PhysRevB.64.195110).
- [3] Feizhi Ding, Frederick R. Manby, and Thomas F. Miller. In: *Journal of Chemical Theory and Computation* (2017). ISSN: 15499626. DOI: [10.1021/acs.jctc.6b01065](https://doi.org/10.1021/acs.jctc.6b01065).
- [4] S. Yoo, B. Domercq, and B. Kippelen. In: *Applied Physics Letters* 85.22 (2004), pp. 5427–5429. ISSN: 00036951. DOI: [10.1063/1.1829777](https://doi.org/10.1063/1.1829777).
- [5] Mark Oxborrow, Jonathan D. Breeze, and Neil M. Alford. In: *Nature* 488.7411 (2012), pp. 353–356. ISSN: 0028-0836. DOI: [10.1038/nature11339](https://doi.org/10.1038/nature11339).
- [6] Mark E. Fornace et al. In: *Journal of Chemical Theory and Computation* 11.2 (2015), pp. 568–580. ISSN: 15499626. DOI: [10.1021/ct5011032](https://doi.org/10.1021/ct5011032).
- [7] A. D. Becke. In: *J. Chem. Phys.* 148 (2018), p. 044112. DOI: [10.1063/1.5012033](https://doi.org/10.1063/1.5012033).
- [8] X. Feng, A. D. Becke, and E. R. Johnson. In: *J. Chem. Phys.* 149.23 (2018). DOI: [10.1063/1.5078515](https://doi.org/10.1063/1.5078515).
- [9] H. M. Rietveld, E. N. Maslen, and C. J. B. Clews. In: *Acta Crystallographica Section B* 26.6 (1970), pp. 693–706. DOI: [10.1107/S0567740870003023](https://doi.org/10.1107/S0567740870003023).
- [10] C. R. Groom, I. J. Bruno, M. P. Lightfoot, and S. C. Ward. In: *Acta. Cryst. B.* 72 (2016), p. 171. DOI: [10.1107/S2052520616003954](https://doi.org/10.1107/S2052520616003954).
- [11] J. Burgos, M. Pope, C. E. Swenberg, and R. R. Alfano. In: *Phys. Status Solidi B* 83 (1977), p. 249. DOI: [10.1002/pssb.2220830127](https://doi.org/10.1002/pssb.2220830127).
- [12] R. J. Charlton et al. In: *The Journal of Chemical Physics* 148.10 (2018), p. 104108. ISSN: 0021-9606. DOI: [10.1063/1.5017285](https://doi.org/10.1063/1.5017285).
- [13] E. Heinecke, D. Hartmann, R. Müller, and A. Hese. In: *The Journal of Chemical Physics* 109.3 (1998), pp. 906–911. DOI: <http://dx.doi.org/10.1063/1.476631>.
- [14] J. Köhler, A.C.J. Brouwer, E.J.J. Groenen, and J. Schmidt. In: *Chemical Physics Letters* 250.1 (1996), pp. 137–144. ISSN: 0009-2614. DOI: [http://dx.doi.org/10.1016/0009-2614\(95\)01393-8](http://dx.doi.org/10.1016/0009-2614(95)01393-8).
- [15] Stuart Bogatko et al. In: *The Journal of Physical Chemistry C* 120.15 (2016), pp. 8251–8260. ISSN: 1932-7447. DOI: [10.1021/acs.jpcc.6b00150](https://doi.org/10.1021/acs.jpcc.6b00150).

- 
- [16] Eugene S. Kadantsev, M. J. Stott, and Angel Rubio. In: *Journal of Chemical Physics* 124.13 (2006). ISSN: 00219606. DOI: [10.1063/1.2186999](https://doi.org/10.1063/1.2186999).
- [17] Stefan Grimme and Maja Parac. In: *ChemPhysChem* 4.3 (2003), pp. 292–295. ISSN: 1439-7641. DOI: [10.1002/cphc.200390047](https://doi.org/10.1002/cphc.200390047).

## 8 Conclusions

In this dissertation we have explored the possibilities for combining embedding methods and linear-scaling density functional theory. We began by looking at the theoretical foundations for both approaches, where the density matrix formalism in particular provides a useful bridge between the two methods. We also considered the practical side of it in Chapter 5 by studying the pentacene in *p*-terphenyl system from first principles with pre-existing electronic structure methods. The limitations for such systems due to the need for explicit host structures to be used to capture the true physics of the system, plus the size restrictions imposed by the hybrid functionals required to provide results of acceptable accuracy, provide solid practical foundations for a joint quantum embedding – linear-scaling DFT approach to doped molecular crystals.

