UNIVERSITY OF NOTTINGHAM

MASTER OF SCIENCE BY RESEARCH THESIS

THEORETICAL PHYSICS

A Hückel Method Application for Studying the Self-Assembly of Phthalocyanine Monolayers

Author
Connor J.B. Fields

Supervisors
Dr. Jannette Dunn
Prof. Philip Moriarty

August 18, 2020



University of Nottingham
UK | CHINA | MALAYSIA

Acknowledgements

Firstly I would like to thank my supervisors Jannette Dunn and Philip Moriarty. Jannette for her support and guidance, which was was exceptional though out the year. Philip I would like to thank for the support during the resubmission process, which was again excellent. I feel personal very fortunate to have had such excellent supervision during my time studying at the University of Nottingham.

I am also incredibly grateful for the support from my family, particularly my parents, as without the unwavering support they have given me, I certainly would not be in the position I am today. My girlfriend has also been incredible throughout my time creating this work, helping me to remain positive and feel like I can achieve what I wanted to with this work.

Abstract

The modern day applications of quantum chemistry offer a vast scope of choices and tools available to researchers working in that field. These ideas range from techniques working close to the quantum fundamentals, i.e. performing wave-function calculations, to those taking purely classical perspective in which molecular properties might be deduced from observing the potential of the system [1, 2, 3, 4]. A relatively old and often overlooked idea for this type of work is the so-called Hückel method. This thesis will argue that under the right conditions this concept is a legitimate and powerful tool that has the potential for studying the behaviour of molecular systems. The idea for this research derives from a paper published by J. Leaf *et al* in 2016 [5], where the Hückel method was used to calculate and investigate the electronic properties of buckminsterfullerene molecules (C_{60}). This work follows on from that research by performing similar Hückel based calculations on a variety of different phthalocyanine molecules. A significant part of the work described in this thesis is devoted to examining how a Hückel method simulation can be applied to aid the interpretation of scanning tunnelling microscope (STM) images. Recognizing this, detailed images will be presented as part of the discussion of the molecular assembly to clarify detail, provide a visualisation of the effects being presented and give insight into the STM images viewed in experiments. Despite the relative simplicity, and, importantly, very low computational cost of the Hückel method (particularly when compared to, density functional theory calculations, for example), the program shows that it can produce simulated data that not only replicate the real-world STM results but predict novel molecular monolayer structures. In this thesis, two key types of molecule, and molecular assembly, are investigated using the Hückel method: benzene (as a prototyping tool) and a collection of phthalocyanine molecules. The adsorption of the latter on a variety of surfaces is explored, using experimental data as a check on the program where it is available. Using these tools, accurate simulations of titanyl phthalocyanine (TiOPc) on a gold surface were obtained, according to the results published by Mannsfeld and Fritz in 2005 [6]. An investigation of the effect of changing the molecular super-lattice structure was also performed showing a small linear change between the minimum molecular angle and inter-molecular distance. The simulations further predicted six angular phases in a comprehensive assessment of each phthalocyanine on three different surfaces, in three different overlayer structures. Included in this discussion is a qualitative comparison with STM images published by Bobaru and Salomon *et al* in 2011 [7]. These again show excellent agreement with simulations performed in this research. Hence demonstrating the validity of these simulation techniques.

Contents

1 Introduction **5**

2 Background **7**
2.1 The fundemental Concepts of Scanning Tunnelling Microscopy . . 7
2.2 Properties of Phthalocyanine 10
2.3 Background of the Hückel Molecular Orbital Theory 11

3 Theory **14**
3.1 Princibles of Simulating Scanning Tunnelling Microscopy 14
3.1.1 Bardeen's Tunnelling Theory 14
3.1.2 Transfer Matrix Elements 16
3.2 Molecular Orbital Calculations 17
3.2.1 Hartree-Fock Method 18
3.2.2 Density Functional Theory 21
3.2.3 Basic Hückel Method 23
3.2.4 Extended Hückel Theory 25
3.2.5 Self-Consistent Extended Hückel Theory 26
3.2.6 Empirical Methods 27
3.3 Orbital Basis 29
3.3.1 Hydrogen-like orbitals 29
3.3.2 Slater-Type Orbitals 30
3.3.3 Gaussian Type Orbitals 31
3.4 Gaussian Type Orbital Manipulations 32
3.4.1 Normalisation of Gaussian Type Orbitals 32
3.4.2 Gaussian Product theory 34
3.4.3 Overlap Integral Calculations 35
3.5 Calculation of Intermolecular Interaction Grids 37
3.6 Orientation Minimization Methods 38
3.6.1 Iterative Minimisation Method 38
3.6.2 Monte Carlo and Metropolis Monte Carlo 39

4 Application the Simulation Program to Benzene **42**
4.1 Molecular Orbital Calculations using the Basic Hückel Method . . 42
4.1.1 Stage 1: Labelling each atom 42
4.1.2 Stage 2: Constructing the Secular Equation 43
4.1.3 Stage 3: Solving the Secular Equation 44
4.1.4 Stage 4: Construction of Molecular Orbitals 45
4.2 Application of the Extended H"uckel Theory to Bezene 45
4.3 Benzene in a Scanning Tunnelling Microscopy Simulation 48

4.3.1 Constant Current vs Constant Height simulation 49
4.4 Simulating 2D Molecular Assemblies 49

5 Simularing Titanyl Phthalocyanine on a Gold Surface 53
5.1 Calculation of Titanyl Phalocyanines Molecular Orbital's 54
5.1.1 Comparison to Experimental Scanning Tunnelling Microscopy Images 56
5.2 Pthalocyanine Grid-Files 57
5.3 Creating a Surface Model 58
5.3.1 Surface Orbital Basis Set 60
5.4 Surface Grid-Files 61
5.5 Surface Calibration 63
5.6 TiOPc Simulation 65
5.6.1 Treatment of errors 67
5.7 Height Dependence 68

6 Comparison of Various Metal Phthalocyanines 71
6.1 Comparison of Different Central Atoms 71
6.1.1 Comparison of the Scanning Tunnelling Microscopy Simulations of Metal Phthalocyanines to Experimental Images . . 74
6.2 Comparison of Intermolecular Spacing 76
6.3 Comparison of Various Super-Lattice Structures 78
6.4 Comparison of Surfaces 84

7 Conclusion 94

Appendix A 97

Appendix B 100

Appendix C 105

Appendix D 107

Bibliography 113

Chapter 1

Introduction

Molecular behaviour is primarily governed by the role of the electron, the particle that essentially underpins all of chemistry. Unfortunately, although modelling electronic behaviour is vital for our understanding of many fundamental processes, a complete mathematical description of all but the simplest electron systems is essentially impossible due to the lack of an exact solution to the Schrödinger equation. This means that approximations (of varying levels of severity) must be adopted to study even rudimentary atomic and molecular arrangements.

With those concepts in mind, the goal of this study is to present an understanding of a system of phthalocyanine molecules from the perspective that arises when utilizing a semi-empirical approach, namely, the Hückel method [8]. Semi-empirical methods aim to tread the middle ground between *ab initio* simulations [3, 4] and empirical analyses [1], by maintaining a quantum mechanical basis but using some empirical parameters to simplify the calculation process [9]. In the past two decades, in particular, the Hückel method has largely been overlooked and considered out of date in the advent new and improved *ab initio* ideas. In this thesis, it is demonstrated that the Hückel method can be an effective (and computationally inexpensive) method for simulating and predicting the electronic structure of systems. In particular, it is shown that simulations of relatively complex molecules (phthalocyanines) on solid surfaces can compete favourably with full-blown *ab initio* approaches such as density functional theory [10].

A key focus throughout this research was to explore how these techniques could be used to simulate and interpret scanning tunnelling microscope (STM) images of the molecular assemblies. Images produced with an STM are fundamentally related to the structures of the electron orbitals being examined and therefore can be used to clarify and understand the images observed in experiments. In this context, the phthalocyanine family of molecules is especially well-suited to the Hückel method approach due to their planar structure, their relative structural complexity (complex, but not intractable for Hückel methods), and the large number of STM studies of 2D phthalocyanine assemblies on surfaces [6, 7, 11, 12, 13, 14, 15].

The structure of the thesis is as follows. Chapters one to three introduce the Hückel method, phthalocyanine and the features of modelling STM. These chapters also discuss the simulation algorithms and attempt to contextualise the Hückel method by discussing alternate modelling techniques. These introductory chapters are followed by an example of the Hückel method being applied to a simple prototypical molecule, benzene, in chapter four. Chapter five then extends this Hückel approach to the titanyl phthalocyanine molecule (TiOPc) and makes a comparison with the predictions of the simulations to STM measurements of TiOPc on a Au(111) surface [6]. Subsequently, in the final chapter of the main body, a variety of phthalocyanine molecules, namely: cobalt phthalocyanine (CoPc), copper phthalocyanine (CuPc), iron phthalocyanine (FePc) and tin phthalocyanine (SnPc) are simulated with the Hückel method and their self-assembly properties on various surfaces are also predicted.

Chapter 2

Background

2.1 The Fundamental Concepts of Scanning Tunnelling Microscopy

Given that one of the primary objectives of this work is to produce simulated STM images, it is first important to outline the key principles of scanning tunnelling microscopy. STM is a combined microscopy and spectroscopy technique that is used universally throughout condensed matter physics and nanoscience because it provides the ability to image, spectroscopically probe, and manipulate surfaces down to the atomic and (sub)molecular level [16, 17]. The development of these techniques originated from researchers inventing ways to utilize particular aspects of the phenomenon of quantum mechanics known as quantum tunnelling.

As the tunnelling phenomenon fundamentally underpins STM it is worth briefly considering where the effect originates. Tunnelling is a process by which quantum scale particles (electrons are the most commonly observed tunnelling particle but other sub-atomic and even atomic scale particles have also been observed [18]) can pass or 'tunnel' through a potential barrier at an energy level that would otherwise be impossible for classical particles. The concept of tunnelling arose out of a need to explain phenomena such as alpha particle radiation and field electron emission that were inexplicable using a classical model of particles [19]. Coinciding with these unexplained phenomena, the advent of quantum mechanics successfully lead to the prediction of a tunnelling mechanism when German physicist Friedrich Hund solved the Schrödinger equation for a double-well potential in 1927 [20]. Predictions of the tunnelling process are a direct result of the Heisenberg uncertainty principle and the wave-particle duality of electrons because these effects produce a non-zero probability of an electron (or any other quantum particle) existing beyond a finite potential barrier [21]. The derivation of this mechanism was a significant feat in the development of quantum mechanics as it was an early example of these new radical ideas being necessary to explain observable phenomena.

A mathematical representation of the wavefunction in a one dimension tunnelling example can be seen in figure 2.1. There are three regions shown in figure 2.1, the first representing the particle propagating in a region before the potential barrier,

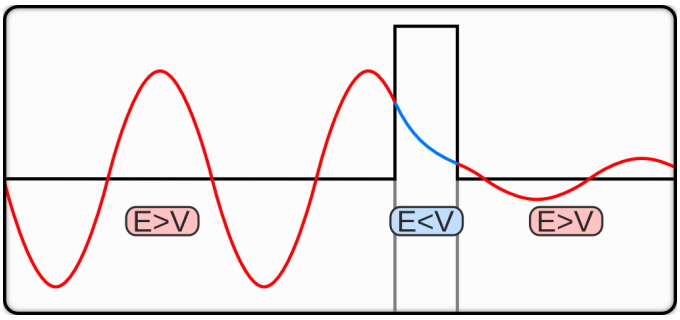


Figure 2.1: A representation of a wavefunction of energy E as it propagates through a barrier into a classically forbidden region. V represents the size of the potential barrier in each region. The size of these values are portrayed by the amplitude of the wavefunction and height of the barrier respectively.

the second showing the particles wavefunction inside the barrier and the third region displays the resultant wavefunction of the particle after passing through the barrier. The most important consideration, in terms of STM, is what happens to the wavefunction inside the potential barrier. For the one dimensional case, described in figure 2.1, the wavefunctions decays exponentially according to [22]:

$$\psi(z) = \psi(0)e^{-\kappa z} \tag{2.1}$$

This simplified tunnelling case describes a wavefunction $\psi(z)$ being initially equal to $\psi(0)$ at the point $z = 0$. The zero point in this instance indicating the boundary between the medium in which the wavefunction is propagating though and the potential barrier. The wavefunction $\psi(z)$ then decays at an exponential rate as it progresses through the barrier (as z increases). The exact rate of this decay is governed by the magnitude of κ, the decay rate. In the model used for the tunnelling problem in this instance, the potential energy barrier is modelled as a one-dimensional rectangle of limited height, corresponding to a decay rate of $\kappa = \frac{\sqrt{2m\Phi}}{\hbar}$, where m represents the mass of the tunnelling object (for STM this is an electron) and Φ is equal to the height of the barrier [22]. The key feature of equation (2.1), in terms of STM, is that the decay is exponentially dependent on the distance z. If equation (2.1) is used to represent a primitive description of the tunnelling between an STM tip and a surface, it is apparent that the probability of tunnelling is incredibly sensitive to the distance between the tip and the sample. Devising STM was a case of developing techniques to measure the tunnelling events on a surface accurately enough to be able to render a topographical image with atomic level resolution.

In the scanning tunnelling microscope, the potential energy barrier arises from the separation between the tip and the sample. Classically this potential barrier is the energy needed for the electron to 'jump' across the gap created by this separation. Figure 2.2 represents an idealised version of what an STM tunnelling process could look like, with the figure displaying a current transferring between an atomically sharp tip and the surface of a sample. The resolution of this system is determined by how sensitive the tunnelling mechanism is and by the amount of tunnelling detected from sites lateral to the central tip position. The tip apex

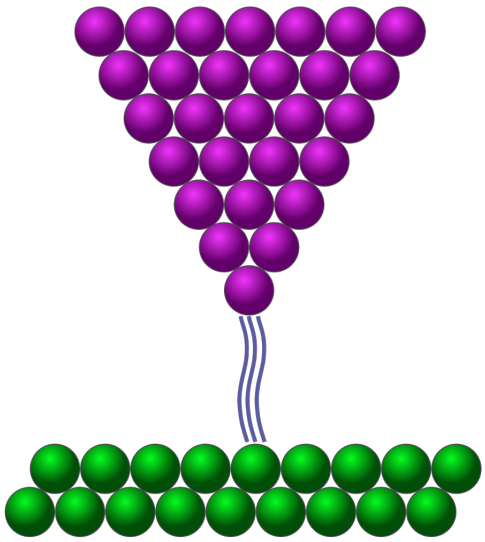


Figure 2.2: A depicting of an idealised version of an atomically sharp STM tip above a surface. The image shows a theoretical idealised version of an STM tip but there are methods by which this is achievable. A common technique involves chemical etching a hard substance such as tungsten or a platinum-iridium alloy [23, 24].

is, therefore, the most important resolution-defining element of the STM set-up and scanning probe microscopists spend a great deal of their time generating atomically sharp tips. A precise "universal" definition of this spatial resolution is difficult to determine as it depends on a significant number of factors related to both the tip and sample properties. However, an approximate model can be derived by assuming that tunnelling is a one-dimensional process and that the STM tip comprises of a single s-type orbital. The procedure from that initial point is to solve for the transmission wavefunction and then use that value to deduce information about the current [25]. Following through such a derivation results in the equation:

$$I(\Delta x) \propto e^{-2\kappa \frac{\Delta x^2}{2R}} \tag{2.2}$$

In this equation, the current I is a function of the lateral spacing Δx and is proportional to an exponential raised to the power of $-2\kappa\frac{\Delta x^2}{2R}$, where κ is the decay constant and R is the tip-surface separation. In terms of predictions, if for example the κ and R were each set to be equal to 1Å respectively, the current will reduce by an order of magnitude at a distance of around 1.5Å away from the centre. This summation is highly simplified and modern STM can achieve resolutions much higher than this (a good lateral and vertical resolution is typically 1Å and 0.1Å respectively [26]) but it provides a good estimate of a typical STM resolution.

A fully functional STM device has several more requirements to be fulfilled before it can be operational. Firstly, a bias voltage needs to be applied to the tip so tunnelling is favoured in a particular direction, otherwise, the process would happen equally in both directions and no overall current would be measured. The position of the tip (in all three Cartesian axes) also needs to be controlled

very precisely, and the microscope needs to be decoupled from the influence of external vibrations. For control, the motion of the tip is achieved via the inverse piezoelectric effect; the tip sits at the end of a tube of piezoelectric material in which sections of the tube are separated by electrodes; by applying a voltage across these different sections independently the tube can be made to bend in different directions. Thus, making it is possible to precisely control the motion of the tip in three dimensions [27]. The device also sits in an ultra-high vacuum with some shock absorbing system to reduce the vibrational noise. Finally, the tunnel current needs to be measured and adjusted via a feedback loop that, along with image generation and spectroscopic control, is provided via a combination of digital signal processing, low and high voltage electronics, and software.

Scanning tunnelling microscopes and their atomic force microscope counterpart, are now ubiquitous in surface science, condensed matter physics, nano-science, and materials science. They have moved far beyond imaging capability to encompass sophisticated spectroscopic (and spectromicroscopic) analyses and the manipulation of matter right down to the single atom, single molecule, and, most recently, single chemical bond limit [28, 29, 30, 31]. The end product is a technique that provides researchers with a way of both visualising and manipulating atomic level structures. The images produced in the later chapters are of course theoretical simulations of STM images, however, it is still important to have an understanding of how they are experimentally generated in order to understand what limitations exist between the two sets and how they vary when compared.

2.2 Properties of Phthalocyanine

Phthalocyanines are a set of large organic molecules consisting of around fifty atoms, that have been a long standing subject of research from both commercial and academic groups. The mostly planar arrangement of atoms in these molecules creates a large π-bonding network that in turn generates novel electronic and magnetic properties. This atomic structure, displayed in figure 2.3 consists of four isoindole units that form a fourfold symmetrical macro-cycle, with each isoindole connected by a nitrogen atom. The centre of this structure is often occupied by a metal atom that is tightly bound into the four most central nitrogen atoms.

The unique properties created by this structure, along with the molecule's stability (phthalocyanine will generally show little to no degradation at temperatures up to $400^{\circ}C$ [32]), and ease of manufacture [33] has made it a key part of many industrial applications. A few of these include its use in industrial dyes [34], detectors for a variety of different gaseous substances [35], components in molecular electronic equipment [36] and spintronic devices [37]. In academia its own properties make it an excellent candidate for work developing and testing theoretical models. Examples of such work have been completed using DFT calculations [38, 39], the Hartree Fock method [40] and Hückel method approaches [41]. Other interests related to the development of particular electronic devices such as photovoltaic cells or organic semi-conductors [42, 43] have made it a com-

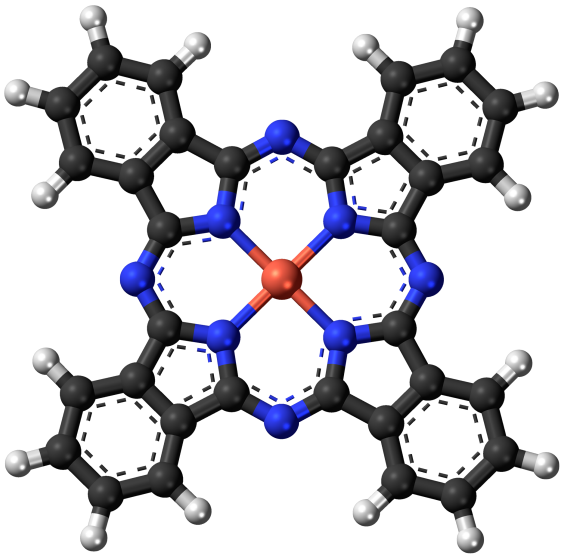


Figure 2.3: Ball and Stick model showing the macrocyclic structure of metal phthalocyanine (MPc) molecules, the colour scheme representation shows grey balls as carbon, blue balls as nitrogen, white coloured balls as hydrogen and the central orange coloured atom is a generic metal atom

mon molecule for study under STM, as understanding the surface effects of these molecules is crucial for improving their function in such devices. The research presented in this work attempts to use a Hückel method approach to provide insight into the interplay of these molecules and aid future applications by offering different ideas on how to discern their behaviour.

2.3 Background of the Hückel Molecular Orbital Theory

The primary objective of this work is to demonstrate that the Hückel method represents an effective, efficient, and accurate alternative to the computationally intensive density functional theory (DFT) methods; which are currently regarded as the standard practice for studies of molecular assemblies on surfaces. Hückel Molecular Orbital Theory (HMO) is a set of semi-empirical methods for resolving the molecular orbital structure of organic molecules and is, therefore, a potential tool for creating molecular models. The ideas that form the foundations behind this theory are presented here, along with their respective nuances, to provide an understanding of where the HMO theory originates and the circumstances in which it is considered effective.