As a preliminary to a full embedding implementation, in Chapter 6 we described a subsystem orbital optimisation strategy in the linear-scaling ONETEP package, with which we confirmed the validity of treating different portions of the system at different levels of accuracy with the ONETEP minimal basis set approach. From there in Chapter 7 we outlined the implementation of embedded mean-field theory (EMFT) in ONETEP, recasting the EMFT formalism in linear-scaling form and adapting the ONETEP optimisation algorithms to consider subsystems operating at different levels of theory. This approach was verified by testing a variety of systems with hybrid functionals embedded in semi-local environments, the largest containing nearly 3000 atoms to demonstrate the power of the method. We observed qualitative and quantitative differences in excitonic behaviour from the inclusion of a full periodic cell of crystalline *p*-terphenyl for the environment of a pentacene molecule compared to implicit solvent and a cluster of nearest neighbour molecules, confirming the scientific value of the quantum embedding method.

As future work, there are a variety of avenues which it is possible to explore with this new method. From an algorithmic perspective, the implementation of EMFT in ONETEP is a restrictive one, as it does not allow for optimisation of the basis functions at the higher level of theory in an embedding calculation, restricting the self-consistent optimisation algorithm to improvements in the density kernel only. Extending this could make the method much more flexible and applicable to a wider variety of systems. Furthermore, more investigations are required to understand the nature of the

breakdown in the self-consistent algorithm associated here with basis set optimisation and in past studies with large differences in exchange-correlation functionals. A deeper understanding of this phenomenon is required before any further adaptation to the ONETEP algorithm for EMFT can be carried out with confidence.

Aside from EMFT, the infrastructural work performed as part of this dissertation to make embedding functional in ONETEP is adaptable to other embedding strategies, such as Wesolowski-Warshel frozen density embedding and the Manby projector method. They have not been implemented yet due to the focus on EMFT, which bears a closer relation to ONETEP via the density matrix formalism, but in principle there is no reason why this cannot be carried out in the future. Computational benefits can be obtained using the subsystem NGWF optimisation routine, as there is now a clearly viable way to reduce the extent to which subsystem orbitals need to be optimised for large-scale calculations, potentially enabling new parallel strategies by optimising components associated with different subsystems separately in a divide-and-conquer fashion.

From the perspective of the room-temperature maser, a greater degree of flexibility is now available for the understanding of complex solute-solvent systems from an embedding perspective. Additional maser active molecules, and potentially host environments, could be tested as viable candidates for an improved maser system. Such approaches may require more sophisticated exchange-correlation functionals, such as optimally-tuned range-separated hybrids (OT-RSH) to be used in combination with time-dependent DFT (TDDFT). In principle both of these methods are compatible with EMFT and linear-scaling but combining them in ONETEP remains an open problem. The use of spin-orbit coupling to examine the triplet splittings would also represent an important future development in understanding the pentacene excited states and their behaviour in explicit host.

Serious consideration should be given to what other kinds of molecular systems may be suitable for this approach, as in this work we have restricted ourselves to a fairly narrow class of molecules and hosts. However there are potential applications for biological systems with large proteins where a small subsystem must be described with a high level of theory but the remainder of the structure is also important and so must be accounted for. Overall, making quantum embedding viable within the linear-scaling family has the potential to open up many new avenues of scientific inquiry.

# A Derivation of the Hartree-Fock Equations

Hartree-Fock Theory is a staple of the condensed matter physicist or quantum chemist. Here we provide the background for a conventional derivation of the Hartree-Fock equations, as a supplement to the embedded Hartree-Fock theory described in the main text. To begin, we will consider the derivation of the HF equations for the case of a closed set of  $N$  mutually orthogonal orbitals  $\{\phi_i\}$ , following the derivation by Lowe [1].

The Hamiltonian for this system is given by

$$\hat{H} = \hat{H}^{core} + \sum_{j=1}^N (2\hat{J}_j - \hat{K}_j), \quad (\text{A.1})$$

where  $\hat{H}^{core}$  is the one-electron Hamiltonian (kinetic plus electron-nuclear interaction),  $J_j$  is the Coulomb term and  $K_j$  is the exchange term:

$$\hat{J}_j |\phi_i\rangle \equiv \langle \phi_j | 1/r_{12} | \phi_j \rangle |\phi_i\rangle, \quad (\text{A.2})$$

$$\hat{K}_j |\phi_i\rangle \equiv \langle \phi_j | 1/r_{12} | \phi_i \rangle |\phi_j\rangle. \quad (\text{A.3})$$

where  $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ . To find the optimum set of orbitals, we must minimise the functional

$$L = \sum_i \langle \phi_i | \hat{H} | \phi_i \rangle = \sum_i \left( 2H_{ii} + \sum_j (2J_{ij} - K_{ij}) \right), \quad (\text{A.4})$$

where the double subscripts indicate matrix elements, with respect to the orbitals subject to the constraint that they remain orthonormal, yielding an equation of the form

$$\delta \left( L - 2 \sum_{i,j} \epsilon_{ij} \langle \phi_i | \phi_j \rangle \right) = \delta \left( L - 2 \sum_{i,j} \epsilon_{ij} S_{ij} \right) = 0, \quad (\text{A.5})$$



where  $\epsilon_{ij}$  are Lagrange multipliers. The various matrix elements then become

$$\delta S_{ij} = \langle \delta \phi_i | \phi_j \rangle + \langle \phi_i | \delta \phi_j \rangle, \quad (\text{A.6})$$

$$\begin{aligned} \delta H_{ii} &= \langle \delta \phi_i | \hat{H}^{core} | \phi_i \rangle + \langle \phi_i | \hat{H}^{core} | \delta \phi_i \rangle, \\ &= \langle \delta \phi_i | \hat{H}^{core} | \phi_i \rangle + c.c., \end{aligned} \quad (\text{A.7})$$

$$\begin{aligned} \delta J_{ij} &= \langle \delta \phi_i | \langle \phi_j | 1/r_{12} | \phi_j \rangle | \phi_i \rangle + \langle \phi_i | \langle \delta \phi_j | 1/r_{12} | \phi_j \rangle | \phi_i \rangle \\ &\quad + \langle \phi_i | \langle \phi_j | 1/r_{12} | \delta \phi_j \rangle | \phi_i \rangle + \langle \phi_i | \langle \phi_j | 1/r_{12} | \phi_j \rangle | \delta \phi_i \rangle, \\ &= \langle \delta \phi_i | \hat{J}_j | \phi_i \rangle + \langle \delta \phi_j | \hat{J}_i | \phi_j \rangle + c.c. \end{aligned} \quad (\text{A.8})$$

$$\begin{aligned} \delta K_{ij} &= \langle \delta \phi_i | \langle \phi_j | 1/r_{12} | \phi_i \rangle | \phi_j \rangle + \langle \phi_i | \langle \delta \phi_j | 1/r_{12} | \phi_i \rangle | \phi_j \rangle \\ &\quad + \langle \phi_i | \langle \phi_j | 1/r_{12} | \delta \phi_i \rangle | \phi_j \rangle + \langle \phi_i | \langle \phi_j | 1/r_{12} | \phi_i \rangle | \delta \phi_j \rangle, \\ &= \langle \delta \phi_i | \hat{K}_j | \phi_i \rangle + \langle \delta \phi_j | \hat{K}_i | \phi_j \rangle + c.c. \end{aligned} \quad (\text{A.9})$$

So

$$\delta J_{ij} = \langle \delta \phi_i | \hat{J}_j | \phi_i \rangle + \langle \delta \phi_j | \hat{J}_i | \phi_j \rangle + c.c. \quad (\text{A.10})$$

$$\delta K_{ij} = \langle \delta \phi_i | \hat{K}_j | \phi_i \rangle + \langle \delta \phi_j | \hat{K}_i | \phi_j \rangle + c.c., \quad (\text{A.11})$$

where  $c.c$  indicates the complex conjugate. Gathering it all together, our minimization equation becomes

$$2 \sum_i \langle \delta \phi_i | \left[ \left( \hat{H}^{core} + \sum_j \left( 2\hat{J}_j - \hat{K}_j \right) \right) | \phi_i \rangle - \sum_j \epsilon_{ij} | \phi_j \rangle \right] + c.c. = 0. \quad (\text{A.12})$$

The variations  $\delta \phi_i$  and  $\delta \phi_i^*$  are independent, so the two halves of (A.12) must independently be zero. In addition, these variations are arbitrary, so the part in square brackets in the first half must equal zero for any variation  $\delta \phi_i^*$ . Hence we obtain

$$\left[ \hat{H}^{core} + \sum_j \left( 2\hat{J}_j - \hat{K}_j \right) \right] | \phi_i \rangle = \sum_j \epsilon_{ij} | \phi_j \rangle. \quad (\text{A.13})$$

We now define the Fock operator as

$$\hat{F} \equiv \hat{H}^{core} + \sum_j \left( 2\hat{J}_j - \hat{K}_j \right), \quad (\text{A.14})$$

such that (A.13) becomes

$$\hat{F} | \phi_i \rangle = \sum_j \epsilon_{ij} | \phi_j \rangle. \quad (\text{A.15})$$

This equation is rather complicated as the solution for one orbital depends on all the

other orbitals, and so must be solved self-consistently. We can convert this to an eigenvalue problem by recasting (A.15) as a matrix equation:

$$\hat{F}\Phi = \Phi\mathbf{E}, \quad (\text{A.16})$$

where  $\Phi$  is the vector of orbitals  $\{\phi_i\}$  and  $\mathbf{E}$  is a matrix of Lagrange multipliers. Now let us post multiply each side of this equation by a unitary matrix  $U$  ( $U^{-1}U = U^\dagger U = \mathbf{1}$ ):

$$\hat{F}\Phi U = \Phi\mathbf{E}U, \quad (\text{A.17})$$

$$\begin{aligned} &= \Phi U U^\dagger \mathbf{E}U, \\ \Rightarrow \hat{F}\Phi' &= \Phi'\mathbf{E}', \end{aligned} \quad (\text{A.18})$$

where  $\Phi' = \Phi U$  and  $\mathbf{E}' = U^\dagger \mathbf{E}U$ . So far, we've achieved nothing since this is exactly the same form as (A.16), but now consider the Coulomb term in the Fock operator (following the derivation of Szabo and Ostlund [2]):