The starting points for the HMO theory is the general quantum mechanical formulation in which the energy of a system is calculated by applying the Hamiltonian, H, of a system to its wavefunction, Ψ, and solving the relevant eigenvalue equation:

$$H\Psi = E\Psi \tag{2.3}$$

For the π-conjugated systems, that are the subject of Hückel method approaches, an exact Hamiltonian is essentially impossible to compute with current techniques. *Ab initio* methods attempt to obtain a close approximate Hamiltonian using the variational method i.e. an iterative minimization process. These methods are applauded for a high level of accuracy but of course still require some approximations, as well as being computationally demanding to complete. The Hückel method is more approximate than those techniques but requires much less computation, thus giving it a potential advantage for use in particular areas of research.

The HMO theory was a process designed as an early quantum mechanical method to study the benzene problem in 1930 [44]. The approach is based on a number of seemingly crude approximations but, as will be demonstrated, this process still has the potential for predicting a number of complex properties. These approximations are described and examined in this section as they are presented in the aptly titled book *Hückel Molecular Orbital theory* by Yates [8]. The first assumption made in the HMO theory is that there exists a separability between π and σ-electrons. Physically, this is akin to stating that π and σ-orbitals do not interact with one another meaning the wavefunction Ψ can be separated into two independent terms ψ_σ and ψ_π such that:

$$\Psi = \psi_\sigma \psi_\pi \tag{2.4}$$

Treating the two systems independently means the energy of each (E_σ and E_π) can be calculated independently and simply combined to obtain the energy of the system E_Ψ:

$$E_\Psi = E_\sigma + E_\pi \tag{2.5}$$

It is further assumed that the terms relating to the σ-bonds are localized around the atoms and can be correctly obtained using a molecular orbital (MO) method, where the interaction is represented with a bond function θ [8]:

$$\psi_\sigma = \Pi_i \theta_{\sigma i} \tag{2.6}$$

Thus, the energy of the σ-electrons is equal to the sum of the bond energies ϵ_i for that system:

$$E_\sigma = \Sigma_i \epsilon_{\sigma i} \tag{2.7}$$

These two assumptions regarding the σ-bonded electrons are reasonable for a large number of organic molecules as the π and σ-electrons only weakly interact with one another. This is demonstrated by considering the bond energy, as this property remains relatively constant between σ bonded atoms whilst changes transpire within the π-bonding network. Thus, it is assumed that the independence of the two systems is justified and that a simple additive approximation is appropriate [8]. The π bonding network represents a large area of delocalization in the electronic structure and therefore requires different handling.

The π electron wavefunction, Ψ_π, can be considered as the product of a set of individual MOs encompassing the entire molecule, thus, each MO in the system can be represented using a linear combination of atomic orbitals method

(LCAO). Disregarding the σ-bonds gives rise to an important feature of a number of Hückel method applications, specifically, it is often only the $2p_z$ atomic orbitals (AO) are used in the LCAO formulation. The justification for this is obtained from the particular circumstances in which Hückel based approaches are commonly used. Generally, these circumstances are where molecules being studied are planar conjugated hydrocarbons, in which the π-bonding orbitals of these molecules, such as ethene or benzene, will only have contributions from the $2p_z$ atomic orbitals (AO). This simplifies the complete π orbital wavefunction, Ψ_π so that it can be described in terms of the MOs as follows:

$$\Psi_\pi = \Pi_i \psi_i \tag{2.8}$$

In which each molecular orbital is described as:

$$\psi_i = \Sigma_{j=1}^{n} C_{ij} \phi_j \tag{2.9}$$

where C_{ij} is the coefficient for the *jth* AO in the *ith* MO and the sum is over all the AOs that contribute to the *ith* MO. Referring back to equation (2.3) each of these MOs will be an eigenfunction of some Hamiltonian operator that can generate the energy E_i of the π-system. Calling this operator H_{eff}, this information can be written as:

$$H_{eff}\psi_i = E_i\psi_i \tag{2.10}$$

What these equations demonstrate is that for a particular molecule a number of the key properties, that result from the π-orbital complex, can be derived by solving the above Hamiltonian in equation (2.10). The method for solving for that Hamiltonian comes from the variational principle, which is utilised to arrive at a secular determinant. Variations of the Hückel method have different levels of approximation for particular terms and result in differing evaluations of the effective Hamiltonian H_{eff}. Many of the specifics of these differences are explained in chapter three, in which the basic, extended and self-consistent versions of the Hückel method are described.

The Hückel method is unfortunately, both widely and incorrectly considered to too crude in its approximations to give accurate results, its use has ,therefore, has been waning in academic circles over the past few decades. Despite this, the method still has the potential as a useful modelling tool across multiple systems, even finding recent applications in work completed with molecular stability, molecular electronics and spintronics [45]. The reason these methods can still be considered useful is because they offer a computationally efficient way to approximate the electron density for large objects, which can often be impractical with more complete techniques. Therefore, when applied appropriately, these methods offer the opportunity of creating flexible and diverse models that are capable of making complex predictions about large scale molecular systems.

Chapter 3

Theory

3.1 Principles of Simulating STM Images

The creation of an accurate STM simulation is, at its most fundamental, an attempt to derive a set of mathematical principles that can predict the observed current between the tip and the sample. The first step in building this model is, therefore, a matter of deducing how to best represent this current. The original argument for STM was based on a model described in the influential 1982 paper by Binning and Rohrer [16]. This model theoretically demonstrated the proof of principle for STM by examining the simple, one dimensional, case of an electron tunnelling through a potential barrier, like the model presented in the last chapter. Whilst their calculations successfully demonstrated that atomic level resolution 'imaging' was possible, the model itself was too crude for even semi-accurate simulations of STM images of molecular assemblies.

A common method for representing the tunnelling effect is Bardeen's tunnelling theory [46]. This theory is highly applicable in regards to STM but it is relatively old and subsequently primitive compared to the modern standard of calculation. Two expansions of this original method were developed as a result, first by Tersoff and Hamman (1985) [47] and later by Chen [48] (1990). Both expansions built on the previous work to include more of the specific geometry and orbital theory involved in STM imaging. All of these ideas will be discussed in this section. It is worth noting before examining these concepts that the overall relationship between the observed current, the tip and sample proves to be relatively simple in terms of generating simulations. Therefore, despite the complexity that arises when examining these concepts, the majority of the computational effort the software required was used in other functions embedded in the program. The word program being a reference to the total collection of functions that comprise the software which was created for this work, an overview of the programs structure is presented in appendix A to show how these functions are connected to each other.

3.1.1 Bardeen's Tunnelling Theory

The idea behind Bardeen's approach to modelling tunnelling phenomena was to define a system with two known Hamiltonians, which can then be used to deduce

a so-called transfer Hamiltonian that describes the change between the two states. For STM these two known Hamiltonians correspond to the electronic states of the tip and sample respectively. The tip and sample states, however, are not necessarily trivial to deduce and so the theory required several assumptions to simplify the derivation of these Hamiltonian operators [46]. The starting assumption is that tunnelling is a single electron process and thus electron-electron interactions are neglected. Generally, ignoring such interactions is a sensible approximation given the low intensity in the tunnelling regime, however, there are certainly circumstances when that approximation is not justified [49]. It was also assumed that the occupational probability of the tip and sample states are independent of each other, that is, coupled states between the two systems are ignored. Finally, the tip and sample are said to both be in electrochemical equilibrium.

Starting with these principles is advantageous as the resulting systems can be modelled by a single electron Hamiltonian:

$$H_{sam} = -\frac{\hbar^2}{2m}\nabla^2 + V_{sam}(\mathbf{r}) \tag{3.1}$$

$$H_{tip} = -\frac{\hbar^2}{2m}\nabla^2 + V_{tip}(\mathbf{r}) \tag{3.2}$$

in which $\frac{\hbar^2}{2m}$ is the kinetic energy operator and the terms, $V_{sam}(r)$ and $V_{tip}(r)$, relate to the potential of the tip and sample states receptively. Considering first the sample state, if the system was governed purely by the sample Hamiltonian, the resulting wavefunction of that system ψ, at a time t, can be given terms of its energy ϵ, as follows [46]:

$$\psi(t) = e^{-it\epsilon/\hbar}\psi \tag{3.3}$$

If the tunnelling effect is 'weak' then the wavefunction for the electrons moving between the two sites is a perturbative parameter and can be approximated by:

$$\psi(t) = -e^{-it\epsilon/\hbar}\psi + \sum_k a_k(t)\phi_k \tag{3.4}$$

where the last term is a sum over all the bound tip states ϕ_k after being corrected by a coefficient $a_k(t)$, which determines the contribution of that state to the systems wavefunction, at time t,. The wavefunction, in the form given by equation (3.4), can then be substituted into the time-dependent Schrödinger equation, which in turn can be resolved to obtain the transmission probability, P, by applying the time-dependent perturbation theory, as demonstrated in [46]. Given a sufficiently large t (for most tip states this turns out to be around a picosecond [50]) the resultant form for P is equal to Fermi's golden rule [46, 51]. Therefore, in terms of the tip and sample state energies, E_ν and E_μ respectively, P can written as:

$$P = \frac{2\pi}{\hbar}\sum_{\mu,\nu}|M_{\mu\nu}|^2\delta(E_\mu - E_\nu) \tag{3.5}$$

where $M_{\mu\nu}$ is known as the transfer matrix. An assumption in this formulation is that the tunnelling current is produced by elastic tunnelling, hence, the energies,

E_ν and E_μ, are represented in equation (3.5) using a Kronecker delta [21].

Given that the current, I, is proportional to the transmission probability i.e. $I \propto P$, a derivation of I can be achieved by performing a sum over the number states available for the electrons to tunnel into. In practice, this involves using Fermi-Dirac statistics to deduce the occupied and unoccupied states of both the tip and the sample after applying a bias voltage. The resultant sum over discrete states can then be transformed into an integral over the density of states [46], which is straight forward to evaluate. The result from those equations, that is essential for the purposes of this work, is simply that:

$$I \propto |M|^2 \tag{3.6}$$

therefore, calculating M offers a solution for deriving the current. To be able to obtain a relationship, for generating simulations, that is as easily implemented as equation (3.6), it has to be assumed that states lie near the Fermi-level. In real terms, this is akin to assuming the STM simulation is taking place in a low temperature environment.

The transfer matrix itself is generally derived by integrating over the wavefunctions of both the tip and the sample, ψ_ν and ϕ_μ respectively, inside the barrier region:

$$M = -\frac{\hbar}{2m}\int_{barrier}(\phi_\mu^*\Delta\psi_\nu - \psi_\nu^*\Delta\phi_\mu)dS \tag{3.7}$$

As the exact tip states are often unknown, resolving equation (3.7) is at best not trivial and a large part of the development of more robust modelling techniques.

3.1.2 Transfer Matrix Elements

Solutions to the transfer matrix M were first successfully derived by Tersoff and Hamman by including more of the specific geometry of STM to build on the foundations described in Bardeen's tunnelling theory. Chen later expanded this theory further by increasing its applicability to more real STM experiments, full derivations of their work can be found in [47, 48].

The important result, from both their ideas, was that by making certain assumptions about the tip states it was possible to relate the transfer matrix M solely to the sample wavefunction calculated at the tip's centre $\mathbf{r_0}$. Both Tersoff and Hamann [47], and Chen [48] concluded that in the case of the terminating tip orbital being s-type (i.e. radially symmetric), M is related to the sample wavefunction at the tip's centre, $\psi(r_0)$ as follows:

$$M \propto \psi(r_0) \tag{3.8}$$

Tersoff and Hamman were able to derive the relationship in equation (3.8) by assuming an s-type orbital at the tip's centre and applying the use of a two dimensional Fourier transform to the wavefunction [47]. A key issue with their formulation is that by assuming an s-type orbital at the tip's centre the equations predict

values of corrugated surface features to be far smaller than the experimental values [52]. Chen improved on that derivation by removing the assumption that the tip state needs to be an s-type orbital. This was achieved by first modelling an s-type orbital at the tip's centre using a spherical modified Bessel function, which can then be solved for the matrix elements by applying a Green's function method. This part of the solution to the tunnelling current agrees with the relationship in equation (3.8) [48]. The advantage of this method, over Tersoff and Hamman, is that it allowed for the development of what is called Chen's derivative rule; a set of principles that are used deduce the relationships for higher order tip states. A table displaying the results of these principles is presented in table 3.1.

Table 3.1: Change in sample wavefunction as effected by change in the tip orbital

Tip Orbital	**Proportionality to** M
s	$\Psi(\mathbf{r_0})$
p_x	$\frac{\partial}{\partial x}\Psi(\mathbf{r_0})$
p_y	$\frac{\partial}{\partial x}\Psi(\mathbf{r_0})$
p_z	$\frac{\partial}{\partial x}\Psi(\mathbf{r_0})$
d_{xy}	$\frac{\partial^2}{\partial x \partial y}\Psi(\mathbf{r_0})$
d_{zy}	$\frac{\partial^2}{\partial z \partial y}\Psi(\mathbf{r_0})$
d_{zx}	$\frac{\partial^2}{\partial z \partial x}\Psi(\mathbf{r_0})$
$d_{x^2-y^2}$	$\frac{\partial^2}{\partial x^2}\Psi(\mathbf{r_0}) - \frac{\partial^2}{\partial y^2}\Psi(\mathbf{r_0})$

In terms of the simulations used to present the research in the later chapters, the images will be presented as if the tip is an s-type orbital. Thus, the relationship between the current and the sample wavefunction used for generating STM images is:

$$I \propto \psi(\mathbf{r_0})^2 \tag{3.9}$$

The validity of simulations, in terms of how well they match with experiments, is dependent on a number of conditions. In this formulation it is assumed that the system is near the Fermi-level, therefore, it is assumed the experiment is performed at a low temperature. This model can also not explain a number of the detailed effects found in STM experiments that are artefacts of coupled states between the tip and the sample, as only the sample wavefunction is being considered. As the simulations in later chapters demonstrate, the images produced from this formulation can and do closely resemble real images observed in real STM experiments.

3.2 Molecular Orbital Calculations

A key consideration for the start of this work was to examine and select appropriate methods for performing molecular orbital calculations. By its very nature full analytical solutions to the Schrödinger equation are either impossible to obtain or impractical to compute, except in a limited number of simple cases. A

consequence of such limitations is that there now exists a vast field of ideas and methods for constructing approximate models of quantum mechanical systems. Given that the Hückel Method was chosen for these calculations it is important to detail why this decision was made and why other methods were rejected. This next section will describe a few common ideas in this field to set out the arguments that concluded that a Hückel theory based approach was both appropriate and advantageous.

3.2.1 Hartree-Fock Method

The Hartree-Fock method (HFM) is a comparatively old method of calculation, having been developed in the late 1920s shortly following the publication of the Schrödinger equation. It is still used today and can be particularly effective when applied in combination with other ideas such as local density-functional theories [53]. The HFM is focused on using the wavefunction of the system to derive its complete Hamiltonian operator and is the earliest example of using the variational method to describe a self-consistent loop [54]. These properties mean that the HFM forms a basis for understanding the ideas used in other more modern methods such as DFT and the self-consistent extended Hückel theory, both of which are detailed in this section. For these reasons, a brief description of the HFM is presented here.

The HFM starts with defining the Schrödinger equation in the Born-Oppenheimer approximation [55], a common approximation used ubiquitously in molecular systems. If defined using one-electron and two-electron operators, $h(i)$ and $v(i,j)$ respectively, the Born-Oppenheimer approximation can be written as:

$$\hat{H} = \sum_i h(i) + \sum_{i<j} v(i,j) + V_{NN} \tag{3.10}$$

The one-electron operator $h(i)$ includes the kinetic energy terms and the electron-nucleus interactions, while the electron-electron interaction and the external potential are represented by the two-electron operator $v(i,j)$ and potential term V_{NN} respectively. The basic idea of self-consistency comes from the variational method and is implemented as follows. A trial Hamiltonian is 'guessed' and used to calculate the energy of the system. The result of that first calculation is used to further derive a variable that alters the Hamiltonian in order to reduce that energy. This process of calculation and alteration can be repeated so that as the system approaches a minimum energy value, meaning that the amount by which Hamiltonian is being altered at each iteration approaches zero. At that point, the algorithm terminates, meaning that the energy predicted by that final Hamiltonian, in turn, predicts that exact same Hamiltonian. It can be concluded that the final energy from that calculation is at least close to being an exact eigenstate of the real Hamiltonian of the system and thus the set of equations is described as being self-consistent. In practice following each of these steps requires going through the following process.

This is started by describing an electronic wavefunction. In the HFM the standard practice is to represent the electrons as non-interacting particles but in a

way as to also satisfy their anti-symmetric properties. Often this is accomplished using a Slater determinant:

$$\Psi(\mathbf{x_1}, \mathbf{x_2}, ..., \mathbf{x_N}) = \frac{1}{\sqrt{N}} \begin{vmatrix} \chi_1(\mathbf{x_1}) & \chi_2(\mathbf{x_1}) & \dots & \chi_N(\mathbf{x_1}) \\ \chi_1(\mathbf{x_2}) & \chi_2(\mathbf{x_2}) & \dots & \chi_N(\mathbf{x_2}) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(\mathbf{x_N}) & \chi_2(\mathbf{x_N}) & \dots & \chi_N(\mathbf{x_N}) \end{vmatrix} \tag{3.11}$$

equation (3.11) represents the wavefunction as the matrix determinant of a set of individual electron spin-orbital functions. These describe each electron independently based on their relative positions and spins where both of these properties are encoded in the variable $\mathbf{x}$. For a full description of exactly how the various important properties of the HF wavefunction are satisfied by the equation (3.11), as well as what properties aren't, the reader is referred to [3]. An important point for consideration in this formulation is to justify the highly unrealistic approach of assuming the electrons are non-interacting. This would, on first consideration, seem incapable of producing accurate results. However, instead of implicitly dealing with each interaction between electrons these are instead handled by being represented as a constant potential. Physically this is suggesting that most of the specific detail of the electron-electron interactions will 'cancel out' so they can be approximated by some average value that is 'felt' by all the electrons. For this reason, the HFM is also described as a 'mean field theory' approach.

Given HF is a variational method approach, the aim becomes to find a version of the electronic spin-orbital functions that minimizes the energy:

$$E = \int \Psi^* \hat{H} \Psi d\tau \tag{3.12}$$

where the Hamiltonian $\hat{H}$ is the same as the one described in equation equation (3.10) and Ψ is equal to the Slater determinant in equation (3.11). Finding this minimum energy solution means finding an expression that minimizes the energy with respect to changes in the orbitals. A property built into the Slater determinant is that it works under the assumption that the orbital expressions χ_i are orthonormal, so it is important that the method used to change the output of these orbitals, as part of the minimization process, also maintains their orthonormality. The mechanism used for resolving this problem comes from Lagrangian mechanics and is subsequently called the Lagrange method of undetermined multipliers.

The full derivation of this process in the HFM is lengthy and complicated so is only outlined in this description but a full analysis can be found in [3]. The main concept is to set up a Lagrangian equation as a functional of the orbitals with the constraint that each orbital is orthonormal:

$$L[\chi_i] = E[\chi_i] - \sum_{ij} \epsilon_{ij}(S_{ij} - \delta_{ij}) \tag{3.13}$$

In this expression ϵ_{ij} are the undetermined multipliers, S_{ij} is the overlap for orbitals i, j and δ_{ij} is the Kronecker delta. This equation creates a variable L that is equal to the energy E when the orbitals are orthonormal (when $i = j$ in the second

term) and zero otherwise. Finding the derivative of this function $\delta L[\chi_i]$ and setting it equal to zero allows, with some work, the derivation of the Hartree-Fock equations:

$$[h(\mathbf{x_1}) + \sum_j J_j(\mathbf{x_1}) - \kappa_j(\mathbf{x_1})]\chi_i = \epsilon_{ij}\chi_i \tag{3.14}$$

The new operators J and κ are the Coulomb and exchange integrals respectively and h is the same one-electron operator described in equation (3.10). In this form equation (3.14) acts as an eigenfunction, with the left hand side representing a Hamiltonian that acts on the orbitals χ, resulting in a series of eigenvalues ϵ_{ij} that represents the energy on the right hand side. This becomes much more apparent if the left-hand side is described in a single operator:

$$f(\mathbf{x_1}) = h(\mathbf{x_1}) + \sum_j J_j(\mathbf{x_1}) - \kappa_j(\mathbf{x_1}) \tag{3.15}$$

thus, equation (3.14) becomes:

$$f(\mathbf{x_1})\chi_i = \epsilon_{ij}\chi_i \tag{3.16}$$

Solving the above expression is the fundamental process of the HFM and can be adapted using a number of different ideas [56, 57]. A common approach involves the Roothaan equations [3], where the orbitals are transformed by a basis function, such as a Slater-type or Gaussian-type orbital function, to produce:

$$\chi_i = \sum_{\mu=1}^{K} C_{\mu i}\chi'_i \tag{3.17}$$

The above equation describes the transformation of the orbitals χ_i into a series of basis orbital functions χ'_i, each fronted by a coefficient $C_{\mu i}$. Equation (3.16) can now be written in a matrix form by using the above transformation. When substituted into equation (3.16) and left multiplied by the conjugate of this basis χ'^* the Hartree-Fock equation becomes:

$$\mathbf{FC} = \epsilon_{ij}\mathbf{CS} \tag{3.18}$$

in which $\mathbf{S}$ is the overlap and $\mathbf{F}$ is the Fock matrix, which has matrix elements equal to the expectation value of the Fock operator when applied to a given orbital:

$$F_{\mu\nu} = \int \chi'^*(\mathbf{x_1}) f(\mathbf{x_1}) \chi'(\mathbf{x_1}) d\mathbf{x_1} \tag{3.19}$$

The loop that leads to a self-consistent expression from the above equations starts with 'guessing' a value of the eigenvectors $\mathbf{C}$ that are used to derive a basis set of orbitals that can solve for $\mathbf{F}$. Feeding $\mathbf{F}$ back into equation (3.18) results in a new set of values for $\mathbf{C}$.