$$\sum_i J'_i = \sum_i \langle \phi'_i | 1/r_{12} | \phi'_i \rangle = \sum_i \sum_{k,l} \langle \phi_k | U_{ki}^* 1/r_{12} U_{li} | \phi_l \rangle = \sum_{k,l} \left[ \sum_i U_{ki}^* U_{li} \right] \langle \phi_k | 1/r_{12} | \phi_l \rangle \quad (\text{A.19})$$

and since

$$\sum_i U_{ki}^* U_{li} = (U U^\dagger)_{lk} = (\mathbf{1})_{lk} = \delta_{lk}, \quad (\text{A.20})$$

we have

$$\sum_i J'_i = \sum_k \langle \phi_k | 1/r_{12} | \phi_k \rangle = \sum_k J_k, \quad (\text{A.21})$$

i.e. a unitary transformation of the HF orbitals has no effect on the form of the Coulomb term in the Fock operator. In addition, the exchange term can be rewritten as

$$\sum_j \hat{K}'_j | \phi'_i \rangle = \sum_j \langle \phi'_j | 1/r_{12} | \phi'_i \rangle | \phi'_j \rangle, \quad (\text{A.22})$$

$$\begin{aligned} &= \sum_{k,l,m} \sum_j \langle \phi_k | U_{kj}^* 1/r_{12} U_{li} | \phi_l \rangle U_{mj} | \phi_m \rangle, \\ &= \sum_{k,l,m} \left[ \sum_j U_{kj}^* U_{mj} \right] U_{li} \langle \phi_k | 1/r_{12} | \phi_l \rangle | \phi_m \rangle, \\ &= \sum_{k,l} U_{li} \langle \phi_k | 1/r_{12} | \phi_l \rangle | \phi_k \rangle, \\ &= \sum_{k,l} U_{li} \hat{K}_k | \phi_l \rangle, \\ \Rightarrow \sum_j \hat{K}'_j | \phi'_i \rangle &= \sum_k \hat{K}_k | \phi'_i \rangle, \end{aligned} \quad (\text{A.23})$$

so the form of the exchange operator is unaffected by the unitary transformation.  $\hat{H}^{core}$  does not depend on the form of the HF orbitals and so remains unchanged. Thus we conclude that the Fock operator is unchanged by the unitary transformation of the HF orbitals i.e.  $\hat{F}' = \hat{F}$ .

All that remains, then, is to consider the Lagrange multipliers. We see from (A.15) that these are the elements of the Fock matrix:

$$\langle \phi_k | \hat{F} | \phi_i \rangle = \sum_j \epsilon_{ij} \langle \phi_k | \phi_j \rangle = \sum_j \epsilon_{ij} \delta_{jk} = \epsilon_{ik}. \quad (\text{A.24})$$

Therefore,

$$\epsilon'_{ij} = \langle \phi'_j | \hat{F}' | \phi'_i \rangle, \quad (\text{A.25})$$

$$\begin{aligned} &= \sum_{k,l} \langle \phi_k | U_{kj}^* \hat{F} U_{li} | \phi_l \rangle, \\ &= \sum_{k,l} U_{kj}^* \langle \phi_k | \hat{F} | \phi_l \rangle U_{li}, \\ &= \sum_{k,l} U_{kj}^* \epsilon_{kl} U_{li}, \end{aligned} \quad (\text{A.26})$$

which implies that  $\mathbf{E}' = \mathbf{U}^\dagger \mathbf{E} \mathbf{U}$ . Since  $\mathbf{E}$  is Hermitian, we can always choose a matrix  $\mathbf{U}$  such that  $\mathbf{E}'$  is diagonal [3]. Thus there exists a set of orbitals such that

$$\hat{F}' | \phi'_i \rangle = \epsilon'_{ii} | \phi'_i \rangle, \quad (\text{A.27})$$

which we can rewrite as

$$\hat{F} | \phi_i \rangle = \epsilon_i | \phi_i \rangle. \quad (\text{A.28})$$

These are the Hartree-Fock equations. Physically, the unitary transformation has no effect; it is simply a mathematical convenience.

## B Derivation of the Embedded Hartree-Fock Equations

To begin, we partition the orbitals of the system into two sets,  $\{\phi_i^A\}$  and  $\{\phi_i^B\}$ , where the former make up the active subsystem and the latter form the stationary environment. For illustrative purposes, we will first consider the case where the A and B orbitals are mutually orthogonal.

The core Hamiltonian term in [A](#) can be split into a sum of A and B orbitals, the latter of which disappears in the variation of the orbital space:

$$\delta H_{ii} = \langle \delta \phi_i^A | \hat{H}^{core,A} | \phi_i^A \rangle + c.c., \quad (\text{B.1})$$

Here we have given the operator the superscript A to make explicit the fact that it only acts on the A orbitals.

The Coulomb, overlap and exchange matrices can be divided into three parts: interaction between two A orbitals, interaction between two B orbitals and interaction between A and B orbitals. Since the B orbitals remain stationary, we need only consider the variation with respect to the A orbitals. First, the case where  $i \in A$  and  $j \in B$ :

$$\delta S_{ij} = \langle \delta \phi_i^A | \phi_j^B \rangle, \quad (\text{B.2})$$

$$\delta J_{ij} = \langle \delta \phi_i^A | \hat{J}_j^B | \phi_i^A \rangle + c.c. \quad (\text{B.3})$$

$$\delta K_{ij} = \langle \delta \phi_i^A | \hat{K}_j^B | \phi_i^A \rangle + c.c. \quad (\text{B.4})$$

where  $J_j^B | \phi_i^A \rangle = \langle \phi_j^B | 1/r_{12} | \phi_j^B \rangle | \phi_i^A \rangle$  and  $K_j^B | \phi_i^A \rangle = \langle \phi_j^B | 1/r_{12} | \phi_i^A \rangle | \phi_j^B \rangle$ . Similar results can be derived for the opposite case ( $i \in B$  and  $j \in A$ ).

For  $i, j \in A$ , the results we obtain are equivalent to those in the case in [Appendix A](#):

$$\delta S_{ij} = \langle \delta \phi_i^A | \phi_j^A \rangle + \langle \phi_i^A | \delta \phi_j^A \rangle, \quad (\text{B.5})$$

$$\delta J_{ij} = \langle \delta \phi_i^A | \hat{J}_j^A | \phi_i^A \rangle + \langle \delta \phi_j^A | \hat{J}_i^A | \phi_j^A \rangle + c.c. \quad (\text{B.6})$$

$$\delta K_{ij} = \langle \delta \phi_i^A | \hat{K}_j^A | \phi_i^A \rangle + \langle \delta \phi_j^A | \hat{K}_i^A | \phi_j^A \rangle + c.c. \quad (\text{B.7})$$

Combining these results, our minimisation condition becomes:

$$2 \sum_{i \in A} \langle \delta \phi_i^A | \left[ \left( \hat{H}^{core,A} + \sum_{j \in A} (2\hat{J}_j^A - \hat{K}_j^A) + \sum_{j \in B} (2\hat{J}_j^B - \hat{K}_j^B) \right) | \phi_i^A \rangle \right] - 2 \sum_{i \in A} \langle \delta \phi_i^A | \left[ \sum_{j \in A} \epsilon_{ij} | \phi_j^A \rangle + \sum_{j \in B} \epsilon_{ij} | \phi_j^B \rangle \right] + c.c. = 0. \quad (\text{B.8})$$

Following the same argument as before, we obtain

$$\left[ \hat{H}^{core,A} + \sum_{j \in A} (2\hat{J}_j^A - \hat{K}_j^A) + \sum_{j \in B} (2\hat{J}_j^B - \hat{K}_j^B) \right] | \phi_i^A \rangle = \sum_{j \in A} \epsilon_{ij} | \phi_j^A \rangle + \sum_{j \in B} \epsilon_{ij} | \phi_j^B \rangle \quad (\text{B.9})$$

Now defining the Fock operator for the whole system to be the quantity in square brackets, we see that the Lagrange multipliers for the B states are given by

$$\epsilon_{ij} = \langle \phi_j^B | \hat{F} | \phi_i^A \rangle, \quad (\text{B.10})$$

so (B.9) can be rewritten as

$$\hat{F} | \phi_i^A \rangle = \sum_{j \in A} \epsilon_{ij} | \phi_j^A \rangle + \sum_{j \in B} \langle \phi_j^B | \hat{F} | \phi_i^A \rangle | \phi_j^B \rangle, \quad (\text{B.11})$$

$$= \sum_{j \in A} \epsilon_{ij} | \phi_j^A \rangle + \hat{P}^B \hat{F} | \phi_i^A \rangle, \quad (\text{B.12})$$

where we have introduced the projection operator

$$\hat{P}^B \equiv \sum_{j \in B} | \phi_j^B \rangle \langle \phi_j^B |. \quad (\text{B.13})$$

Finally, we obtain a new form for the HF equation in the embedded regime:

$$(1 - \hat{P}^B) \hat{F} | \phi_i^A \rangle = \sum_{j \in A} \epsilon_{ij} | \phi_j^A \rangle. \quad (\text{B.14})$$

A unitary transformation can be performed to cast (B.14) as an eigenvalue equation:

$$(1 - \hat{P}^B) \hat{F} | \phi_i^A \rangle = \epsilon_i | \phi_i^A \rangle. \quad (\text{B.15})$$