The key feature of HFM is that it is *ab initio* and complete. This means that it can derive an accurate form of the Hamiltonian for a complex system from first principles, therefore deducing everything that might want to be known about

that system. The main difficulty with the implementation is that each iteration of the loop is highly taxing computationally. For a molecule consisting of just ten electrons (methane for example), a single iteration requires performing calculations of a hundred positions in the Fock matrix which itself is a set of operators acting on a matrix consisting of a large basis set of orbitals. It is for this reason that the complexity of the computation increases rapidly with the increasing electron number, N. The factor itself varies depending on the approach but is generally within in the range of N^5 to N^7 [4].

Since the advent of the HFM numerous other wavefunction based methods have been developed, these are collectively known as the Post-Hartree-Fock methods (PHF) and such examples include the configuration interaction method, coupled cluster theory or Møller–Plesset perturbation theory [58, 59, 60]. The implementations of these vary, however, they tend to build on the HFM by focusing on improving the electron correlation relation in systems instead of increasing algorithmic efficiency. While the increase in accuracy is beneficial the need for a more computationally efficient method created a new field of ideas centred around using density calculations as a replacement for wavefunction methods.

3.2.2 Density Functional Theory

Density functional theory (DFT) is currently the most popular and widely used computational method for molecular calculations, so much so that it is practically considered to be *de rigeuer* as a theoretical tool across the fields of condensed matter physics, surface science, and nanoscience. DFT, however, was not used during this work and given its reputation the reasons for this are worth exploring, as they provide insight into the advantages of a Hückel method approach.

To discuss why DFT wasn't used it must first be understood what DFT is and what its function would be in a project of this kind, thus, a brief description of the concepts behind DFT and an application that could have been used in this project are presented here. The conceptual idea is based on two theorems set out by Hohenburg and Kohn in 1964 [61]:

1. *The ground state energy of a many-electron system is a unique functional of the electron density and so can be entirely derived from the electron density*

2. *The ground state energy can be found variationally i.e. the density functional which minimises the total energy is exactly the density of the ground state*

The derivation of these two principles is relatively complicated and so won't be shown in full detail here (detailed versions can be seen in [3]). Importantly, the Hohenburg and Kohn theorems demonstrate that there exists a method for calculating the ground state energy that only requires knowledge of the system's electron density. Crucially, this means that the calculation does not require any representation of the wavefunction itself. This facilitates a significant lowering of the computational cost when compared to HF and PHF methods, as it drastically reduces the dimensionality of the problem. Wavefunction handling requires consideration of a system that contains 3N dimensions compared to just the three

spatial dimensions associated with the density, thus producing a substantial decrease in complexity. The concession for the advancement in efficiency is related to the completeness. While HF and PHF methods are only limited by the accuracy in accounting for each term correctly, DFT has inherent limitations such as the self-correction error that creates problems in predicting band gap energies and other phenomena [62]. Remedying these errors often necessitates going beyond DFT or applying it as a hybrid with other methods [63, 64]. Despite these limitations, it is currently the most powerful and widely used method available.

Applications of DFT vary widely but share the commonality of all being variations of approximating the kinetic energy and electron-electron interaction density functional. From the first Hohenburg and Kohn theorem it is demonstrated that the energy is a functional of the density, that can itself be written in terms of the electronic operators that are also functionals of the density. Under this formulation the Born-Oppenheimer approximation [55] is defined as follows:

$$E[\rho] = T[\rho] + V_{ee}[\rho] + V_{ext}[\rho] \tag{3.20}$$

where the right-hand side shows kinetic, electron-electron and external interaction functionals respectively. The methodology for solving equation (3.20) is derived from the second theorem which in the context of the above expression says that if there exist functionals for those terms, minimisation of said functionals will result in the correct energy expression. However, those are non-trivial to find and are still the subject of much ongoing research, examples of which can be seen in [65, 66]. The method that will be discussed here is the Khon - Sham approach [67]. This is among the most popular approaches and it contains many similarities to the Hartree-Fock method shown in the previous section and so makes for excellent comparative analyses.

The method by which Khon and Sham make their approximation to the necessary functionals is with the addition of non-interacting electronic orbitals, such that the density is:

$$\rho(\mathbf{r}) = \sum_{i}^{N} \mid \phi_i \mid^2 \tag{3.21}$$

details of these derivations can be found in [3]. It can be shown, using the orbital definition of the density in equation (3.21), in conjunction with the energy functional relation in equation (3.20), that the ground state energy is an eigenvalue of the following function:

$$\big[-\frac{1}{2} \bigtriangledown^2 + v_{eff}[n^k(\mathbf{r})] \big] \phi(\mathbf{r})_i = \epsilon_i \phi(\mathbf{r})_i \tag{3.22}$$

In equation (3.22), the ground state energy is minimised, so long as the orbitals are satisfactory (that is, they satisfy the equation in question, in this case equation (3.22)). The above equation bears a significant resemblance to the HF equation (and the common expression of the Schrödinger equation) but with the difference that the potential component of the Hamiltonian is defined in terms of a density. The potential v_{eff}, which is a combination of the external potential, Coulomb interactions and the exchange-correlation function is, in this format, much simpler

to solve than in the HF formulation as it only spans three-dimensional space for each orbital. In this process there is again a self-consistent loop where the density is guessed and used in equation (3.21) to find an orbital set, that is then used to find a new density. The orbital set can be adjusted according to this correction in the density value. The loop is repeated until the orbital basis converges and predicts the density used in the Khon - Sham equation (equation (3.22)). Formally the idea scales theoretically as N^3 with additional electrons but current methods have found exploits that put the factor closer to N^1 [68]. This is a vastly significant improvement compared to the wavefunction methods discussed earlier.

Even with the scaling improvements outlined above, current methods are still largely impractical for systems containing many thousands of atoms, like those being studied in this investigation. While possibilities such as using a periodic DFT calculation are potentially applicable, a Hückel method approach is simpler and much more efficient to process. The key advantage is flexibility, by utilizing the HMO theory for calculations it enabled a large number of comparative and investigative calculations (displayed in chapter six), that would otherwise be too time-consuming and cumbersome to obtain.

3.2.3 Basic Hückel Method

In this section, so far, the discussion has been limited to two types of *ab intio* method, which are, by their very nature, more complex than any Hückel based application. The simplicity of a Hückel method approach derives from several approximations (some of which are outlined in chapter two), that are comparatively crude when contrasted to the ideas behind wavefunction or DFT methods. As was already alluded to, this work argues in favour of this simplicity by suggesting that when a Hückel method approach is applicable its efficiency allows for a level of flexibility lost with other techniques.

The most easily applied Hückel approach, the basic Hückel method (BHM), is a linear combination of atomic orbitals method (LCAO) [8]. The general definition of a molecular orbital Ψ^μ, in the LCAO method, for a given orbital basis χ, is given as:

$$\Psi^\mu = \sum_{i=1}^{N} c_i^\mu \chi_i^\mu \tag{3.23}$$

where μ indicates which MO is being constructed, i which particular AO is contributing a given MO and c refers the coefficient that determines the contribution of each AO to the μth MO. The first of the simplifications made in the BHM was detailed in the previous chapter. If the BHM is limited to planar conjugated hydrocarbons, then only the p_z AO contributes to the π-bonding network, and therefore the rest can be disregarded. Equation (3.24) is consequently simplified to:

$$\pi^\mu = \sum_{i=1}^{N} p_z^\mu c_i^\mu \tag{3.24}$$

The total wavefunction Ψ^{μ} has been reduced to considering only the π-bonding MO wavefunction π^{μ}, and χ is now a fixed singular p_z orbital. Determining π^{μ} is achieved through the resolution of the coefficients c. As established in the previous chapter, each of these MO's will be an eigenfunction of some Hamiltonian $\hat{H}$. For molecular systems the overlap is often a necessary part of the molecular orbital function, therefore, the standard Hamiltonian eigenfunction can be adjusted to include this term as follows:

$$\hat{H}\pi^{\mu} = E\pi^{\mu}\hat{S} \tag{3.25}$$

Equation (3.25) represents the eigenfunction of the Hamiltonian $\hat{H}$ with respect to the molecular orbitals π^{μ}, whilst including the overlap of these orbitals $\hat{S}$. It is assumed in all Hückel method techniques that these functions are described using matrix formulation, meaning $\hat{H}$ and $\hat{S}$ become H_{ij} and S_{ij} respectively. In the BHM S_{ij} is approximated to be equal to:

$$S_{ij} = \int p_z^i p_z^j d\tau \text{ is equal to } \begin{cases} S_{ij} = 1 & \text{if } i = j \\ S_{ij} = 0 & \text{if } i \neq j \end{cases} \tag{3.26}$$

thus, S_{ij} is now the identity matrix and is cancelled out, this simplification is among the most egregious assumptions of the BHM. The validity of work when neglecting this term is particularly dependent on circumstance and greatly limits the number of systems this technique can be applied to [8].

Given that in equation (3.24) the orbital function p_z is a singular independent object it is equivalent on both sides and can be removed from the expressions. Equation (3.25) can, therefore, be written entirely in terms of the coefficients c_i^{μ}:

$$H_{ij}c_i^{\mu} = E_i c_i^{\mu} \tag{3.27}$$

Solving for these coefficients and the subsequent energies is now reduced to a relatively straight-forward eigenvalue problem, assuming the Hamiltonian can be constructed in a simplistic fashion. In fact, a fundamental component of the computational efficiency of the BHM is the simplicity of the Hamiltonian created in this scheme. Hückel first made the approximation that any Hamiltonian integrals will vanish if the atoms are not nearest neighbours. It was further assumed that any non-zero term of the Hamiltonian matrix is described using two constants: α, which is the Coulomb integral that represents the diagonal elements in the matrix; and β, which is known as the resonance integral and is assigned to the matrix elements representing nearest neighbours [8]. Physically these terms represent the energy difference between the observed delocalised system and a theoretical system in which the electron in question is tightly bound to a nucleus. These assumptions can be expressed mathematically as follows:

$$H_{ij} = \int \psi^i \hat{H} \psi^j d\tau \text{ is equal to } \begin{cases} H_{ij} = \alpha & \text{if } i = j \\ H_{ij} = \beta & \text{if } i, j \text{ refer to adjacent atoms} \end{cases} \tag{3.28}$$

This is again a particular crude approximation of the complex interactions governing real molecular systems. However, in the particular molecules where this is

applied, the BHM represents these systems remarkably well and examples of its application to benzene and C_{60} molecules can be found in [69, 5]. A crucial aspect of the performance of the BHM algorithms results from it not being necessary to explicitly calculate the values of α and β, thus no integrals need to be evaluated to derive the Hamiltonian. The only operation that is strictly necessary is a diagonalization routine, which is performed on the Hamiltonian to solve the terms in the eigenfunction. In modern day computing, the implementation of such a function is trivial. A consequence of not performing these integrals is that the exact energy levels can't be expressed numerically but must instead be expressed as a function of α and β, therefore this method is incapable of determining an exact energy of the system. Instead, the BHM yields the number of energy levels, the number of degeneracies, and the prominence of each AO in a given MO (this being the functionality of the LCAO coefficients, c_i^μ). In terms of creating simulations, the shape and structure of the MO is fundamentally determined by the coefficients c_i^μ. Therefore, these drawbacks do not limit the ability of researchers to study molecular systems using the BHM, as long as the system under investigation is selected appropriately.

3.2.4 Extended Hückel Theory

The extended Hückel theory (EHT) attempts to improve on the limitations of the BHM by incorporating a number of the terms either left out or grossly oversimplified in the BHM. By building on the calculation method in this manner the theory becomes more robust and applicable to real-world materials. The primary alteration in which the EHT builds on the BHM is that all valence state AOs are taken into account, including the orbitals associated with hydrogen atoms and those bonded in the σ orbital systems. By incorporating these components into the calculation the theory can be accurately applied to a much greater number of molecules. When using the BHM researchers are limited to planar conjugated hydrocarbons, whereas the EHT can be applied to an entire series of organic molecules. Examples of the use of the EHT in modelling studies can be found in [70, 71, 72]. Consequentially, an adjustment must be made in regards to the overlap matrix. Consideration of systems where mixing occurs between σ-orbitals and π-orbitals necessitates being able to represent how much of such mixing may occur. Thus, a full and complete overlap matrix of each possible combination of AOs must be included in the calculation, in which each term is equal to:

$$S_{ij} = \int \psi^i \psi^j d\tau \tag{3.29}$$

The matrix element ij in the overlap matrix is given by taking the integral of the product of the ith and jth AO (represented by ψ in equation (3.29)) over all space. Furthermore, increasing the scope of the Hückel calculations to include more AOs also necessitates that the EHT Hamiltonian is constructed with a more detailed approach, properly considering the contribution from each orbital. Maintaining simplicity is still an important part of this procedure, and so, similar Coulomb and resonance integrals are approximated. Following a similar line of reasoning to the BHM, the EHT evaluates the terms in the Hamiltonian by using a tight binding approximation, in which each term is equal to the energy required for the

valence electron to become delocalised. Therefore, the Hamiltonian is a representation of the energy difference between a tightly bound system and an energetic system capable of generating the π-orbital network present in organic molecules. Mathematically this achieved by first setting the diagonal elements, H_{ii}, equal to an appropriate valence state ionization potential (VSIP). The resonance terms, H_{ij}, are then assigned values derived from a combination of these potentials and calculated according to the Wolfsberg-Hoffman approximation [73]:

$$H_{ij} = K\frac{H_{ii} + H_{jj}}{2}S_{ij} \tag{3.30}$$

Equations (3.29) and (3.30) describe each new entry into the Hamiltonian as being the average of the two VSIP values, H_{ii} and H_{jj}, for the orbitals multiplied by their overlap, S_{ij}, and by some constant scaling factor K. Wolfsberg and Hoffman deduced that the best results are found when K is taken to be 1.75 [73]. It should be noted, given that Hückel based calculations are considered in terms of relative values (VSIP being a relative potential), the results are not very sensitive to the exact value of K being used. Calculation of the respective molecular orbitals, ψ^{μ}, follow the same procedure as the BHM, with the only exception being that the secular equation is the version described in equation (3.25), which also includes the overlap. This method is an improvement on the BHM by being more flexible and applicable to systems including those with multiple elements and non-planar structures. The only significant cost of the EHT, when compared to the BHM, is the calculation of the overlap matrix, as that requires a significant number of integrals to be evaluated. The increase in computation is not significant and essentially negligible when compared to HF, PHF or DFT methods. All of the ideas in those methods require the evaluation of multiple intricate observables, each of which is associated with a significantly higher computational cost than that of the overlap calculation. As will be seen, the process of running these algorithms is straightforward and can be completed in a matter of seconds in many cases.

3.2.5 Self-Consistent Extended Hückel Theory

The EHT is, in many regards, a useful and powerful tool but there are a number of important criticisms. Chief amongst the issues with applying an EHT method, to model real systems, is that it does not explicitly treat the electrostatic repulsion between the electrons in any specific way. The most complete technique that is applied to correct this issue (whilst fundamentally remaining a semi-empirical approach), is to require the EHT Hamiltonian to become self-consistent. The self-consistency in this context is found by making a correction to the valence state ionization potential values (these being the same VSIP values used in the EHT calculation), calculating a new Hamiltonian, and using this corrected Hamiltonian to deduce a new correction to the VSIP values. Again, just as with the HF methods, this process is repeated until the Hamiltonian converges. Multiple versions of this method exist but the most common approach was presented in 1973 in [74]:

$$H'_{ii} = H_{ii} - a_i q_I \tag{3.31}$$

The corrected Coulomb interaction term H'_{ii} is derived from reducing the original value H_{ii} by the amount $a_i q_I$. The a_i term is an adjustable parameter that is related to the ionization energy of the orbital and q_I is the partial charge [75] of the entire atom. The off-diagonal elements are then calculated using the standard EHT theory equation (3.30).

Other versions of the SC-EHT tend to build on this variant of the theory and so the discussion of this particular method is applicable to its extensions [75]. Although SC-EHT theory was considered for use in the research presented in this thesis, the decision was taken to instead focus on the "standard" EHT for two reasons: (i) the extended Hückel theory has been successfully applied to molecules and studies of a similar nature [76, 77], concluding it is highly likely to be adequate in of itself; (ii) calculation of the partial charges required for the SC-EHT method is far from trivial and requires a computationally expensive (but often far from accurate) method such as Mulliken population analysis or electronegativity equalization method [78, 79].

Out of all of the molecular orbital calculation methods described in this chapter the only one used consistently was the EHT. The EHT was deemed to have a good balance between accuracy, applicability for research, and a low computational cost associated with it. A DFT or SC-EHT program would have been useful as a comparative measure, however, the time taken to derive an efficient implementation would have been too great for it beneficial. If the project was extended, the development of these routines would form part of the continued research in order to evaluate the accuracy of an EHT routine when compared to other work in the field.

3.2.6 Empirical Methods

The discussion has so far reviewed a number of quantum mechanical and semi-empirical approaches for modelling molecules. There also exists a range of fully empirical ideas for representing molecular structures [1]. Such a method was used for deducing the angular orientation of titanyl phthalocyanine (TiOPc) molecules in a 2*OO*5 study by Mannsfeld and Fritz [6]. This paper will be used as a comparative study in chapter five, so it is worth discussing the details of that method in order to better understand the nature of the comparison. In addition, the benefits and limitations of their chosen method are applicable to a number of other empirical techniques.

The particular but also common approach used for the Mannsfeld and Fritz paper [6] was to create a model of the Van der Waals forces via the Lennard-Jones potential (LJ):

$$V^{LJ}(\mathbf{r}) = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] \tag{3.32}$$

Equation (3.32) describes the potential between two atoms as a function of the interatomic separation, r, a graph representing the function is presented in figure 3.1. For values of r less than the equilibrium separation, r_0, the repulsive term, $(1/r)^{12}$, dominates. For values of $r > r_0$, the attractive term, $(1/r)^6$, instead dom-

inates the interaction. The σ and ϵ terms are experimentally fitted parameters for a given material.

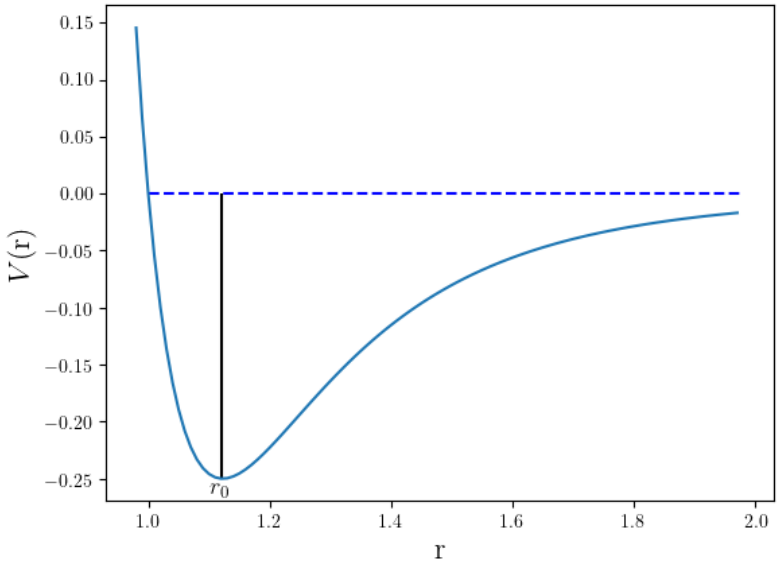


Figure 3.1: A demonstration of the r dependence of a Lennard-Jones potential. The parameters σ and ϵ were arbitrarily chosen to be set to 1 and $\frac{1}{4}$ respectively as these clearly display the nature of the relationship. The minimum point r_0 represents the bond length in this model.

The advantage of using this system is that, like the Hückel method, the computational effort required is minimal, particularly when compared to the *ab inito* methods discussed in this section. However, these types of models become less successful when applied to circumstances that require more specific detail, such as studying the effects of defects [1]. This and similar functions also have limitations related to the requisite assumptions about the physics. For instance, the LJ function is only applied isotropically and the repulsion term is calculated as simply the square of the attraction term, while in reality this function would be an exponential. Therefore, whilst technically being physically related to the Pauli-exclusion principle, as it represents the repulsion that comes from the overlapping electron clouds, it is not mathematically derived from those principles. It is from this basis that semi-empirical methods, such as the EHT, can be considered advantageous over many empirical ideas. Whilst Hückel based applications are highly simplistic versions of the quantum mechanical fundamentals, they are still rooted in those principles and so are generally considered to be more representative of real systems.