## B.1 Generalised Embedded Hartree-Fock Equations

We now consider the general case of a set of orbitals  $\{\phi_i^A\}$  that is not orthogonal to the fixed set of orbitals  $\{\phi_i^B\}$ . A result identical to (B.9) can be derived for a set of orbitals  $\{\tilde{\phi}_i^A\}$  that have been orthogonalised to the  $B$  orbitals by means of the Gram-Schmidt approach, following the example of Phillips and Kleinman in their pseudopotential approach [4]

$$|\tilde{\phi}_i^A\rangle = (1 - \hat{P}^B) |\phi_i^A\rangle, \quad (\text{B.16})$$

where the projection operator is defined as

$$\hat{P}^B = \sum_{j \in B} |\phi_j^B\rangle \langle \phi_j^B|. \quad (\text{B.17})$$

Then the generalised version of (B.9) is

$$\left[ \hat{H}^{core,A} + \sum_{j \in A} (2\tilde{J}_j^A - \tilde{K}_j^A) + \sum_{j \in B} (2\hat{J}_j^B - \hat{K}_j^B) \right] |\tilde{\phi}_i^A\rangle = \sum_{j \in A} \epsilon_{ij} |\tilde{\phi}_j^A\rangle + \sum_{j \in B} \epsilon_{ij} |\phi_j^B\rangle, \quad (\text{B.18})$$

where the new Coulomb and exchange operators are defined as

$$\tilde{J}_j^A |\tilde{\phi}_i^A\rangle \equiv \langle \tilde{\phi}_j^A | 1/r_{12} | \tilde{\phi}_j^A \rangle |\tilde{\phi}_i^A\rangle, \quad (\text{B.19})$$

$$\tilde{K}_j^A |\tilde{\phi}_i^A\rangle \equiv \langle \tilde{\phi}_j^A | 1/r_{12} | \tilde{\phi}_i^A \rangle |\tilde{\phi}_j^A\rangle, \quad (\text{B.20})$$

where we have included a tilde to indicate that these operators are formed from projected orbitals. Defining the quantity in square parentheses in (B.18) to be the Fock operator  $\tilde{F}$ , we get

$$\tilde{F} |\tilde{\phi}_i^A\rangle = \sum_{j \in A} \epsilon_{ij} |\tilde{\phi}_j^A\rangle + \sum_{j \in B} \epsilon_{ij} |\phi_j^B\rangle. \quad (\text{B.21})$$

Just as before, we can express the Lagrange multipliers in the second sum as elements of the Fock matrix, reducing the above equation to

$$(1 - \hat{P}^B) \tilde{F} |\tilde{\phi}_i^A\rangle = \sum_{j \in A} \epsilon_{ij} |\tilde{\phi}_j^A\rangle, \quad (\text{B.22})$$

which is identical to the result in (B.14).

Performing our unitary transformation, the result is

$$(1 - \hat{P}^B) \tilde{F} |\tilde{\phi}_i^A\rangle = \epsilon_i |\tilde{\phi}_i^A\rangle, \quad (\text{B.23})$$

and substituting the form for  $|\tilde{\phi}_i^A\rangle$  into our HF equation, we get

$$\left(1 - \hat{P}^B\right) \tilde{F} \left(1 - \hat{P}^B\right) |\phi_i^A\rangle = \epsilon_i \left(1 - \hat{P}^B\right) |\phi_i^A\rangle. \quad (\text{B.24})$$

This can be rewritten as

$$\left(\tilde{F} + \hat{V}^{WR}\right) |\phi_i^A\rangle = \epsilon_i |\phi_i^A\rangle, \quad (\text{B.25})$$

where we have introduced the Weeks-Rice style pseudopotential [5],

$$\hat{V}^{WR} \equiv -\hat{P}^B \tilde{F} - \tilde{F} \hat{P}^B + \hat{P}^B \tilde{F} \hat{P}^B + \epsilon_i \hat{P}^B. \quad (\text{B.26})$$

This completes our derivation of the embedded Hartree-Fock equations.

## C Embedding data structures in ONETEP

As part of the implementation of freeze-and-thaw embedding (F+T) and embedded mean-field theory (EMFT) in ONETEP, a significant amount of changes were made to the underlying data structures to handle these features. Fundamentally, this reflects the fact that, prior to this work, the ONETEP algorithm was designed specifically to work with a single quantum system described uniformly at a single level of theory. Other forms of embedding, such as implicit solvation and QM/MM, are already available but can be considered as external to the quantum density-functional theory (DFT) part of the calculation. Implementing quantum embedding required substantial changes to the underlying parallel strategies, sparse matrix structures and other data structures that are pervasive throughout the code, comprising a significant part of the work discussed in this thesis. In this appendix we briefly discuss the most important data structure changes that have been implemented and are now available within the academic release version of ONETEP. A detailed description of how to run embedding calculations in ONETEP is available with the ONETEP documentation provided with academic version 5.2 and later

### C.1 New data types

- **SPAM3\_EMBED** — modified version of traditional SPAM3 sparse matrix type to allow for subsystem matrices.

```

type, public :: SPAM3_EMBED
    integer :: mrows
    integer :: ncols
    type(SPAM3), allocatable :: m(:, :) ! Array of SPAM3's
    character(len=30) :: structure
    logical :: iscmplx ! TRUE if matrix is complex
    type(SPAM3), pointer :: p ! Pointer to %m(1,1)
end type SPAM3_EMBED

```



This contains a two-dimensional array of SPAM3 matrices `SPAM3_EMBED%m`, of size `(mrows × ncols)`, the elements of which are the subsystem matrices. For embedding with two subsystems, the structure contains four matrices stored in the form `SPAM3_EMBED%m(1:2, 1:2)`, with matching row and column indices referring to diagonal matrices and alternating indices being off-diagonal terms, while for a non-embedding calculation with one subsystem, there is only a single matrix accessed via `SPAM3_EMBED%m(1, 1)`. Each SPAM3 has its own unique structure code, composed of the `SPAM3_EMBED` structure and a subsystem identifier e.g. the structure code for the Hamiltonian of subsystem 1 would be `H11`. The pointer `SPAM3_EMBED%p` is a shorthand for parts of the code that have not yet been interfaced with the new embedding structures, acting as a pointer direct to `SPAM3_EMBED%m(1, 1)`.

- **SPAM3\_EMBED\_ARRAY** — modified version of `SPAM3_ARRAY`, a generalised version of `SPAM3` that allows for the use of multiple spin channels and k-points.

```

type, public :: SPAM3_EMBED_ARRAY
    integer :: mrows
    integer :: ncols
    integer :: num_spins
    integer :: num_kpoints
    type(SPAM3_EMBED), allocatable :: m(:, :)
    character(len=30) :: structure
end type SPAM3_EMBED_ARRAY

```

`SPAM3_EMBED_ARRAY%m` is a `(num_spins × num_kpoints)` array of `SPAM3_EMBED` matrices, each of which are themselves composed of `(mrows × ncols)` `SPAM3` matrices. Accessing matrices associated with different spin channels and k-points is unchanged compared to `SPAM3_ARRAY`. For example, with two spin channels and one k-point, there are two `SPAM3_EMBED` matrices available in the format `SPAM3_EMBED_ARRAY%m(1:2, 1)`, while subsystem matrices are stored within each `SPAM3_EMBED` in the same manner as discussed above. The embedding code has been verified to work with multiple spin channels but not with multiple k-points.

- **REGION** — new type within `MODEL`, containing all information specifically related to the subsystem, such as `REGION%ELEMENT`, the elements list which contains information relating to the atoms in the subsystem. The model contains `MODEL%nsb` regions, the data for which are stored as `MODEL%REGION(1:nsb)`. All structures contained within `REGION` are also still present in `MODEL`, but have

been altered to refer exclusively to the whole system. So for example for two subsystems, `MODEL%ELEMENT` is an amalgamation of the subsystem element lists `MODEL%REGION(1)%ELEMENT` and `MODEL%REGION(2)%ELEMENT`. For one subsystem, the regional quantities are merely pointers to the full system versions.

## C.2 New module

- **sparse\_embed\_mod.F90** — module designed for manipulation of `SPAM3_EMBED` and `SPAM3_EMBED_ARRAY` types so that the subsystem structures can be passed on to `sparse_mod.F90`. For example, the subroutine `sparse_embed_product` takes two input matrices of type `SPAM3_EMBED`, `amat` and `bmat`, and returns the output matrix `cmat`. The relevant code for this operation is:

```
!=====
do jsub=1,ncols
  do isub=1,mrows
    ! Sum up all contributions from subsystem products
    do ksub=1,nsup
      call sparse_product(cmat%m(isub,jsub), &
        amat%m(isub,ksub), bmat%m(ksub,jsub))
      call sparse_axpy(tmp_mat%m(isub,jsub), &
        cmat%m(isub,jsub),1.0_DP)
    end do
  end do
end do
! Copy tmp_mat back to cmat
call sparse_embed_copy(cmat,tmp_mat)
!=====
```

The three `do...enddo` loops cycle over all subsystem matrices, passing the relevant matrices to the subroutine `sparse_product` for the sparse matrix multiplication. The cumulative result is stored in the local variable `tmp_mat`, calculated by adding the product matrices together using `sparse_axpy`. Finally the resulting matrix is copied back to `cmat` and returned. If there is only one subsystem in the calculation, then the relevant matrix structures are passed directly to `sparse_mod.F90`.