Summarising, the computational methods described in this chapter, thus far, each represent a trade-off between accuracy and computational expense. It is possible to have a highly accurate, high computational cost (and therefore low capacity) simulation method that performs well for very small system sizes, but this would

quickly become intractable for even modest increases to the number of atoms. Hückel method approaches lie in the centre of this spectrum of computational methods, leading to a particular advantage when it comes to modelling large systems such as the monolayer molecular assemblies comprising of thousands of electron orbitals, that are the focus of this thesis. *Ab initio* methods remain exceptionally computationally expensive and Hückel methods represent a pragmatic, and often powerful, trade-off between accuracy and cost.

3.3 Choices for the Orbital Basis

So far, and particularly in equations presented section 3.2, the atomic orbitals (AO) have been expressed as single variables. Notating them in that way is essential for those explanations to be readable, however, it does then beg the important question, what are those orbital bases. Three possibilities, that could have been applied to the MO calculations, are explored in this section.

3.3.1 Hydrogen-like orbitals

A possible choice, originally presented by Schrödinger, is to use the orbital functions derived from solving the Schrödinger equation for the hydrogen atom [80]. These are described as follows.

A common formulation for expressing AOs is to separate the components of the orbital function, $\chi(r)$, into its radial and spherical components, $R(r)$ and $Y_l^m(r)$ respectively, to create the following expression:

$$\chi(r) = R(r)Y_l^m(r) \tag{3.33}$$

The spherical component is represented by the relevant spherical harmonic. For instance, a p_z orbital, for which the quantum numbers l and m are equal to 1 and 0, is said to equate to the spherical harmonic Y_0^1 which is equal to:

$$Y_0^1 = \sqrt{\frac{3}{4\pi}}\frac{z}{r} \tag{3.34}$$

The radial component $R(r)$ is more complicated and derived directly from the quantum mechanical properties of the orbital. In general, for hydrogen-like orbitals, $R(r)$ is equal to:

$$R_{nl}(r) = \sqrt{Z'^3\frac{(n-l-1)!}{2n[(n+1)!]^3}}e^{-\frac{Z'r}{2}}(Z'r)^l L_{n-l-1}^{2l+1}(Z'r) \tag{3.35}$$

This equation represents $R(r)$ in terms of the quantum numbers, n and l; the effective atomic number, Z'; and a function, $L(r)$, which is a Laguerre polynomial. The effective atomic number is evaluated as an adjustment to the atomic number, Z, by performing the following calculation:

$$Z' = \frac{2Z\mu}{nm_e a_0} \text{ with } \mu \text{ being the reduced mass } \mu = \frac{m_n m_e}{m_n + m_e} \tag{3.36}$$

This above expression is described in terms of the radial distances r, the principal quantum number n, the Bohr radius a_0, the mass of the electron m_e, and the nuclear mass m_n. To demonstrate the kind of calculation that would be required to simulate this orbital type, a hydrogen-like p_z orbital is described as follows:

$$P_z(r) = R_{21}(r)Y_1^0(r) = \frac{r^2}{12\sqrt{6}} Z'^{\frac{7}{2}} e^{-\frac{Z'r}{2}} L_0^3 \sqrt{\frac{3}{4\pi}} \frac{z}{r} \tag{3.37}$$

The major advantage of using orbitals of this type is that they are exact eigenfunctions of the one-electron system [80]. This makes them highly representative of real-world data and a good way to approximate electrons in molecules and other elements. However, its usefulness is limited by its complexity, and in particular, by the presence of radial nodes. These features represent stationary points where the wavefunction is equal to zero. Whilst nodes have important ramifications in turns of electronic behaviour [81], their inclusion makes integration costly to perform and so these orbitals are rarely used for molecular calculations.

3.3.2 Slater-Type Orbitals

A more common approach is to represent an AO as a so-called Slater-type orbital (STO), which is defined as follows. Remaining in the spherical harmonics model, the form of an STO can be split using the same formulation described in equation (3.33). The advantage of this scheme is that for a different orbital function only the radial component changes. For an STO, this radial part is given by:

$$R_n(r) = Nr^{n-1}e^{-r\zeta} \tag{3.38}$$

where ζ is a constant related to the effective charge of the orbital and the normalization N is defined as:

$$N = (2\zeta)^n \sqrt{\frac{2\zeta}{(2n)!}} \tag{3.39}$$

Remaining with the example of a p_z orbital for comparative purposes, an STO with $R_2(r)$ and Y_1^0 will be described as:

$$\chi(r) = R_2(r)Y_1^0 = 4\zeta^2 \sqrt{\frac{6\zeta}{96\pi}} z e^{-r\zeta} \tag{3.40}$$

In many regards the functions for the hydrogen-like orbital and STO are similar. In both cases the p_z orbital function has a similar exponential dependence, both containing a term equal to e^{-kr}, k in this context being some constant. The most significant omission of the STO is the Laguerre polynomial. The exclusion of this function was an essential objective for researchers when designing the STO function, as this term contains the information relating to the radial nodes, that, as mentioned in the last section, make integration difficult to perform. Integration of STO functions are, therefore, substantially simpler, whilst still maintaining a number of key features; notably Kato's cusp condition at short range and the exponential decay at long-range [3].

The omission of these radial nodes can have implications on the short-range accuracy of the STOs. The long-range behaviour of this model, however, is generally considered accurate and is therefore useful for a great number of orbital calculations. The most common argument against using STOs is that despite being a simplification, they are still, in a number of ways, too complicated for large scale integration. For a while STOs were somewhat antiquated, although they have recently found a new popularity in many *ab initio* calculations (where a high computational cost is generally expected) because of their high-level accuracy coupled with improved algorithms for evaluating multi-centre integrals [82, 83]. However, during this project Gaussian type orbitals (GTO) were the preferred means of defining orbital functions. The reasons for choosing to use GTOs in this research are described in the following section.

3.3.3 Gaussian Type Orbitals

To understand the features of a full GTO a reasonable place to start is with a so-called un-contracted Gaussian primitive (GP) [3], the form of which is:

$$G(x,y,z) = (x-R_x)^l(y-R_y)^m(z-R_z)^n e^{-\alpha(r-\mathbf{R})^2} \tag{3.41}$$

This formulations of a GP defines the value of the orbital function as being equal to the product of four terms. The first three are the Cartesian variables x, y and z, after being corrected for the orbitals centre position, R_x, R_y and R_z, and raised to the power of the respective x, y and z component of angular momentum, denoted l, m and n. The fourth term defines the exponential component and is raised to the power of the orbital radius, r^2 (again corrected for the orbital's centre $\mathbf{R}$), multiplied by a value α, which is some constant.

A single GP has a poor level of accuracy, particularly when compared to the STO function in equation (3.38). To improve their performance a full GTO is defined as a sum of GP functions that can be more accurately matched to the relevant STO:

$$G(x,y,z) = N\sum_{i=0}^{n} d_i(x-R_x)^l(y-R_y)^m(z-R_z)^n e^{-\alpha_i(r-\mathbf{R})^2} \tag{3.42}$$

In this expansion the normalization, N, and a term dubbed the contraction coefficient, d, are also added as part of the sum to complete the STO equivalence. For completeness it is necessary to add that the radial and spherical separation have not been mentioned in this section, as in this formalization of a GTO the spherical information is not given in terms of a straightforward spherical harmonic. Instead the information regarding orbital shape is described in the first three terms whilst the factor of π (that is normally necessary when calculating orbital geometry) is contained within the normalization. At first glance this expression would appear to be more complex than the STO function but the real benefit of working with GTOs lies within expanding the r^2 term and following the algebra to a natural simplification. Working in Cartesian coordinates the radial term is equal to:

$$r^2 = (\sqrt{x^2+y^2+z^2})^2 = x^2+y^2+z^2 \tag{3.43}$$

This means it is possible to separate equation (3.42) into three one-dimensional components:

$$G(x,y,z) = N\sum_{i=0}^{n} d_i(x-R_x)e^{-\alpha(x-R_x)^2}(y-R_y)e^{-\alpha(y-R_x)^y}(z-R_z)e^{-\alpha_i(z-R_z)^2} \tag{3.44}$$

This now represents a significant reduction in the computational cost needed for integration, as many one dimensional integrals are substantially easier to complete than a single complex three dimensional calculation. In fact, as will be seen in section 3.4, these can be solved using formulas that are straight forward to evaluate. There is still a question of how many GP functions are necessary to accurately represent a full GTO. Theoretically the more GPs that are added to the function the more accurate it becomes, however, these have diminishing returns (the correction added from each GP becomes smaller and so is less significant) and so the improvement becomes insubstantial beyond a certain point. The actual definition of that point is chosen by consensus and the modern standard is to use six GP functions in the creation of a GTO, the subsequent orbital function is commonly refereed to as the STO-6G basis set. To give a demonstration of the effect of increasing the number of GPs, figure **??** displays the STO for hydrogen with along with three fitted GTO orbitals, one representing the best fit for a single GP, the second using three GPs and the third utilising six. The graph shows the increasing accuracy of the GTO with the increasing number of GPs, and also demonstrates the long range accuracy of the type of GTO function used in this research.

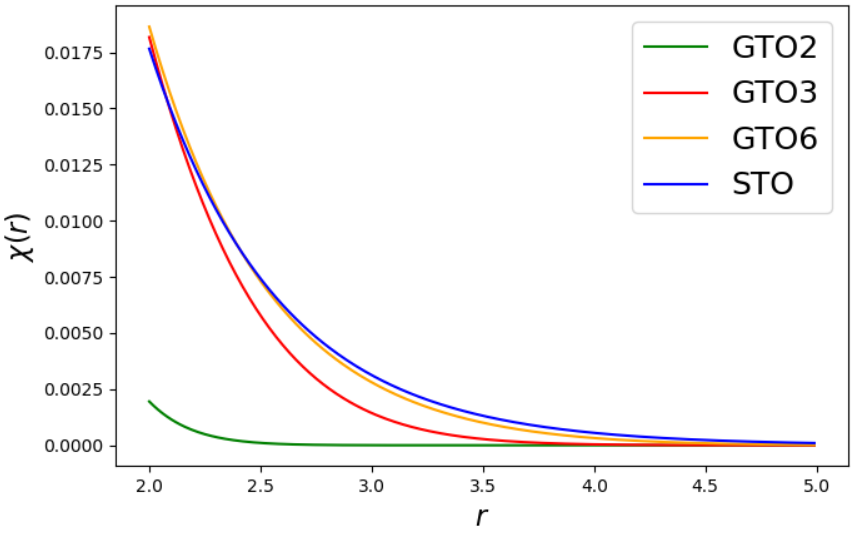


Figure 3.2: A graphical comparison of the accuracy of three separate s1 hydrogen GTOs against an s1 hydrogen STO. Each GTO is constructed with a different number of GPs. This displays the increased equivalence that occurs when a greater number of GPs are used in constructing a GTO.

3.4 Guassian Orbital Manipulations

To maximise the potential efficiency increases available when selecting GTOs, it is first necessary to consider the behaviour of these functions when combined and integrated. This section will examine the calculations that are performed when manipulating GTO functions and present a picture of the computation that is implemented when completing orbital calculations.

3.4.1 Normalisation of GTOs

The coefficients for GTO functions are designed so that if each individual uncontracted GP is normalised, the whole function will also be normalised. The reason for this construct is simply to guarantee suitable contributions from each GP. The result is that any program involving a GTO simply has to consider the normalisation for each GP individually and the entire function is automatically normalised, following the use of well established coefficients. Starting from the general definition of normalisation, if the orbital function is denoted $\chi(r)$ and its conjugate $\chi^*(r)$, normalisation defines the condition that the integral of their product, over all space, is equal to one.

$$\int \chi(r)\chi^*(r)d\tau = 1 \tag{3.45}$$

Using the definition of a GP centred at the origin, $\mathbf{R} = 0$, equation (3.41) can replace the χ terms in equation (3.45) to give:

$$\int N^2 x^l y^m z^n e^{\alpha r^2} x^l y^m z^n e^{-\alpha r^2} d\tau = 1$$

$$=> \int N^2 x^{2l} y^{2m} z^{2n} e^{-2\alpha r^2} d\tau = 1 \tag{3.46}$$

At this stage the separability of the Gaussian equations can be utilized to arrive at the following expression:

$$N^2 \int x^{2l} e^{-2\alpha x^2} dx \int y^{2m} e^{-2\alpha y^2} dy \int z^{2n} e^{-2\alpha z^2} dz = 1 \tag{3.47}$$

Given that each integral is performed independently, a closer examination of just the x component results in the following equation:

$$\int x^{2l} e^{-2\alpha x^2} = \frac{(2l-1)!!\sqrt{\pi}}{(4\alpha)^l \sqrt{2\alpha}} \tag{3.48}$$

The full derivation of this particular integral can be found in [84]. Given that the z and y components are functionally similar, their results can be extrapolated from the expression above:

$$N^2 \frac{(2l-1)!!\sqrt{\pi}}{(4\alpha)^l \sqrt{2\alpha}} . \frac{(2m-m)!!\sqrt{\pi}}{(4\alpha)^m \sqrt{2\alpha}} . \frac{(2n-n)!!\sqrt{\pi}}{(4\alpha)^n \sqrt{2\alpha}} = 1 \tag{3.49}$$

therefore:

$$N = \left(\frac{2}{\pi}\right)^{\frac{3}{4}} \frac{2^{(l+m+n)}\alpha^{(l+m+n+3)/4}}{\left[(2l-l)!!(2m-m)!!(2n-n)!!\right]^{\frac{1}{2}}} \tag{3.50}$$

Following the qualitative description of the construction of a normalised GTO at the start of this section, a fully normalised GTO is simply:

$$G(x,y,z) = \sum_{i=0}^{n} N_i d_i (x-R_x)^l (y-R_y)^m (z-R_z)^n e^{-\alpha_i (r-\mathbf{R})^2} \tag{3.51}$$

Equation (3.51) describes the normalised GTO orbital as the sum of a set of normalised GPs, when combined with the contraction coefficients d_i. Although this expression is complete it is still possible to apply a further simplification to increase the efficiency of implementing this calculation in the program. Noticing that only the contraction coefficients d_i and the exponents α_i vary across the summation it is possible to rewrite the expression as follows:

let

$$N' = \left(\frac{2}{\pi}\right)^{\frac{3}{4}} \frac{2^{(l+m+n)}}{\left[(2l-l)!!(2m-m)!!(2n-n)!!\right]^{\frac{1}{2}}}$$

therefore:

$$G(x,y,z) = N'(x-R_x)^l (y-R_y)^m (z-R_z)^n \sum_{i=0}^{n} \alpha_i^{(l+m+n+3)/4} d_i e^{-\alpha_i (r-\mathbf{R})^2} \tag{3.52}$$

3.4.2 Gaussian Product theory

A significant component of polyatomic-molecule calculations is the necessity of computing complex integrals of orbital products with various operators. The previous section, which discussed the orbital choices, argued in favour of applying GTOs in this program for the explicit reasons that they simplify these integral calculations. A major factor for this simplicity is as a result of the Gaussian Product theory (GPT) [85]. GPT is a mechanism for reducing the multi-centred integrals that arise from orbital calculations to being based around a single point. Simply put, the theory postulates two arbitrary angular momentum GPs centred on positions $\mathbf{R}_1$ and $\mathbf{R}_2$ can be a written as a single GP centred on $\mathbf{R}'$ [86]:

$$\chi(r)_{\mathbf{R_1}} \chi(r)_{\mathbf{R_2}} = K \chi(r)_{\mathbf{R}'} \tag{3.53}$$

The process followed in demonstrating the relationship in equation (3.53) is initiated by examining the product of two s-type GP functions, i.e. the angular momentum component is equal to 0, $L = n = l = m = 0$. Under these conditions it is concluded that the RHS of equation (3.53) is equal to $Ke^{-\gamma(r-\mathbf{R}')^2}$, where K, γ and $\mathbf{R}'$ have the form:

$$K = e^{\frac{\alpha_1 \alpha_2 (\mathbf{R}_1 - \mathbf{R}_2)^2}{\gamma}} \quad , \quad \gamma = \alpha_1 + \alpha_2 \quad \text{and} \quad R' = \frac{\alpha_1 R_1 + \alpha_2 R_2}{\gamma} \tag{3.54}$$

Removing the assumption that the angular momentum necessarily be zero, $L_1 = n_1 + m_1 + n_1$ and $L_2 = n_2 + m_2 + n_2$. The above relationship for γ and $\mathbf{R}'$ can be used to derive a general formula for the product of two GPs:

$$\chi(r)_{\mathbf{R_1}}\chi(r)_{\mathbf{R_2}} = KI_xI_yI_ze^{-\gamma(r-\mathbf{R}')^2} \tag{3.55}$$

Where the terms I_x, I_y and I_z are equal to:

$$I_x = \sum_{i=0}^{l_1}\sum_{j=0}^{l_2}(x - R'_x)^{i+j}\binom{l_1}{i}\binom{l_2}{j}(R_{x_1} - R'_x)^{l_1-i}(R_{x_2} - R'_x)^{l_2-j} \tag{3.56}$$

$$I_y = \sum_{i=0}^{m_1}\sum_{j=0}^{m_2}(y - R'_y)^{i+j}\binom{m_1}{i}\binom{m_2}{j}(R_{y_1} - R'_y)^{m_1-i}(R_{y_2} - R'_y)^{m_2-j} \tag{3.57}$$

$$I_z = \sum_{i=0}^{n_1}\sum_{j=0}^{n_2}(z - R'_z)^{i+j}\binom{n_1}{i}\binom{n_2}{j}(R_{z_1} - R'_z)^{n_1-i}(R_{z_2} - R'_z)^{n_2-j} \tag{3.58}$$

Although those expressions for I_x, I_y and I_z would appear to add complexity to the equations, the usefulness of GPT and the ability to rewrite the product of the two GPs becomes apparent when examining the more complicated integrals calculations.

3.4.3 Overlap Integral Calculations

The overlap calculations are the most computationally strenuous process performed in the Hückel method program. In fact, the advantages gained from choosing a Gaussian based formulation are almost entirely a result of needing to perform this operation. This section will outline how the GTOs simplify this procedure.

To detail these simplifications, it is reasonable to start from the general definition of the overlap for two orbital functions $\chi(r)_{\mathbf{R_1}}$ and $\chi(r)_{\mathbf{R_2}}$, which takes the form:

$$S = \int \chi(r)_{\mathbf{R_1}}\chi(r)_{\mathbf{R_1}}d\tau \tag{3.59}$$

The $d\tau$ term is referring to the integral being over all space. It is feasible using modern integral functions to evaluate the overlap with just the use of equation (3.59) and some orbital basis function. Even for relatively simple molecules, however, the computational effort for performing such calculations can easily become excessive. It is therefore useful to analytically simplify this expression. Using the GPT described in the previous section it was possible to obtain a relatively simple expression for the overlap of any two orbitals, based on their position and type alone [87]. Deriving this formula is relatively uncomplicated, thus starting from the definition of GPT, equation (3.55) can be used to transform the overlap expression:

$$S = \int \chi(r)_{\mathbf{R_1}}\chi(r)_{\mathbf{R_2}}d\tau = K\int I_xI_yI_ze^{-\gamma(r-R')^2}d\tau \tag{3.60}$$

Again the separability of the Cartesian components can be used to simplify equation (3.60):

$$S = KS_xS_yS_zd\tau \tag{3.61}$$

where:

$$S_x = \int I_x e^{-\gamma(x-R'_x)^2}d\tau \quad , \quad S_y = \int I_y e^{-\gamma(y-R'_y)^2}d\tau \qquad \text{and} \quad S_z = \int I_z e^{-\gamma(z-R'_z)^2}d\tau \tag{3.62}$$

and each component is a functionally similar expression equal to:

$$S_x = \int_{-\infty}^{\infty} \sum_{i=0}^{l_1}\sum_{j=0}^{l_2}(x-R'_x)^{i+j}\binom{l_1}{i}\binom{l_2}{j}(R_{x_1}-R'_x)^{l_1-i}(R_{x_2}-R'_x)^{l_2-j}e^{-\gamma(x-R'_x)^2}dx \tag{3.63}$$

It can be noticed that many of the terms in equation (3.63) do not depend on x and, therefore, do not have to be included in the integral, as such these are factored out to create:

$$S_x = \sum_{i=0}^{l_1}\sum_{j=0}^{l_2}\binom{l_1}{i}\binom{l_2}{j}(R_{x_1}-R'_x)^{l_1-i}(R_{x_2}-R'_x)^{l_2-j}\int_{\infty}^{-\infty}(x-R'_x)^{i+j}e^{-\gamma(x-R'_x)^2} \tag{3.64}$$

Note that any odd value of $i+j$ will result in an anti-symmetric integrand about the x-axis and, therefore, the integral term will equal 0. Letting $i+j=k$ will lead to the expression:

$$S_x = \sum_{i=0}^{l_q1}\sum_{j=0}^{l_2}\binom{l_1}{i}\binom{l_2}{j}(R_{x_1}-R'_x)^{l_1-i}(R_{x_2}-R'_x)^{l_2-j}(\frac{\pi}{\gamma})^{\frac{1}{2}}\frac{(k-1)!!}{(2\gamma)^{k)/2}} \tag{3.65}$$

for all even values of k. Similar results for S_y and S_z are also obtained by following the same procedure. Returning to the definition of a full contracted GTO in section 3.3, equation (3.44), the expression for the overlap can be expanded to represent a full GTO as follows:

$$S = N_1N_2\sum_{i=0}^{n}\sum_{j=0}^{n}Kd_id_j\int I_xI_yI_ze^{-\gamma_{ij}(r-R')^2}d\tau \tag{3.66}$$

Hence following the overlap is calculated as being equal to:

$$S = N_1N_2\sum_{i=0}^{n}\sum_{j=0}^{n}Kd_id_jS_{x_{ij}}S_{y_{ij}}S_{z_{ij}} \tag{3.67}$$

in which each overlap term $S_{x_{ij}}$, $S_{y_{ij}}$ and $S_{z_{ij}}$ are defined according to equation (3.65). The result in equation (3.67) is now entirely general in giving the overlap of any two GTOs. Although the expression itself appears complex, in computational terms, it is in fact straight forward and can be solved precisely using basic mathematical operations. Thus, this method is advantageous over the use of any technique requiring more intricate or complex solving functions.

3.5 Calculation of Intermolecular Interaction Grids

The previous sections have been focused on exploring a range of the methods, ideas and theories for how to calculate and model molecular orbital properties. What is yet to be described is how this molecular information will be used to generate simulations and replicate molecular monolayers. The final sections of this chapter will hence detail the various ideas used for constructing molecular simulations.

The fundamental principle behind the simulations performed in this thesis is that each individual molecule will manoeuvre itself to minimize the orbital overlap between neighbouring molecules. This research seeks to use this basis to make predictions about the orientations and arrangements of different molecular and surface combinations. Thus, this section describes how the intermolecular overlap was estimated so it could be assigned to the program as a cost function.

The overlap between two electron orbitals, in a given region of space, is taken to be the integral of the combined electronic wavefunctions of the two molecules in that same region (equation (3.29)). Deriving such a function for entire MOs (and not just AOs) would require an immense amount of computation as it would necessitate knowing the full wavefunction of the system. However, since this program assigns the overlap to the procedure as a cost function, it is not necessary to know the exact overlap, only to have a function that scales correctly i.e. high valued outputs are produced when the orbital overlap is large and smaller valued outputs are produced when this property is reduced. From this consideration, the proposed method of approximating the intermolecular overlap between two adjacent molecules is as follows. Denoting the two molecules A and B, the MOs derived from the Hückel method are used to evaluate the contribution to the electron density from both A and B at a series of points on a grid placed between them. The product of each of those contributions, at each point on that grid, is calculated and stored into a variable. Finally, the sum of each data point in that variable is calculated and that total is taken to be the cost of that given molecular pairing. Detonating each grid points (i, j), the cost, C, of the interaction between the molecules A and B, in a given orientation, is described expressed as being equal to:

$$C = \sum_i \sum_j A(i,j)B(i,j) \tag{3.68}$$

where $A(i, j)$ and $B(i, j)$ represent the electron density contribution at each grid point and the total interaction is calculated as the sum of the product of each contribution. Whilst this is a relatively crude approximation of the real overlap, this maintains the correct scaling features that make it an applicable cost function. At points on the grid where the contribution to the overlap from the molecules A and B are both higher, the value of the point on the plane where they intersect will also be higher; vice versa for points where the contribution from each molecule is lower. Thus, the cost of each calculation scales according to the overlap.

Even in this reduced scale, continually repeating this calculation of the overlap (equation (3.68)) during simulations would become computationally taxing. As a result, for the simulations to be implemented efficiently, each possible overlap value needed to be calculated and stored in a so-called grid-file beforehand. The process of which requires generating a set of orientations for a pair of molecules and using equation (3.68) to create an easily accessible look-up table for each possible grid value. The exact process this involves is dependent on the molecule being investigated, as such, the details of the grid-files made in this research are described, in detail, when discussing the molecules themselves.

This section describes an almost identical procedure to that which was used in the 2016 paper by J Leaf *et al* [5]. The work described in that paper forms the foundation on which this research is based. In their research, these methods were being applied to layers of buckminsterfullerene (C_{60}), instead of the planar phthalocyanine monolayers being investigated for this thesis. Thus, a key distinction between the two programs is that researchers preforming the calculations on C_{60} molecules had to consider all the possible rotations in three dimensions, whereas this study assumes the molecules rotate in the x-y plane about their centres. Therefore, in some regards, this version is less computationally intensive when compared to the calculations that were required to simulate C_{60}. Despite these differences, the program those researchers created demonstrated that applying this approximate overlap calculation to the system as a cost function has the potential to be a useful and powerful tool for generating simulations. This assertion is demonstrated by both the simulations performed in their work and those presented in the later chapters of this thesis.

3.6 Orientation Minimization Methods

This next section will conclude the discussion in this chapter by exploring the final component necessary to complete the simulation program, the simulation techniques themselves. The previous sections have, so far, discussed the molecular properties and how they are calculated but, to generate simulations, these must all be combined into a set of algorithms that can, in theory, represent molecular behaviour.

The general principle in defining how molecules self-assemble is derived from the assumption that they are attempting to reside in some minimum energy state. Therefore, the process of finding the final orientation is a case of modelling that interaction energy and deducing the molecular positions that minimise this interaction. The various depths and complexities of how these simulations are implemented can have an important impact on how the molecular complexes behave.

3.6.1 Iterative Minimisation Method

When initially constructing the simulation a basic iterative minimisation algorithm was implemented to represent the minimisation process. The main reason for its inclusion was as a simple mechanism for testing other parts of the pro-

gram before moving on to more sophisticated ideas. However, displaying its results and its shortcomings are still worthwhile as they provide insight into why a more involved technique was needed. The principle operation of the algorithm was to take a set of randomly ordered molecules on a surface and iterate over the monolayer, selecting each molecule in turn. The size of the interaction could then be checked, for that chosen molecule, in all available alignment positions. The program would then be able to determine the position that minimises the interaction of that selected molecule with its nearest neighbours. That minimum position would then be the new position of the molecule and the programme would progress onto the next molecule. In the event that no new minimum is found after iterating through all possible azimuthal angles, no change would occur. The programme would then run, continuing to iterate over each molecule, before repeating over the whole layer enough times to reach some overall minimum set of orientations for the monolayer i.e. until the simulation is unable to find any new positions that reduce the interaction between any molecules.

The advantages of this idea are its speed and simplicity. The algorithm is easy to implement and is generally reduced to a minimum position in fewer than five iterations over the entire monolayer. Whilst this idea is computationally favourable there are two fundamental problems such an approach. Firstly, on a practical level, this idea will often fail to find a global minimum and instead reside in a local minimum where it is incapable of improving. The second key fault with this algorithm is related to the theoretical interpretation of the simulation it produces. The motion of molecules on a surface when settling into alignment positions is not iterative but naturally chaotic. Therefore, the model should display some degree of randomness, as would be observed with real molecular systems. With these considerations a different mode of simulations was necessary.

3.6.2 Monte Carlo and Metropolis Monte Carlo

The final section in this chapter will discuss the chosen computational methods for simulating the self-assembly of molecular monolayers. These are the so-called Monte Carlo algorithm and its extension the metropolis Monte Carlo algorithm [88]. Monte Carlo algorithms are employed in a large number of simulation techniques from cash-flow projections algorithms to numerical integration [89, 90]. All implementations of the algorithm follow the same broad principles; a large sampling of some variable with probabilistic outcomes is tested and used to predict an overall trend or outcome of these random events. In this body of work the probabilistic event is the orientations of a set molecules on the surface. Conceptually, the motion of a plane of molecules positioning themselves into a low energy configuration resembles that of a canonical ensemble, in which each microstate of the system has a probabilistic outcome. Monte Carlo algorithms are generally considered an appropriate choice for this kind of molecular simulation because they are good at mimicking the process of a canonical ensemble [5].

The 'standard' Monte Carlo simulation can be applied to a molecular monolayer, as a tool minimize the interaction with respect to their orientations, as follows. A molecular monolayer is initiated and the orientations are set to be in some ran-

dom configuration. A molecule is then selected, at random, by the algorithm and angularly repositioned by rotating it by some random amount. The cost of this new position is then tested against the previous value; if the interaction potential has been reduced, this new position is kept, otherwise the program takes no action. At this point a new molecule is selected and the process is repeated until enough iterations of the algorithm have been performed to reduce the system into some minimum state.

Generally, this is an effective tool for representing canonical ensembles, however, real systems are prone to random fluctuations as a result of thermal energy, that in turn cause small increases in interaction potential. Thus, this version of a Monte Carlo algorithm lacks some details which would make it more representative of real molecules. Whilst that concept is described as a difference in theoretical interpretation, there is a practical basis for improving on that particular simulation routine, as the inability of the simulation to increase the cost of the interaction make it prone to getting caught in local minima. In an effort to overcome these shortfalls the metropolis Monte Carlo algorithm differs from the original method by instead of resoundingly rejecting any values that increase the cost function, they are dealt with probabilistically. Using the definition of a canonical ensemble [91], the probability of transition between microstates is given by:

$$P = e^{-\frac{\Delta E}{k_B T}} \tag{3.69}$$

This function defines a probability by taking the ratio of the difference between, the final and initial states, ΔE, and $k_B T$ is a scaling factor used in thermodynamics as a representation of some basic energy unit of the system. More generally, the relationship $\frac{E}{k_B T}$ is a representation of the entropy per molecule of the system in natural units and it is from this definition that the relationship between these factors and the probability of transitioning between microstates is derived, a more detailed description of the thermodynamic perspectives of these ideas can be found in [91]. The result of equation (3.69) is that the probability of transitioning into a new given microstate decreases exponentially the larger the value of the energy of the new state.

The implementation of the metropolis Monte Carlo algorithm is as follows: starting with the same procedure as the standard Monte Carlo routine, a monolayer is generated with randomly oriented positions. Then a molecule is selected at random and rotated by some random amount; if the cost of this new interaction is less than the cost of the original, the new position is selected. Under the metropolis Monte Carlo algorithm, if the interaction of this new position is greater than the original, the choice of whether it is kept is handled probabilistically. The transition probability of that current microstate is calculated according to equation (3.69) and compared to a random event. In the program used for this thesis, this event was a number, r, randomly generated between 1.000 and 0.000 that represented a probability. If the probability value generated by r is less than the probability of transferring into this microstate, the new state is kept i.e. if $P(\Delta E) > r$, otherwise the original state is maintained. Again the program terminates after a chosen number of iterations is implemented.

By creating a probabilistic selection process it allows the system to have small increases to the cost function, improving the chances of settling on a global minimum. An important note on this model is that given the energy value ΔE is not the precise energy of the monolayer but a relative measurement that scales similarly, the metropolis temperature in equation (3.69) is, therefore, not the real temperature of the system but an arbitrary parameter that has to be set. The drawback of this method is that it takes more iterations to find a stable position and so is more computationally demanding in terms of the time it takes to run. Implementations of both versions of the Monte Carlo algorithm were created, and so discussed here but only the metropolis Monte Carlo algorithm was used for the results on display the later chapters, as it was the more robust technique that produced more consistent results.

Chapter 4

Application of the Simulation Program to Benzene

All of the simulation algorithms used in this research were written, from scratch, purely for this body of work. Given that the software required the construction of many components, the evolution of the program had a number of developmental stages. At the expansion of each new set of functions the molecule benzene was used to test the proficiency and effectiveness of each new section of code. Benzene is a simple molecule containing just twelve atoms; six carbons and six hydrogens in a six-fold aromatic planar structure. It also has the property of being well researched with its features assiduously documented. In combination these features made benzene an ideal candidate for testing various components of the simulation program. In this chapter the results of these testing phases will be presented, primarily because they demonstrate that the simulation algorithms are correctly implemented. A secondary reason is to detail each subsequent component of the software and explain its role within the program, with clarity, before discussing the research itself.

4.1 Molecular Orbital Calculations of Benzene using the Basic Hückel Method

The initial phase of development began with applying a BHM routine to benzene. Many of the stages are exactly the same for both the BHM and EHT programs and bear distinctly similar results when applied to benzene. It was, therefore, useful to create an algorithm for performing BHM calculations as a forerunner to the EHT routine, acting as a prototyping tool for the functions that were necessary to create in both programs. It also has the benefit of being able to provide a check on the answers outputted by the EHT functions, as these are expected to produce similar results.

4.1.1 Stage 1: Labelling each atom

The BHM Hamiltonian is constructed so that each orbital's interaction, with any other orbital, is a separate matrix element. This means before it is possible to construct the Hamiltonian it is imperative to consider how to label each atomic

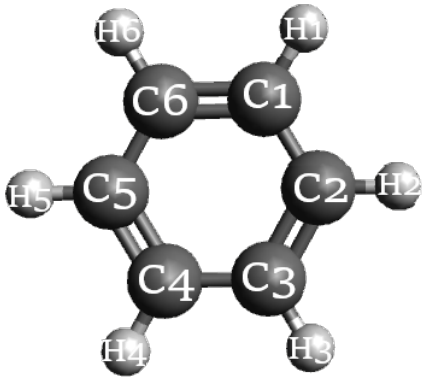


Figure 4.1: A ball and stick diagram of a benzene molecule showing the label of atoms used in the BHM and EHT programs.

orbital so it is possible to state which interaction term is being examined.

Figure 4.1 shows a ball and stick model, of benzene, displaying how the atoms have been labelled in this program. The hydrogen atoms are also labelled here as these will be used in the next section. For a molecule such as benzene, this step is trivial to complete but it is still necessary and can, in other circumstances, be more complex. An example of a more intricate labelling system would be one used for performing BHM calculations on C_{60} (a molecule that does generate accurate results in a BHM program, despite, not being planar [92]) as this is significantly more complex and necessitates more careful deliberation. In terms of the code, this labelling system translates to creating a series of functions that determine how the Hamiltonian matrix is constructed. In this instance, this process was completed by generating a two dimensional array in which each element of the array represented each term of the Hamiltonian.

4.1.2 Stage 2: Constructing the Secular Equation

For the BHM, the secular equation is as defined in equation (4.1):

$$H_{ij}c_i^{\mu} = E_i c_i^{\mu} \tag{4.1}$$

in which H_{ij} is the Hamiltonian and, c_i^{μ} and E_i are the LCAO coefficients and the energy states respectively. In equation (4.1) the Hamiltonian is needed to solve for the rest of the terms and so the key objective in this step is completed by evaluating H_{ij}. Computationally, this step is the process of defining the values for the two dimensional array generated to represent the Hamiltonian in step one. For the BHM, the Hamiltonian is defined according to equation (3.28), in which each element is equal to α for diagonal elements, β when defining the interactions between neighbouring atomic orbitals and zero otherwise. Using figure 4.1 as a reference, it can be established that the six carbons, labelled c1 - c6, will have nearest neighbours at $c(i+1)$ and $c(i-1)$; but will loop back to one when the number becomes greater than six and back to six when it is less than one. Thus, equation (4.1) can be written as:

$$\begin{bmatrix} \alpha & \beta & 0 & 0 & 0 & \beta \\ \beta & \alpha & \beta & 0 & 0 & 0 \\ 0 & \beta & \alpha & \beta & 0 & 0 \\ 0 & 0 & \beta & \alpha & \beta & 0 \\ 0 & 0 & 0 & \beta & \alpha & \beta \\ \beta & 0 & 0 & 0 & \beta & \alpha \end{bmatrix} c_i^\mu = E_i c_i^\mu \tag{4.2}$$

There are modules theoretically capable of resolving the above eigenvalue problem in its current form but normally this equation is simplified to speed up the calculation. By using the substitution $E' = -\frac{\alpha - E_i}{\beta}$ all algebraic terms can be removed from the Hamiltonian matrix, producing:

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} C^\mu = E'_\mu C^\mu \tag{4.3}$$

The α and β terms have simply been encapsulated into the transformed energy term E' to create an entirely empirical equation.

4.1.3 Stage 3: Solving the Secular Equation

The power of a BHM calculation lies in the efficiency of this calculation because the only real step, in terms of computation, is the diagonalization of the Hamiltonian in equation (4.3). Even for large molecules, this step could be performed in much less than a second on modern processors. Completing the necessary computation will output a series of eigenvectors and eigenvalues, which are equal to the solutions for the altered energy E' and the size of the components for the molecular orbitals C^μ respectively. Resolving for the energies, E, is now simply a case of rearranging the substitution applied in equation (4.3) and using the answers for E', the results are displayed in table 4.1.

Table 4.1: Table displaying solutions that result from the BHM program when applied to benzene

μ	E'	E	C	MO
1	2	$\alpha + 2\beta$	$\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}$	$HOMO - 1$
2	1	$\alpha + \beta$	$(-\frac{1}{2\sqrt{3}}, \frac{1}{2\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{\sqrt{3}})$	$HOMO$
3	1	$\alpha + \beta$	$\frac{1}{2}, \frac{1}{2}, 0, -\frac{1}{2}, -\frac{1}{2}, 0$	$HOMO$
4	-1	$\alpha - \beta$	$-\frac{1}{2}, \frac{1}{2}, 0, \frac{1}{2}, -\frac{1}{2}, 0$	$LUMO$
5	-1	$\alpha - \beta$	$(-\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{1}{\sqrt{3}}, -\frac{1}{2\sqrt{3}}, -\frac{1}{2\sqrt{3}}, \frac{1}{\sqrt{3}})$	$LUMO$
6	-2	$\alpha - 2\beta$	$(\frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{1}{\sqrt{6}})$	$LUMO + 1$

From table 4.1 it can be seen there are four unique energy levels, with two two-fold degenerate states lying at the HOMO-LUMO gap. The true numerical value

of those energy states can not be deduced from this calculation as the values of α and β are unknown. However, it is known that β always has a negative value, thus, the observation that each energy level differs by multiples of β allows the deduction of their order and degeneracy. Each carbon atom contributes one electron to the π network and each energy level requires two electrons to be considered full. The six electrons, therefore, occupy the three lowest energy states (the states in which β is an additive to the energy), the HOMO-LUMO gap consequently lies between third and fourth state.

4.1.4 Stage 4: Construction of Molecular Orbitals

It was mentioned in the previous section that the eigenvectors C in the BHM formulation represent the size of the components of the MOs. In mathematical terms, this value acts as a coefficient that dictates the prominence of the contribution of each AO to a particular MO in a LCAO method. Thus, using the LCAO method function (equation (3.23)) and the GTO formulation for a p_z orbital (equation (3.42), in which $l = m = 0$ and $n = 1$), the MOs, Ψ^μ, in a BHM routine are given by:

$$\Psi^\mu = \sum_{n=1}^{N} c_n^\mu G_{P_z}(x, y, z) = \sum_{n=1}^{N} c_n^\mu N \sum_{i=1}^{6} d_i (z - R_{zn}) e^{-\alpha_i (r - \mathbf{R}_n)^2} \tag{4.4}$$

In the above expression each MO, μ, is given as the sum of a series of AOs, represented by a GTO, and a coefficient, c_n, that determines the size of the contribution of the nth AO, to the μth MO. By taking the values for the exponent, α, and contraction coefficient, d, for a carbon p_z AO from basis set exchange [93, 94, 95, 96], the following variables can be created:

$$\alpha = 30.49723950, 6.036199601, 1.876046337, 0.7217826470, 0.3134706954, 0.1436865550$$

$$d = 0.0037596966, 0.0376793698, 0.1738967435, 0.4180364347, 0.4258595477, 0.1017082955$$

From the knowledge that each bond length in benzene is 1.4Å, it is possible to model each atomic position, $\mathbf{R}_n$, and draw a series of isosurfaces representing each MO. Such a plot is displayed in figure 4.2. The MO's are arranged in terms of energy from lowest to highest, with the isosurface images on the same level representing degenerate states. The blue sections show areas of negative polarity and red represents positive polarity.