### C.3 Major datatype modifications

- The local (pseudopotential), Hartree and exchange-correlation potentials are collectively represented on the ONETEP fine grid as a single quantity, `lhxc_fine`. For EMFT, this is now an array consisting of the total potentials required for each subsystem separately, `lhxc_fine(1:nsub)` where `nsub` is the number of subsystems, such that the high- and low-level calculations can be performed with different exchange-correlation potentials. The low-level term is thus the uncorrected full potential, while the active region quantity contains the final, corrected EMFT potential as discussed in Chapter 7.
- Each subregion is allocated its own basis of NGWFs, with `ngwf_basis(1:nsub)` now being an array of the type `FUNC_BASIS` with length equal to the number of regions `nsub`. For the case of one subsystem, the data is then accessed simply via an array of size 1. Similarly, the `ngwfs_on_grid(1:nsub)` quantity is an array of type `FUNCTIONS` and length `nsub`, containing each subsystem's NGWF data as distributed on the grid, which are parallelised for each region separately across the MPI framework (see below). This also applies to related `FUNC_BASIS` quantities such as non-local projectors (`nl_projectors(1:nsub)`) and the joint conduction-valence NGWF basis (`joint_ngwf_basis(1:nsub)`).
- High-level data structures for quantities such as the density kernel (`DKERN`), Hamiltonian (`NGWF_HAM`) and overlap matrices (`NGWF_REP`) which contain `SPAM3` or `SPAM3_ARRAY` matrices have been replaced with the types `SPAM3_EMBED` and `SPAM3_EMBED_ARRAY`. Conventions for accessing and manipulating these matrices are the same as prior to the embedding implementation. This facilitates calculation of total energies and forces, but has not been done for more specialised features such as couplings and DFT with Hubbard U, which still use the non-embedding structures.

### C.4 Parallel strategies

In a normal ONETEP calculation, atoms are distributed across the available MPI processes according to a 'parallel strategy'. This determines how resources such as matrix elements will be spread across the MPI environment in order to reduce the communication between nodes and maximise the efficiency of the calculation. Details on maximising parallel efficiency are available via the ONETEP documentation and website.

As part of the embedding infrastructure, each subsystem is given its own parallel strategy, contained in the quantity `MODEL%REGION%par`. This contains all information

relating to the distribution of resources across the MPI nodes available to the calculation, which are determined by the parameter `PARALLEL_SCHEME`. There are three settings for the distribution of resources during an embedding calculation:

- **NONE:** All subsystems are treated completely independently, with atoms distributed across all available processors as though the other subsystems do not exist. The number of MPI processes cannot be greater than the number of atoms in the smallest subsystem. For example, if there are 8 processors available then each will hold atoms and data from all subsystems, though the calculation will fail if any subsystem has less than 8 atoms (or possibly slightly more if the space-filling curve is in use). This is the default setting for testing but is not recommended for practical calculations due to the constraint on the number of processors.
- **SENATE:** Nodes are partitioned evenly between all subsystems. For example, if there are 8 processors and 2 subsystems, then each will be allocated 4 processors, regardless of the number of atoms in each subsystem. Unlike the **NONE** setting, there is no upper bound on the number of processors which may be used, so user discretion is advised.
- **HOUSE:** Divides the processors proportionally between all subsystems, with a minimum of 1 processor per subsystem. For example, if we have two subsystems consisting of 15 and 5 atoms each, then with 8 processors each subsystem will be allocated 6 and 2 nodes respectively. At a minimum all subsystems are granted 1 processor — if we had two subsystems with 1 and 100 atoms in our 8 processor example, then they will receive 1 and 7 processor respectively. Like **SENATE**, there is no upper bound on the number of processors that can be allocated and finding a sensible setting is left to the user. This is the recommended setting for running calculations, the others are mainly of use for testing.

In addition to MPI parallelisation, ONETEP utilises open multi-processing (OpenMP) for memory sharing via multiple threads across a single node. The OpenMP parallelisation has not been affected by the implementation of embedding.

### C.4.1 Data structure changes for parallel strategies

In addition to the creation of regional parallel strategies via `MODEL%REGION%par`, there are several other ways in which the handling of parallel strategies within ONETEP have been altered.

- Deprecation of the global variable `pub_par`. Previously this was used throughout the code to extract parallel strategy information, which is unsuitable for using

multiple such strategies. Only used in parts of the code that are not yet compatible with embedding.

- `MODEL%par` is now a pointer to `MODEL%regions(1)%par`. Like `pub_par`, this is now used only in modules that have not been integrated with embedding structures.
- Individual sparse matrices hold pointers to two parallel strategies, `SPAM3%row_par` and `SPAM3%col_par`, for the strategies used to construct the elements along the rows and columns of the sparse matrix. For diagonal matrices, these are the same.

## C.5 Embedding algorithm

- For freeze-and-thaw embedding, the NGWF gradient is calculated in blocks for each region's NGWFs, such that for the frozen NGWFs there is no need to calculate the gradient coefficients at all. The EMFT NGWF gradient depends on the Hamiltonian operator corresponding to each region (see Chapter 7). This is implemented via selective application of the `lhxc_fine` potential for each block of regional NGWFs.
- Exchange-correlation calculations with EMFT require the use of multiple types of exchange-correlation functionals. For semi-local functionals, this involves a straightforward change of the functional to be used to generate the potential. Hybrid calculations require initialisation of the spherical wave resolution of identity (SWRI) structures. With embedding, the SWRI process can be activated for a specified active subsystem for hybrid-in-semi-local embedding. This required few changes to underlying data structures, since the active subsystem quantities can be passed to `hf_exchange_mod.F90` without reference to the full system. Currently there is no facility for off-diagonal terms between subsystems to be calculated with exact exchange, limiting embedding to hybrids in a semi-local environment.