4.2 Application of Extended Hückel Theory to Benzene

Working with a system that is replicable using the BHM is incredibly convenient as its efficiency means almost all of the computation needed for producing the images in figure 4.2 is used by the isosurface function itself. The work in this project was not able to apply a BHM routine to the simulations as certain assumptions

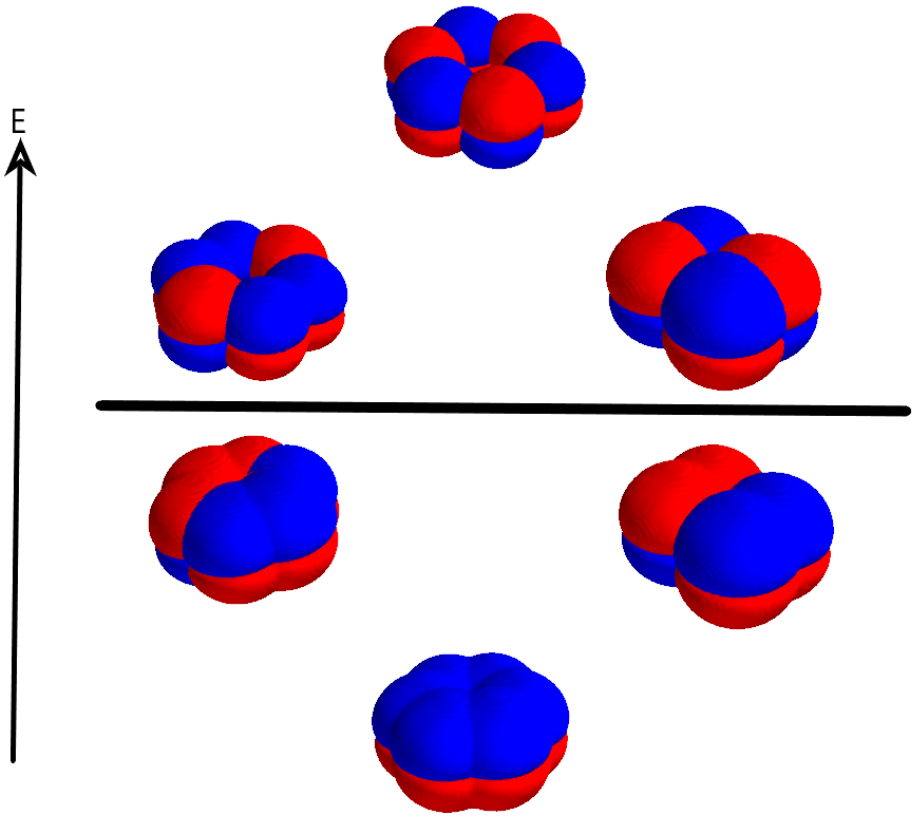


Figure 4.2: Isosurface plots of the molecular orbitals of Benzene as calculated from the BHM. These are arranged from the lowest energy to highest energy with the line indicating the HOMO-LUMO gap. The value of the isosurface value, for each MO function, was arbitrarily chosen to best represent the shape of the MO for each plot.

limit its use to hydrocarbons only. For the phthalocyanine molecular calculations, this meant applying the extended version of the BHM.

The EHT application follows an almost identical process to the BHM, so much so, significant parts of the code were implemented in both programs. The differences between the two occur predominately with the construction of Hamiltonian, where the EHT methodology is more rigorous and involved than it was for the BHM. Primarily this step now requires the inclusion of all valence orbitals and the explicit calculation of the overlap integrals [8]. Recalling from the previous chapter, each element of the EHT Hamiltonian is calculated according to equation (3.30):

$$H_{ij} = K\frac{H_{ii} + H_{jj}}{2}S_{ij} \tag{4.5}$$

For this function, K is taken to be 1.75 [73] (this value of K was used for all calculations in this thesis) and the VSIP terms for carbon and hydrogen, H_{ii} and H_{jj}, are given as:

$$H = -0.500$$
$$C^{2s} = -0.768$$

$$C^{2p} = -0.419$$

These values were taken from Skinner and Pritchard [97] and here they are presented in atomic units.

The overlap matrix S_{ij} is by far the most complex term to calculate in this process. The explanation for this can be seen by examining the calculation of a single element of this matrix. Using figure 4.1 as a reference, the example to be investigated here, will be the overlap between the H2 *s*-orbital and the C2 *s*-orbital (elements S_{28} and S_{82} in the formulation used in this program). By setting the centre of the molecule to being the origin, C2 and H2, in angstroms, are centred at $(1.212, 0, 0)$ and $(2.312, 0, 0)$ respectively. Inputting these values, along with the orbital angular momentum information, into the overlap equation (equation (3.67) from the previous chapter) results in the following calculation:

$$S_{17} = N_1 N_2 \sum_{i=0}^{n} \sum_{j=0}^{n} d_i d_j K S_x S_y S_z \tag{4.6}$$

$$S_{17} = N_1 N_2 \sum_{i=0}^{n} \sum_{j=0}^{n} d_i d_j e^{\frac{\alpha_1 \alpha_2 ((23.12)-(12.12))^2}{\gamma}} (\frac{\pi}{\gamma})^{\frac{3}{2}} \tag{4.7}$$

For the STO-6G basis set, that is used for all overlap calculations in this thesis, the above computation requires thirty six different calculations to be completed and summed. Individually each calculation is still relatively efficient and easy to complete but to deducing the whole overlap matrix is costly as this needs to be computed for every matrix element. For benzene, which has thirty valence orbitals, the resulting overlap matrix has dimensions 30x30, which equates to 900 positions in the matrix that must be evaluated. By choosing two *s*-orbital types, equation (4.7) has been reduced to one of the more simplistic versions of this calculation found in the routine. The calculations involving *p*-orbital types have extra factors, another double sum and two choice functions that must be computed, increasing the number of necessary individual calculations for each element by a factor of two in these cases. Despite the added complexity in this methodology, the overlap calculation is still quick to complete when applied to benzene, only taking around three seconds to compute the entire matrix. In proportion to the rest of the program, however, the overlap accounts for roughly 99.9% of all required computation in terms of time taken.

Once that information has been derived the Hamiltonian can be constructed and diagonalized with relative ease. The complexity of the functions required for the overlap calculation limits the possibilities for finding efficient tricks or exploits in the code. The Hamiltonian, on the other hand, can be derived easily by combining a consideration of the combinatorics of the different VSIP values with the effective use of concatenation modules. Considering equation (4.5) with specific attention paid to the terms $H_{ii} + H_{jj}$, it can be seen that as there is a limited set of VSIP values for a given molecule (three for benzene) so there will be a narrow number of possibilities that sum can produce. Completing this step of the program can be done with great efficiency by devising a system to block fill each of possible value of this sum using concatenate functions. The full Hamiltonian

can then be easily calculated by multiplying by the necessary constants and taking the product of each value with each respective entry in the overlap matrix. Finally, using the same routine as the BHM program, an eigenvalue finding function is employed to diagonalize the Hamiltonian and produce both the energy levels and molecular orbital coefficients.

Implementing this program, for benzene, produces an output of thirty energy levels, each with its own set of thirty MO coefficients. These matrices are too large to comfortably display in this section, for completeness, a version of the important results from this calculation can be found in Appendix II, here they will be discussed. A number of important results can be gleaned from the information these provide. First of all, there is almost no mixing between the p_z AOs and the rest of the orbitals, demonstrating consistency with the BHM hypothesis. Secondly, very similar orbital results to the BHM orbitals in figure 4.2 can be found amongst the results in the EHT routine. Their distinctness is a consequence of the p_z orbital separation, the MO values that resemble the BHM MO terms have a similar value for each p_z orbital term but are zero otherwise. The opposite is true of those MOs that uniquely result from the EHT calculation, those levels are zero for each p_z orbital coefficient and non-zero otherwise. Finally, if the HOMO-LUMO gap is worked out by filling the MO levels up with the number of valence electrons ($(6 * 4) + 6 = 30$ for benzene), it is found to lie between the same two states derived from the BHM. Thus, it can be concluded that these methods are showing consistent results and are therefore a testament to the accuracy of the EHT program.

4.3 Benzene in an STM simulation

The next stage of implementation, in terms of the development of the program, was to create the necessary functions in order to simulate STM images. The start of the previous chapter discussed the ideas behind modelling STM currents (section 3.1) and the function selected to represent it, equation (3.9). This equation states that, for an s-type atomic tip, the observed current is proportional to the square of the sample wavefunction calculated at the tip's centre i.e. $I \propto \psi(\mathbf{r_0})^2$. Although the exact value of the current is unknown in that equation, the function is still applicable to generating simulations as the production of a contour plot is only dependent on the variable in question having the correct proportions at each point. Thus, this relative current, $\mathbf{I}'$, will create a plot similar to the true current assuming the proportionality is correct. $\mathbf{I}'$ is therefore given by:

$$\mathbf{I}' = \sum_{i}^{n} \Psi_i^{\mu}(\mathbf{x}, \mathbf{y}, \mathbf{z})^2 \tag{4.8}$$

In which $\Psi_i^{\mu}(\mathbf{x}, \mathbf{y}, \mathbf{z})^2$ denotes the square of the sample wavefunction for each MO i at a particular energy level μ. To accurately represent the sample wavefunction the degenerate states must be added together, n, is therefore the number of degenerate states.

4.3.1 Constant Current vs Constant Height simulation

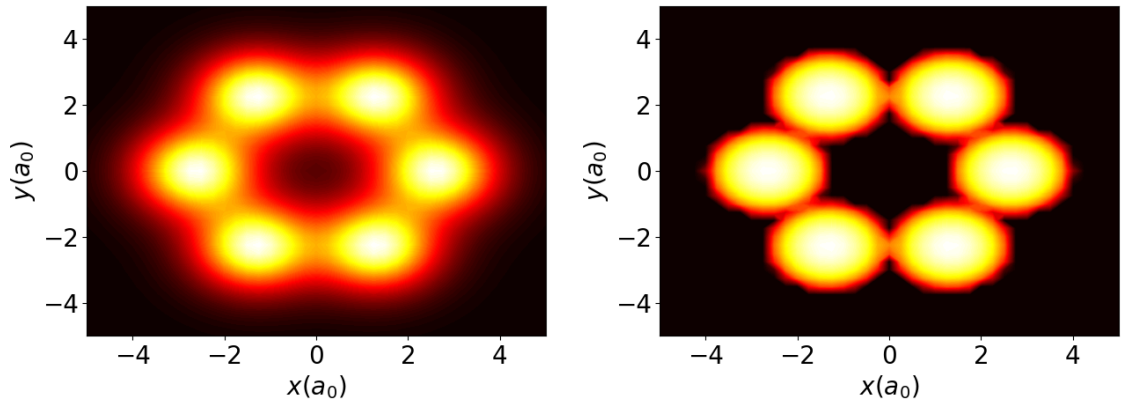
At this point, it is necessary to discuss the two modes of conducting STM experiments as each one would change the method required to resolve (4.8). The two modes are constant current and constant height respectively. Constant current mode describes an STM experiment in which the tip changes height in order to maintain a constant current, whereas constant height mode is when the height of the STM tip remains stable whilst the current is measured. From the simulation perspective a constant height mode is calculated by deriving the current according to equation (4.8) at each point $\mathbf{x}$ and $\mathbf{y}$ with z remaining constant. The constant current mode is more tricky to determine as this requires implementing some solving function to determine z at each point $\mathbf{x}$ and $\mathbf{y}$ for a given value of $\mathbf{I}'$. A comparison between the two different types of calculations is displayed in figure 4.3, for each mode both a singular and molecular monolayer image is presented. From an experimental perspective, constant current mode is generally the preferred option, as this provides more detail. However, in terms of simulations, those solving functions are costly and so it is worth considering if they are necessary.

For the duration of this study, there are no molecules being investigated that are not planar in structure, therefore, the difference in detail between simulations is minimal. The extra computation required for performing constant current simulations was, as a result, deemed unnecessary and constant height mode calculations were instead exclusively performed in chapters five and six. Comparing the two images of benzene as a singular molecule figures 4.3a and 4.3b, it is noticeable that the two images are distinct from each other however on the scale of monolayers those differences become negligible, as can be seen in figures 4.3c and 4.3d. If the research focused on non-planar molecules this choice would not be possible (as was the case for the work presented in [5]), however, in this instance, it is reasonable and appropriate to approach the simulations in constant height mode.

4.4 Simulating 2D Molecular Assemblies

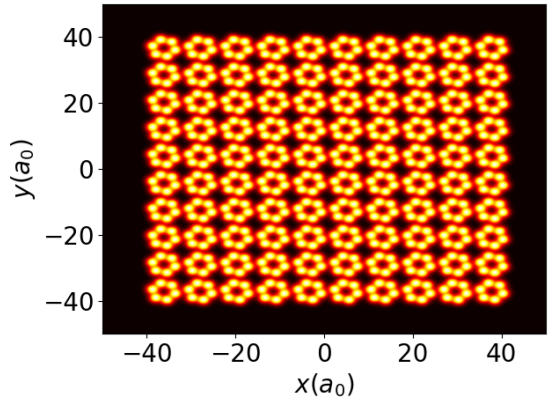
The last set of functions in the program are those related to the surface simulations itself, namely, the metropolis Monte Carlo algorithm. By using the EHT calculations from section 4.2 it was possible to demonstrate this procedure and provide an example of its effectiveness as a minimization method.

Before being able to perform any simulations the monolayer is itself must be initialized. A key part of the research presented in chapter six is a series of investigations into how different intermolecular structures on the surface affect the orientation of the molecules. For that to be possible it was important to be able to select the positions of each molecule on the surface. Given that the maximum distance between two of the atoms in a benzene molecule is around 4.7Å, a reasonable choice for the intermolecular spacing is to centre each molecule on a square grid of points, each separated by 5Å. Once the centre positions have been ar-

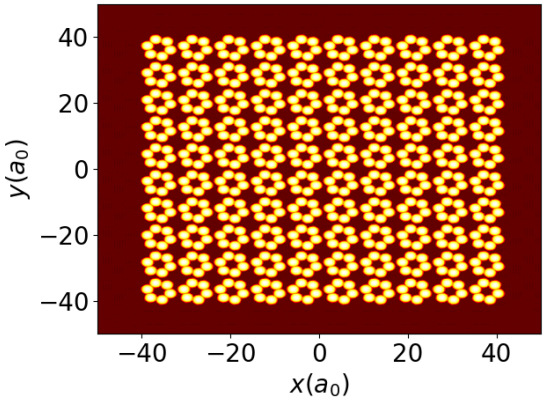


(a) Constant height

(b) Constant current



(c) Constant height



(d) Constant current

Figure 4.3: Depiction of images comparing constant current and constant height simulations of benzene, both singularly and as part of a molecular monolayer. The constant current simulations were obtained height value of $z = 2.5$Å, the current in the constant current simulations was fixed at $2.2mA$. Both sets of images are greatly resemblant of the images obtained from STM experiments of benzene, thus, these images can be considered consistent with real world data [98].

ranged the set of MOs occupying a particular energy level (usually the HOMO or LUMO, in this case the HOMO) can be selected to represent each molecule. The monolayer construction is then simply a case of translating a copy of this MO function to each centre position to create a surface of molecules. For benzene, either the HOMO or LUMO are entirely constructed out of p_z orbitals, and so, despite the planarity of benzene, the distribution of a p_z orbital largely protrudes out of the plane of the surface. This means that the overlap between those orbitals on neighbouring molecules is likely to also be out of the plane. With this consideration the size of the interaction grids placed halfway between each molecule (see section 3.5) was set to have a size equal to 5Å$\times$5Å to cover all possible contributions to the molecular overlap. Figure 4.4 displays an example of the grid situated between two benzene molecules.

A useful method for aiding the efficiency of the program can be found by considering the symmetry of the molecule in question. The six-way symmetry of benzene means, for a given intermolecular spacing, the molecular rotation only

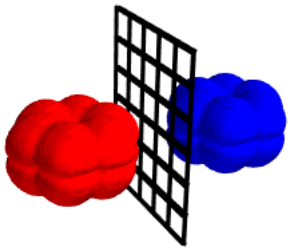
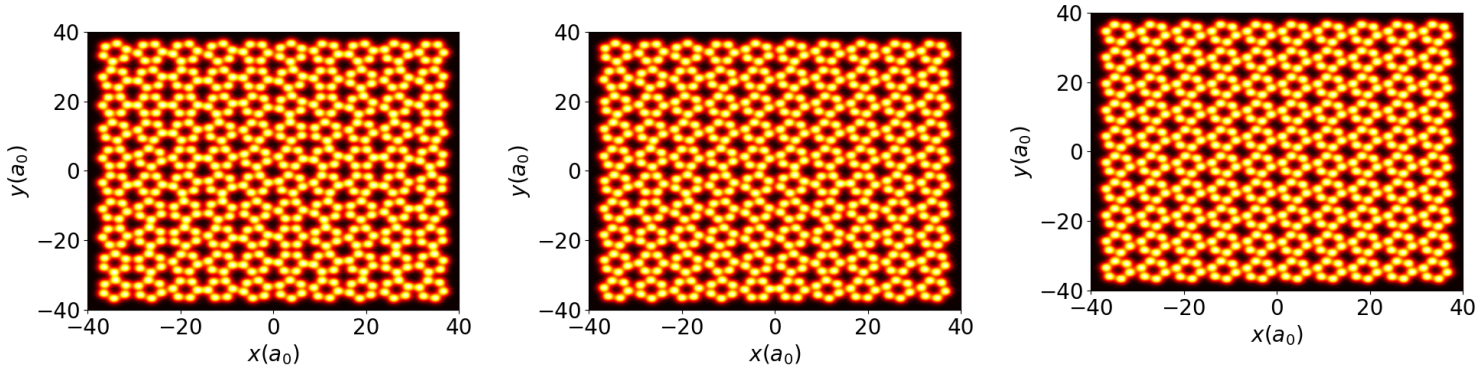


Figure 4.4: A 3D model representing the placement of a set of grid points in the middle of a pair of benzene molecules in which the distance between the two molecules, as well as the height and width of the grid is 5Å

needs to be calculated up to 60° before the grid values will repeat themselves, reducing the number of calculations by a factor of thirty-six. To create the grid-file the program examined all possible pairs of orientations from 0° to 60° at 1° intervals, creating a grid-file consisting of 3600 entries. This grid-file, along with the surface construction described above, were used to produce the simulations displayed in figure 4.5.

The images of the benzene molecules themselves, in figure 4.5, portray a set of stages in the minimization process, figure 4.5a displays the random ordering at the start of the simulation. The second image, figure 4.5b, depicts a set of partial ordered molecules, found after performing a hundred thousand Monte Carlo iterations. The final image of the benzene monolayer, figure 4.5c, displays the monolayer after the minimization algorithm has performed ten million cycles and the resulting molecules are fully ordered into a minimum orientation. Before discussing the effectiveness of this routine it must be noted that the preceding example is likely to be an unrealistic model of benzene molecules on a surface. Whilst the exact behaviour of benzene when being bound to a surface is dependent on a number factors, such as the temperature of the experiment and the strength of the surface interactions, it is common for benzene rings to form bonds [99, 100], which is a process not accounted for in this model. Despite this discrepancy in the physics, it is possible to use the results of the two graphs in figure 4.5 to comment on the performance of the minimization process. In the final state each molecule is orientated at an angle of 14° or 13°, indicating that the system has found some minimum orientation for the molecules that lies between 14° and 13°. This assertion is validated by the graph in figure 4.5e. This graphs shows what the total interaction energy of the monolayer, $E(\theta)$, would be if every molecule was orientated at an angle θ i.e. $E(30)$ is the sum of all the interactions between each benzene molecule, when each benzene has been rotated to 30°. Figure 4.5e shows that the minimum point of this calculation occurs at $\theta = 14°$, indicating that the set of angles found by the metropolis Monte Carlo simulation represents the orientation that minimises the interaction potential of

the system. The graph on the left hand side shows how this minimum is reached; as the number of iterations increases the cost function decreases exponentially mimicking a conical ensemble. Hence it can be concluded that the algorithms are implemented correctly and are, therefore, applicable to systems that favour being orientated into configurations that minimises the intermolecular overlap.



(a) Benzene molecules in a random orientation

(b) Benzene molecules in a partially ordered orientation

(c) Benzene molecules in an ordered state

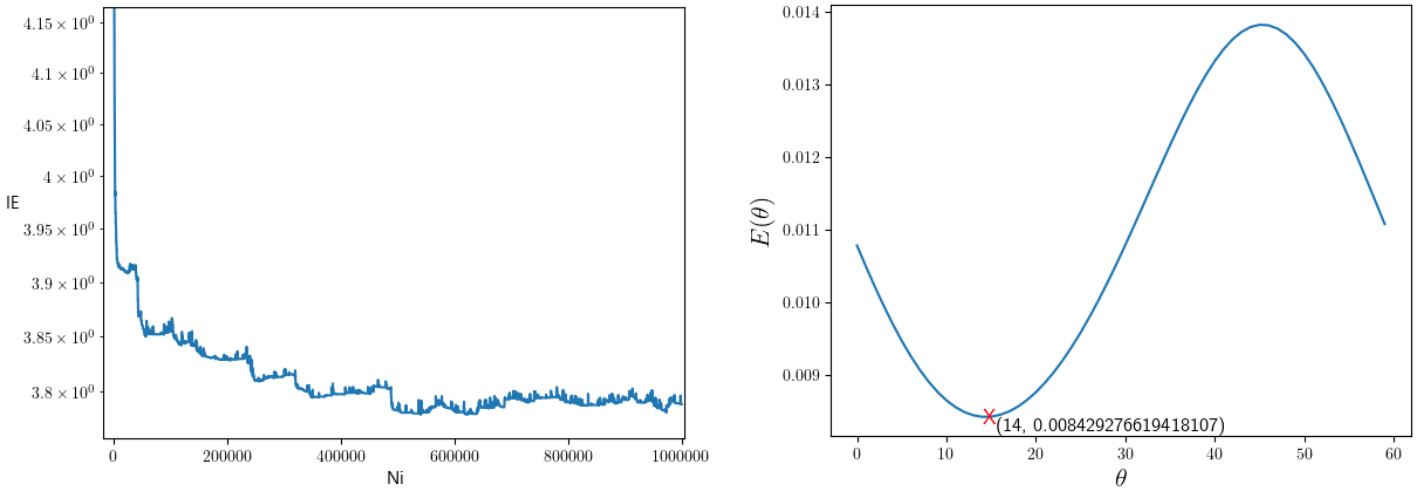
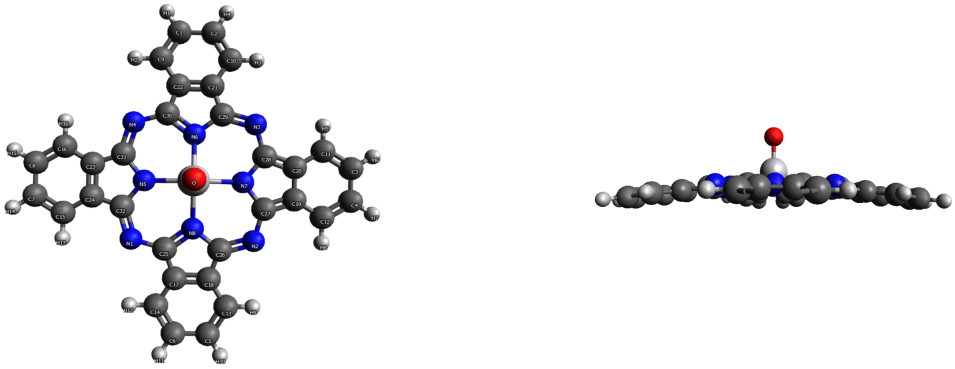


Figure 4.5: The top presents a series of STM simulations of a benzene monolayer as the molecules move from a disordered to an ordered state, over the course of the simulation. The bottom row displays two graphs, the first showing the interaction values against the number of iterations demonstration the exponential decay of the interaction. The second displaying the interaction values for the molecules against a given angle (assuming neighbouring molecules also occupy that same angle) with an X marking the minimum on the curve

Chapter 5

Model of Titanyl Phthalocyanine on a Gold Surface

The previous chapters have, so far, described the processes this research used to model molecular properties and how these properties can be utilised to implement a monolayer simulation. Here, those techniques will be applied to the more complex phthalocyanine molecules by examining, in detail, the case of titanyl phthalocyanine (TiOPc). Initially, this chapter will focus on the results of applying the EHT to TiOPc before examining the crucial step of including a surface interaction and discussing various intricacies of that process. Finally, this study will combine those models and simulation techniques to investigate the self-assembly of TiOPc molecules on a gold surface. Then a comparative analysis shall be presented between the ideas in this model and those used by Mannsfeld and Fritz in 2005 [6].



(a) Top view

(b) Side elevation

Figure 5.1: A ball and stick model portraying TiOPc in two elevations, one viewing from above and one from the side. The left hand image shows the overall shape formed by the molecular arrangement including the four-fold symmetry. The right hand image of the side elevations and displays the curvature across the molecule as well as the position of the oxygen atom directly above the titanium.

5.1 Calculation of TiOPc MO's

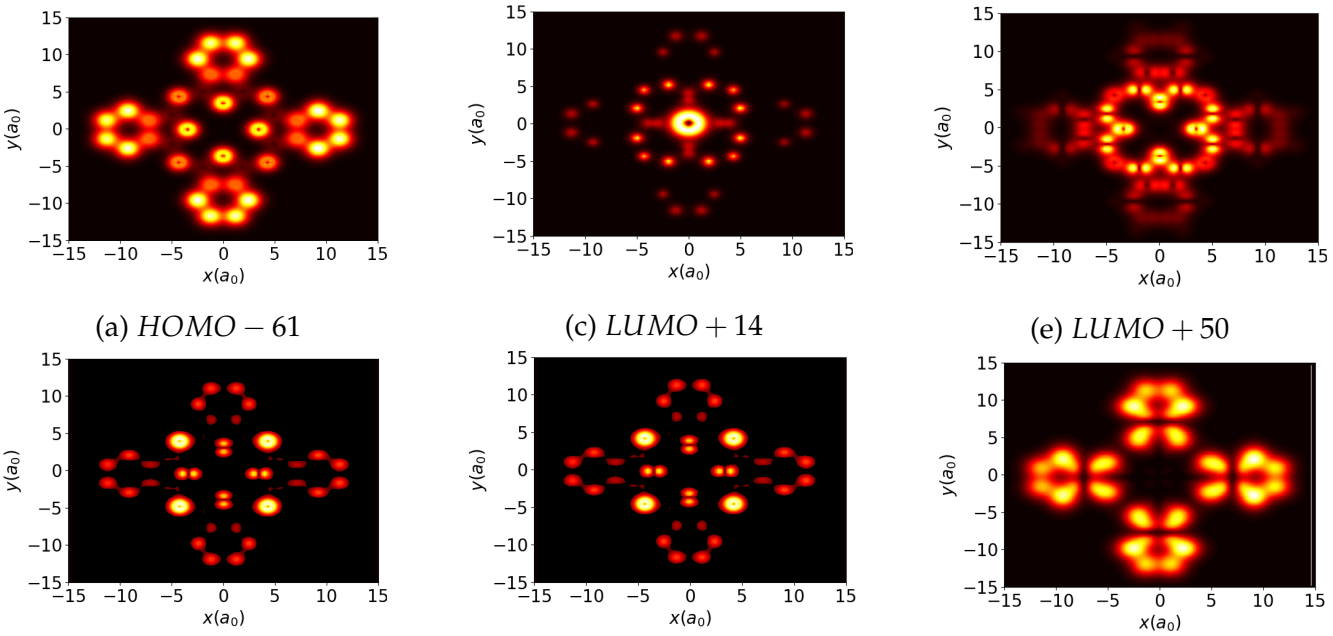
Applying the EHT to the phthalocyanine molecules generally follows the same procedure that was applied to benzene in the previous chapter. However, there are significant additions in this case. These are: the increased complexity that comes from the growth in the number of orbitals being processed, phthalocyanine having at least 176 valence orbitals compared to benzene's 30; the incorporation of a larger variety of AOs, including those from other elements; and the inclusion of d orbitals in the calculation, which are necessary when considering heavier metallic elements. These do not affect the principle method of the EHT calculation, instead, they simply increase the computation required to complete it. As before, each atom is labelled creating a system by which the relevant matrices are initialized. These can then be combined with information relating to the atomic structure to derive the overlap matrix, using the equations described in section 3.4. This calculation, like it was with the benzene example, is the most complex and computational expensive part of the routine, accounting for 99.9% of all the time taken. In this case, however, when applied to the larger phthalocyanine molecules, the program required approximately two minutes to complete, a significant increase from three seconds that was required to calculate the overlap for benzene. The Hamiltonian of the system is then easily assembled and diagonalized using the same functions that incorporate the overlap matrix into the Wolfsberg-Hoffman approximation, equation (3.30).

Considering the particular EHT parameters of TiOPc, it is found there are 189 valence orbitals necessary for calculations: nine orbitals are bound to the titanium atom, five d, three p and one s; four attributed to the oxygen atom; 32 associated with the eight nitrogen atoms bonded into the centre of the molecule; each of the thirty-two carbon atoms add four each to the valence shell, making their total contribution 128 orbitals; and each of the 16 hydrogen atoms contributes one orbital to the valence shell. Unlike the example of benzene, the inter-atomic spacings of TiOPc are not readily available and so had to be calculated before it was possible to deduce the overlap. The inter-atomic distances were found for TiOPc (and the other MPcs described in this thesis) by constructing an interactive model using the Avogadro chemistry program (application version 1.2) [101]. After building the molecular model inside the program a geometric optimization was then performed using the Merck molecular force field [102] to determine some minimum state positions of the atoms. Once completed, these results were used to determine the overlap matrix. The Hamiltonian VSIP values for the carbon and hydrogen orbitals are valued at the same level given in chapter four. The newly added nitrogen, oxygen and titanium VSIP values were found in the appendix of *Electrons and Chemical Bonding by Harry B Gray* [103] and scaled into atomic units. This meant deriving a conversion factor by scaling the hydrogen VSIP value in [103] from 13.6 eV to -0.5 and then taking the product of that factor ($\frac{5}{186}$) with each of the other VSIP values to produce:

$$H = -0.500, \quad C^{2s} = -0.768, \quad C^{2p} = -0.419$$

$$N^{2s} = -0.938, \quad N^{2p} = -0.482, \quad O^{2s} = -0.584, \quad O^{2p} = -1.19$$

$$Ti^{2s} = -0.269, \quad Ti^{2p} = -0.141, \quad Ti^{3d} = -0.346$$



(a) $HOMO - 61$ (c) $LUMO + 14$ (e) $LUMO + 50$

(b) $HOMO - 16$ (d) $LUMO + 30$ (f) $LUMO + 62$

Figure 5.2: A display of a selection of the molecular orbitals generated by running the EHT program on TiOPc. The figures are presented as 2D STM images obtained from running a constant height simulation at $z = 3$Å.

The EHT Hamiltonian is then produced by inputting these parameters into equation (3.30), this is then diagonalized in order to solve for the MOs.

The product of this particular EHT routine is a diverse set of electronic states for TiOPc. There exists 189 states in total that are reduced to 141 unique energy levels after incorporating degeneracies. The HOMO and LUMO states occupy the $70th$ and $71st$ energy levels respectively. The range of configurations varies with the electron density, in some cases being more concentrated around the central regions, while in others towards the edges of the molecule. A sampling of various MO's that outline this variety is displayed in figure 5.2.

Such a diverse range of states is a result of the complicated π-bonding network that allows for a large amount of delocalization of the electrons. As for the validity of these states, there are few publications on the exact formation of the free-standing MO's for TiOPc, derived from the EHT or otherwise. Instead, the following arguments are presented in favour of the statement that these MOs are a 'good' model of a TiOPc molecule. Firstly the program is demonstrably correct when applied to benzene and given that phthalocyanines are a suitable candidate for the EHT procedure, it follows that these states are likely to be representative of the system. A second assessment is that the images produced from the density calculations closely resemble real STM images seen in experiments, arguably a superficial quality but a useful indication of their validity. Finally, and crucially, as the sections below will demonstrate the model produces simulations that reflect real experimental data, a feat that would not be possible if the system did not

resemble the real states of the molecules. In a continuation of this study, the correctness of these states would be more effectively checked with the aid of some more robust procedures, ideally DFT and the Self-consistent extended Hückel method.

5.1.1 Comparison to Experimental STM images

It has been mentioned that the STM simulations of TiOPc, and other MPc molecules, resemble those generated from STM experiments. This section aims to demonstrate that point by providing a detailed comparison between a set of experimental STM images of TiOPc, and those generated by the simulation program being used for this research. For this comparison a simulation of a single TiOPc occupying the HOMO, as calculated from the EHT routine, is generated in figure 5.3, using constant height mode at a z value of $z = 2$Å. The z values were chosen arbitrarily in order to generate images that had the best resemblance to those found experimentally.

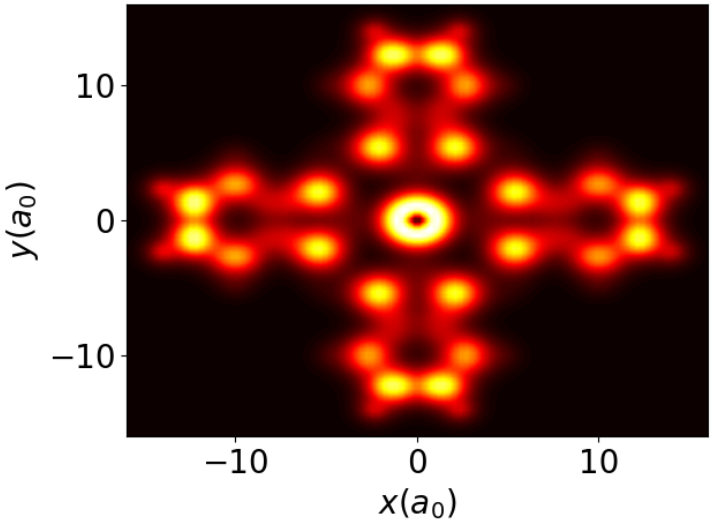


Figure 5.3: STM simulation of a single TiOPc molecule occupying the HOMO, as derived from the EHT calculation. The simulation performed in constant height mode at $z = 2$Å. Ideally the STM images taken in [6] would be portrayed here to present a direct comparison between the EHT model and experimental images, however, copyright limitations prevented a reproduction of the image.

The experimental images of TiOPc, that are presented in the Mannsfeld and Fritz paper [6], display two modes of TiOPc, one with a bright centre and one with a dark centre. This difference is likely a result of the direction of the oxygen atom; where there are bright spots, the oxygen atom is pointing away from the surface; where the centre is dark, the oxygen atom is pointing towards the surface, as discussed in the paper in question [6]. Each phthalocyanine in these simulations has been modelled as if the oxygen atom is facing away from the surface, which in turn explains the bright spot in the centre of figure 5.3. The experimental images display the rest of the molecule as a comparatively dull cross, the experimental resolution is not high enough to generate any of the specific detail that is on display in figure 5.3. In the case that experiments could generate that level of detail,

the dark patches in the middle of each isoindole arm would likely not be observed in reality, which creates a significant discrepancy between the simulation and the experimental images. However, both sets of images have a bright centre, from which four symetrical arms are extended, creating the cross shape that is seen in real STM images. It can, therefore, be concluded that these images are, in general terms, a reasonable approximation of those found experimentally [6].

Whilst the EHT has a number of advantages it is limited by not accounting for the electrostatic repulsion in any specific way. In phthalocyanine molecules the orbital structure would, in reality, give rise to a number of special properties, such as the d orbital splitting predicted by crystal field theory [104]. These properties are not currently within the scope of the program and is an inherent limitation on the predictive power of the model, the inclusion of these effects would likely result in an image that produces a more representative simulation of TiOPc in an STM experiment. A self-consistent routine, as described at the end of section 3.2, could be used to account for this and other phenomena but, as will be seen in the monolayer simulations, this current method is still a powerful tool for large scale applications.

5.2 Intermolecular Interaction Grid-Files for Pthalocyanine

Recalling the basis for how the overlap is estimated and minimized, this section will describe the algorithm that calculated and stored the grid-files that were used for every simulation of phthalocyanine. The routine for calculating the grid-files was constructed so that that the intermolecular interactions would be determined over a range of distances. This was so it would cover all the possible values that were needed for the analysis presented in chapter six. The process for this calculation was as follows.

The molecule is initiated in the central position at the origin, as displayed in figure 5.4. The initial grid starts parallel to the y-axis on the line $x = x_0$, x_0 being some distance in the positive x direction away from the centre of the molecule. Each grid consisted of 100×100 evenly spaced points encompassing a plane of area 15Å$\times$15Å. The electron density, as derived from EHT calculation described in the previous section, was then computed at each point on that grid for that single molecule and stored. That grid was subsequently rotated 1° about the centre of the molecule, this new grid was then calculated in the same manner and also stored. Considering the four-fold symmetry of phthalocaynine (shown in figures 5.2 and 5.3), this is repeated until the grid has been rotated 90° about the molecule to account for all unique grid values.

At this stage, having calculated all possible grid values for single degree increases, the total set of all possible pairs of grids is determined. The process for this starts by selecting each pair of grids in turn. When selected, each corresponding point on that grid-pair is multiplied and the sum of that set of products, as is described in equation (3.68), is stored into a two-dimensional array of size 90×90.

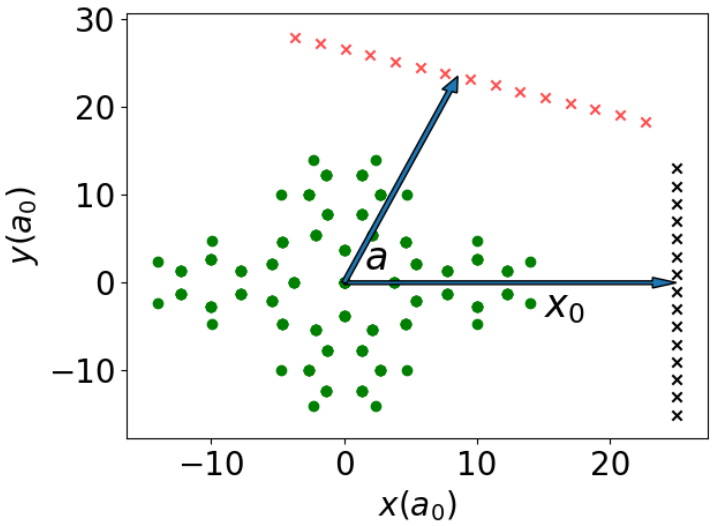


Figure 5.4: Depiction of the calculation process for the phthalocyanine grid-files, in which an initial grid is a distance x_0 away from the centre. A grid rotated a° about the centre of the molecule is also displayed to demonstrate the process of obtaining subsequent grids.

After all grid pairs at a distance x_0 have been evaluated, the two-dimensional array these grid value pairs is being stored in, is incorporated into a much larger array so the first index of this larger array corresponds to the distance away from the centre of the molecule. The distance of the initial grid, x_0, is then increased by some increment, Δx, for the final set of grid-files used in this research Δx was equal to 0.1Å. The algorithm would then complete another iteration of the grid pair calculation but now starting at the distance $x_0 + \Delta x$ away from the molecule's centre. This is then repeated until a significant amount of distance had been covered and all eventualities for modelling were accounted for. For these molecules, x_0 ranged from 5Å to 30Å . For the latter, the grid values had all fallen below the 30 decimal place precision and rounded to zero. It was thus concluded that this was a reasonable point to end the procedure.

5.3 Creating a Surface Model

The simulations thus far have been missing a key consideration: the surface. None of the molecules displayed in this research exists as a monolayer in isolation but has to be bound to a surface to be studied, particularly with regards to the use of STM techniques. These interactions can, and often do, play a significant role in determining the behaviour of the molecular monolayers. This poses two questions, namely, how can the surface be modelled and what is the most effective way to deduce the surface-molecule interactions.

The more theoretically taxing consideration, in terms of the two surface issues being examined, is the creation of the surface model itself. There are numerous different approaches to creating a surface model that include the use of concepts like plane-wave solutions and topology [105, 106]. However, such ideas would

add significant levels of complexity that would, in turn, diminish the advantages of using a Hückel based approach. Instead, the choice was made to consider an orbital based model, in which each atom in the plane of the surface is represented by an orbital function. Modelling the surface with this idea is consistent with the modelling processes used for the molecular monolayer and maintains the computational efficiency that allows for the simulation parameters to remain flexible. As was the case when considering the molecule-molecule interactions, the surface functions are assigned to this simulation as a cost function to be minimized. From this basis, it was decided to model each surface orbital function as a singular valence state p_z orbital extending above the plane of the surface. This approximation of the surface orbitals is based on two assumptions. Firstly, that there is little mixing between p_z orbitals and the other orbitals present on the surface and, therefore, they do not contribute to bonding in the plane. Basing the surface orbitals entirely on p_z orbitals also requires it to be postulated that these orbitals are the largest singular contributor to the electron density directly above the plane. The picture being painted, in this surface model, is one where the mixing between d, s and other p orbitals results in the electronic density contribution, from these orbitals, being approximately equal to a constant across the surface. This creates the circumstances in which only the p_z orbitals fundamentally contribute to the differences in electronic density around each atomic position on the surface. Thus, in terms of cost, it would only be the p_z orbitals that have an impact on the simulations. This model is crude in its simplicity and does ignore many complex surface effects, such as the herringbone reconstruction for gold surfaces [107]. These are, however, long range and subtle effects that, in theory, will have a relatively minor impact on the simulations. More generally, how well this approach works in practice can be assessed from the quality of the results produced by the monolayer simulations.

These approximations create a surface that is simple to describe mathematical and quick to model computationally. It just requires information regarding the bond lengths of surface atoms and the crystal structure to create an array of centre points to represent each surface atom. Referring back to the idea that these interactions are assigned to the program as a cost function, a full three dimensional model of values accounting for every point in the position space is not necessary. Instead, all that is required is an interaction grid between the surface and the molecules, as was performed in calculating the intermolecular interaction grid-file. Note that this grid doesn't change at any point in the simulation program, as the surface does not move or change shape, it just needs to be calculated once and stored. To make this efficient, the grid was calculated by placing a p_z orbital at each atomic centre with the advantage being taken of python's in-built vectorised operations to quickly deduce the density contribution of that single orbital at every point on the grid. This calculation was then repeated for each atom in the surface, with the value of each grid point above the surface being continuously evaluated at each iteration over the surface atoms. If each grid point is represented by x_i and y_j and the constant height is represented by z_0, the contribution from the surface to each grid point, $S_{x_i y_j z_0}$, can be calculated as follows:

$$S_{x_i y_j z_0} = \sum_k \chi_k(\mathbf{x}, \mathbf{y}, z_0) \tag{5.1}$$

where χ_k is an orbital function that represents the valence state p_z AO of whichever element comprises the surface. The orbital function coordinates, $\mathbf{x}$ and $\mathbf{y}$, are labelled in a different notation to the surface grid points to indicate they represent the position of every point on the surface grid simultaneously. The size of the grid only needs to cover slightly more space than all the possible positions of the molecule as the exponential term of the orbital function causes these values to quickly decay towards zero. The efficiency of the program is dependent on the number of points in the grid and the number of surface atoms. Given that the size of the surface is determined by the molecule, the number of atoms in that surface is essentially fixed. The number of grid points is, however, selected by considering the balance between the improvements in accuracy against the decreases in computational efficiency that result from increasing to the number of grid points. In the grid-files that were used to generate simulations, the grid was chosen to be of dimensions 100×100, which is large enough to determine the distribution in the surface orbitals accurately, whilst remaining quick to calculate. The actual time taken to compute the surface grid, using equation (5.1) and the given grid parameters, was around a second. Whilst that particular calculation can be performed with extreme ease, this is just the first step of deducing the molecule-surface interaction grid-file. The selection of the number of points is the biggest factor in deciding the amount of computation required for calculating the surface. This choice, however, also impacts the subsequent stages of the program. While this calculation is straight forward and could easily be performed for a much larger array of points, the latter stages of the surface-molecule grid routine could quickly become impractical. It was from consideration of how to perform those calculations, later in the program, that the number of points used to model the surface was derived.

5.3.1 Surface Orbital Basis Set

By choosing an orbital model of the surface it was necessary to give some consideration to the particular functions that would represent the relevant surface orbitals. For the comparative analysis performed in section 5.6, it was necessary to simulate a gold surface to replicate the conditions of the experimental data this model is being tested against. To remain consistent an STO-6G minimal basis set would have ideally been selected, as these have been exclusively chosen the work thus far. For gold, however, there are issues with making this choice, as no gold STO-6G coefficients (or any other STO-nG coefficients) are published in the basis set exchange or other sources, this is likely due to issues regarding accuracy. These minimal basis sets allow for little electron correlation in their formulation [108] and as the atomic structures increase in size these effects become more significant. These accuracy issues are the probable reason why no STO-nG basis sets are posted on the basis set exchange beyond the fifth period.

Upon examination of the surface construction equation (equation (5.1)), it can be observed that no integration (or complex mathematical formulation) is involved

at any point. Therefore, all the previous advantages of using an STO-6G function are no longer applicable. For these reasons, the surface orbitals can easily be represented using a standard STO functions instead. For the case of gold, the valence p_z orbital are in the $6p$ shell [109], which was found to produce an STO function equal to:

$$S_{6P_z}(Au) = \frac{1.316146}{\sqrt{\pi}} r^5 e^{2.481152r} \tag{5.2}$$

It can be recalled that the STO-6G basis sets are designed to replicate the original STO, leading to the conclusion that these are, in many regards, a more suitable choice of orbital function that create a better representation of the surface.

5.4 Surface-Molecule Grid-File Calculations

By creating a surface model it is now possible to incorporate the interaction between the surface and the molecules into the program. For algorithmic efficiency, the surface-molecule interactions need to be pre-calculated at each possible orientation. This is for the same reason that an almost identical calculation is performed to create a grid-file for the intermolecular interactions, to make the Monte Carlo algorithm viable. The calculation itself is an altered version of the grid-file calculation from section 3.5 (equation (3.68)):

$$E = \sum_i \sum_j A_{x_i y_j z_0} \times S_{x_i y_j z_0} \tag{5.3}$$

Functionally this equation is identical to the intermolecular interaction calculation, but the notation has been altered to reflect that it is being used for the surface calculation. The term $A_{x_i y_j z_0}$ represents the contribution to the calculation from the molecule at each grid point, with $S_{x_i y_j z_0}$ denoting the input from the surface orbitals. The algorithm that computes the function was also implemented differently. Previously the molecule stayed at a fixed location and the grid changed position, for the surface grid-file procedure, this application was reversed. The grid points remained in a fixed location whilst the molecule was cycled through different positions and orientations in order to capture as many unique interaction values as possible. These positions were designed to reflect all possible combinations within a symmetry regime of the surface structure. All the surfaces used for this study had an FCC (111) structure, and so, a representation of how the structural symmetry can be accounted for is shown in figure 5.5.

The figure describes a triangular set of points based within atomic positions on the surface. The number of points in the figure is limited to ten to clearly show the pattern, the real calculation featured many more positions. Crucially, by centring the molecule within that symmetry regime, it is possible to account a large range of all the possible interactions that could occur between the molecule and the surface. This method, like the intermolecular grid-file calculation described earlier, is dependent on the number of calculated points being high for the simulations to be accurate. For the set of surface grid-files used for generating simulations, the number of points, positions, and orientations used were as follows. The electron

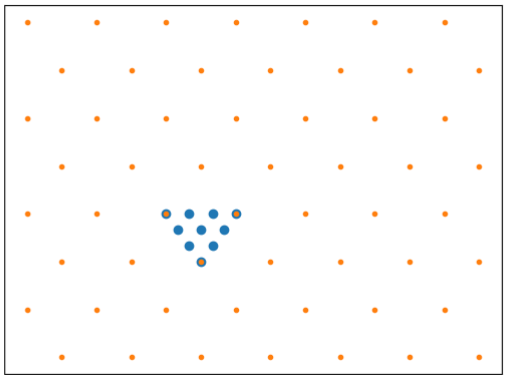


Figure 5.5: Representation of a possible layout of ten surface points that cover the symmetrical pattern of the FCC(111) surface.

density was calculated on 30×30Å grid containing 100×100 grid points, for a molecule positioned at 625 locations within the symmetry regime, in 90 different angular orientations.

For the example of a gold surface, which was used in the simulations later in this chapter, the (111) surface has an atomic spacing on the surface equal to 2.88Å. This number was found by taking the unit cell length for gold at 4.07Å [110], using Pythagoras theorem to obtain the length of the (111) side of the unit cell and halving that value to account for the FCC structure. Noticing that 625 is a triangle number of base 35, the distance between adjacent surface points is $\frac{2.88}{35} = 0.0822$Å. The scale of this number is important because when running the simulations the molecules are positioned on the surface according to a separate set of functions and ideas. Therefore, the molecules are unlikely to lie exactly on one of those points where the interactions were calculated. To account for this discrepancy the molecular centre positions were judged as being at the same locations as the nearest surface point used in the program. This created some error in the position where the interaction was calculated and the true position of where the molecule is, making it important to maximise the number calculable position. For the parameters used to generate a gold (111) surface the maximum error in the position of a molecule, randomly placed onto the surface, is 0.0581Å. This was found by noting that the maximum distance away from a calculable position is the centre point of an equilateral triangle with sides of length 0.0822Å. Given that value is much less than 0.1Å it is a safe assertion to say that approximating the molecule centre positions in this way will have little impact on the overall interaction values.

The number of points calculated in total to create this grid-file was 5.265×10^8, the product of 100×100 grid points, 90 angular orientations and 625 surface positions. Whilst a large amount of calculation the program was terminated on a laptop running with an Intel i7-7700HQ CPU processor in approximately four hours. In total the research required this set of functions to be computed fifteen times, each combination of three surfaces and five molecules used for the research, in-

dividually totalling a run time of sixty hours. However, the quad-core feature of the processor meant that three or four of those grid-file calculations could be efficiently run simultaneously. Thus, the time needed to create these files was reduced to being approximately between fifteen and twenty hours.

5.5 Surface Calibration

A consequence of neither the surface or the molecule interaction potentials being an exact calculation is that combining these two sets of values requires some form of the calibration process. Each singular interaction is dependent on a number of factors inherent in the program that are not real measurable events, such as the number of points used to calculate the grid. There are also variables, like the molecular height above the surface, that are reasonable estimates and not exact (discussed in section 5.7). Therefore, a mechanism for calibrating the surface-molecule and the intermolecular interactions is needed to create effective simulations. A number of ideas for this process will be presented and discussed in this section.

A reasonable sounding option, which was initially considered for the simulations, was to just normalize both sets of values using a min-max scaling feature:

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{5.4}$$

This describes the transformation of each individual value of a set, x, into a new value, $\hat{x}$, so every value in this new set is proportionately scaled between 0 and 1. By normalising each set of grid values, both the molecule and surface interactions are converted to having the same relative cost. Each set in this calibration now ranges from zero, at the minimum cost position, to one, at the maximum cost position. Hence, this removes the need to consider the actual calculated values.

The issue with using this feature becomes apparent when the scale of each set is observed. In the intermolecular interaction grid-file, the maximum value, x_{max}, is a factor of 10^8 times larger than its minimum value, x_{min}. For the surface-molecule interactions, that same calculation of the ratio between the maximum value and minimum value, x'_{max} and x'_{min} respectively, is equal to just 2 (the notation x' is being used to distinguish that these values relate to the surface interaction). The intermolecular interactions, therefore, reduce on a much larger scale than the interactions that result from the overlap of surface orbitals. By normalising both sets, each has an identical range which causes the surface potentials to be overcompensated for.

Given the difference in scale, between the molecular and surface interactions, creates a discrepancy when normalizing each set, a solution would be to maintain the range of each function, in terms of proportion, and set the minimums to be equal. The process is easily performed by a simple transformation that, for the purposes of this work, will be dubbed the surface-molecule scaler minimum

calibration method (SMCM):

$$\hat{x} = fx, f = \frac{x_{min}}{x'_{min}} \tag{5.5}$$

For two data sets where, the intermolecular interactions range from $x_{min} = 10^{-11}$ and $x_{max} = 10^{-3}$, and the surface interactions range from $x'_{min} = 2 \times 10^{-5}$ to $x'_{max} = 4 \times 10^{-5}$, the above transformation could be applied to the surface values so they become: $x'_{min} = 10^{-11}$ and $x'_{max} = 2 \times 10^{-11}$. From a consideration of the cost to the angular positions, this would prioritise the intermolecular interactions until the low energy regime is reached by the minimization process. The justification for applying this method of calibration is simply that it is capable of producing simulations that match experimental results, as is demonstrated in the following chapter. The issue with applying the SMCM calibration is that different surface constructions, that would normally have entirely different interaction potentials, become entirely similar, making it impossible to compare the effects of different surfaces. Therefore, a different more thoughtful process was necessary.

To make the cost function more effective at comparing surfaces, the idea was to perform some transformation on the surface data so it would span a certain amount of the intermolecular interactions values. In other words, this transformation would still leave both minimums equal, $x_{min} = x'_{min}$, but the maximum value of the surface interaction would be some amount between the intermolecular grids maximum and minimum value i.e. $x'_{max} = \hat{x}$ where $x_{min} < \hat{x} < x_{max}$. A function capable of meeting the criteria for performing that particular transformation is expressed as follows:

$$\hat{x} = (x - x_{min})\frac{(\hat{x}_{max} - \hat{x}_{min})}{(x_{max} - x_{min})} + \hat{x}_{min} \tag{5.6}$$

This takes a value, x, that lies between x_{max} and x_{min} and transforms it into a new value, $\hat{x}$, so its relative position is maintained in the new set of values lying between $\hat{x}_{max}$ and $\hat{x}_{min}$. Notice in this function that $\hat{x}$ is used and not x'. This is because x' has been used to denote surface values and in this function, $\hat{x}$ need not necessarily be the calculated surface values. These can, in fact, be chosen to represent any set required to be utilized for the research.

The key distinction with this new scaling method is that $\hat{x}_{max}$ can now be chosen to suit the nature of the surface interaction i.e. the range of the surface grids can be adjusted according to their varying strengths. For instance, if the research was comparing gold and silver surfaces and say a gold surface was measured to be twice as strongly interacting as silver, that could be accounted for by doubling the span of the gold surface compared to silver. This makes the transformation described in equation (5.6) a more effective way of drawing such comparisons. Although this provides an improved ability for incorporating a realistic surface interaction into the simulations, the choice of how to exactly fix the surface cost to the alignment cost is still somewhat subjective. This is because the nature of the types of calculations used are not exact, therefore, the 'best amount' for a given surface to span, is not easily quantifiable from the theory alone. For this reason, the next chapter uses a number of the simulation techniques along with experimentally gathered data to arrive at some reasonable spanning values.

5.6 TiOPc Monolayer Simulation and Experimental Comparison

With a method of combing the surface and intermolecular interactions in place, it is possible to use this set of simulation techniques to generate results that can be compared to experimental data. Testing the model, in this manner, would determine the validity of using this methodology and allow for the program to be adapted to make predictions about previously unknown systems. A 2005 study by Mannsfeld and Fritz [6] investigated the self-assembly of TiOPc by gathering both experimental and theoretical data in order to test the validity of an empirical approach as a simulation method. The methodology, used in that research, was functionally identical to the work performed here, making it an excellent choice for comparative analysis. Both programs used a pre-calculated grid-file as a look table to perform an efficient metropolis Monte Carlo algorithm, in order to generate simulations. The only significant difference being the modelling technique; this research is applying a Hückel method approach instead of the Lennard-Jones model utilized for the calculations in their work. To make a comparison between the two sets of data, the molecules are initiated according to the parameters used in [6]. The parameters in their work were, the angle of orientation between the surface and molecular monolayer (ψ), the spacing between adjacent molecules (b) and the angle between the rows of adjacent molecules (Γ). These were: 30°, 13.9Å and 90° respectively, as shown in figure 5.6.

These conditions create an overlayer structure that is incommensurate with the surface atoms. The resulting surface interactions are, therefore, mostly distinct for each molecule and the subject to changing entirely with small translations of the monolayer. As described in section 5.4, the program is designed to be able to handle issues that arise from this type construction by approximating the centre position of the molecule to the nearest surface point where the interaction was calculated.

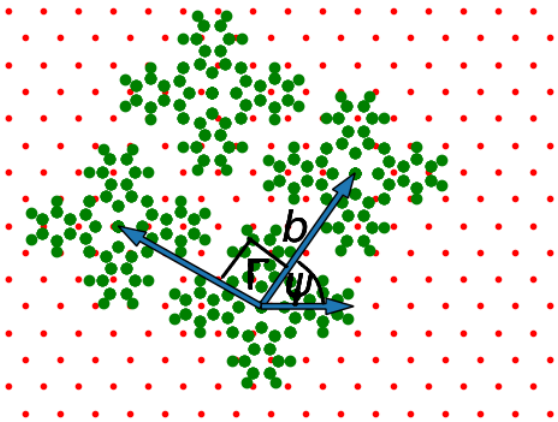
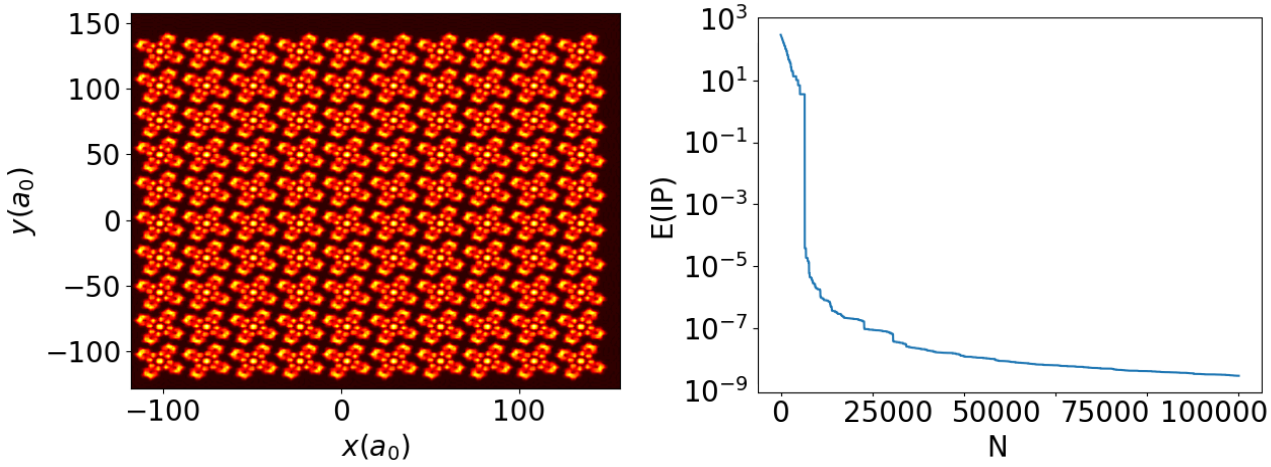


Figure 5.6: Topological representation of the simulation parameters described in [6], Γ represents the angle between rows, ψ the angle between the monolayer and the surface, and b is the intermolecular spacing.

A monolayer consisting of a 10×10 grid of molecules was constructed to those parameters, the height was measured from the central metallic atom and set to 5Å above the surface. The value of the height was chosen to be some reasonable amount. For molecules on the edge of the grid, the simulation was designed to use periodic boundary conditions i.e. the first row of molecules are simulated as if they lie between the next row and the last row of the monolayer. A similar condition is applied to the first and last column so that each molecule on the edge of the monolayer is simulated as if it has four adjacent molecules, like those that are placed in the centre. This is a standard technique for simulating small groups of molecules that would normally reside as part of a much larger set.

The surface calibration was adjusted to span up to a maximum value of $1000\times$ the surface maximum value that results from using the SMCM calibration i.e. if x'_{max}, in the SMCM calibration, is equal to 2×10^{-11}, then the new maximum of the surface interaction values is equal to, $\hat{x}_{max} = 2 \times 10^{-8}$. This value was a subjectively chosen because it produced correct results. A more detailed and robust investigation into these calibrations is presented in the next chapter. The metropolis Monte Carlo algorithm was run for 5×10^5 iterations at a metropolis temperature set to $T = 0.001$, under which the program produced the following results.



(a) (b)

Figure 5.7: STM simulation of the minimum position of TiOPc on a Gold Surface (a), whilst the graph (b), shows the process of this minimization by plotting the value of the TiOPc interaction against the number of iterations. The steepness of the graph against a log scale indicating a particularly rapid descent. The parameters of the STM simulation are identical to those used in figure 5.3.

Figure 5.7a depicts an image of a TiOPc monolayer in a constant height STM simulation. The calculation and simulation used the HOMO level of the TiOPc as predicted by the EHT. The graph in figure 5.7b shows how the interaction potential decreasing as a function of the number of iterations, clearly demonstrating the expected exponential decay. The final position shows each molecule at a consistent angle, lying on either 29° or 30° indicating the true minimum of the

Table 5.1: Results of ten completed simulations of TiOPc on a gold surface, so the spanning maximum is the maximum gold value in the SMCM calibration $\times 1000$.

Simulation Number	**Mean Angle**
1	29.64
2	29.73
3	29.58
4	29.42
5	29.78
6	29.65
7	29.62
8	29.56
9	29.44
10	29.70
Simulation Result	29.6 ± 0.2

system lies somewhere between the two angles. The Mannsfeld and Fritz paper measured an experimental value of $28 \pm 2^{\circ}$, meaning the results of the Hückel simulation lie within experimental parameters.

5.6.1 Treatment of errors

Although the angle of each molecule was iterated though by single degrees, information about repeated simulations allows for a better prediction of the 'best' angular orientation and the error in the system. Stated in last paragraph, the angle of each molecule in figure 5.7a lies either on 29° or 30°, however, the mean of these angles, at that minimum position, remains remarkably consistent across multiple simulations. This indicates that this average between angular positions could be taken as a better evaluation of the 'true' minimum of the system. The calculation of this new minimum was completed by performing ten simulations, recording the mean angle of each simulation, then taking the average and range of those results to represent the minimum interaction orientation and its error respectively. The outcome of each of those ten simulations, along with the overall minimum angle, are displayed in table 5.1.

By collecting the results of those simulations into table 5.2, a comparison of the experimental and theoretical results can be completed. A final note on each individual simulation, being a random orientation process, the program occasionally became trapped in some local minimum position. For this reason, a criterion was designed to separate simulations that were successful from those that failed to find a global minimum. For a simulation to be considered complete, the standard deviation in the orientation between all molecules had to be less than two, $\sigma < 2$. Each simulation, used in predicting the value in table 5.2, met that condition. Whilst describing similar results, there are a number of key distinctions between this mode of calculation and the one presented by Mannsfeld and Fritz in [6]. By using an approximate classical measurement (the electric potentials) as the simulated property, their model required additional assumptions in the physics

Table 5.2: Comparison of results between the Hückel theory and Lennard-Jones Simulations performed in this research and the Mannsfeld and Fritz paper, along with experimental data from the same paper [6].

Hückel Simulation	**Lennard-Jones Simulation**	**Experiment**
$29.6^\circ \pm 0.2^\circ$	$27^\circ \pm 0.1^\circ$	$28^\circ \pm 2^\circ$

compared to the semi-empirical Hückel method. This model is, thus, offered as a potential improvement in terms of being a more complete theory that has similar levels of computational efficiency. However, some discrepancy lies within the completeness of the results displayed in both sets of research. For this work, the parameters about the overlayer structure were taken from the Mannsfeld and Fritz paper and used to predict the orientation, whereas, in their paper, they predicted the overlayer structure. Using the model to make these predictions is in its developmental stages, with the current implementation discussed in appendix I. For reasons detailed in the appendix section, the program, in its current form, was not capable of correctly identifying a minimum set of intermolecular structures. There is no reason this set of data could not be used for that calculation and this would be a natural expansion of the functionality of the program.

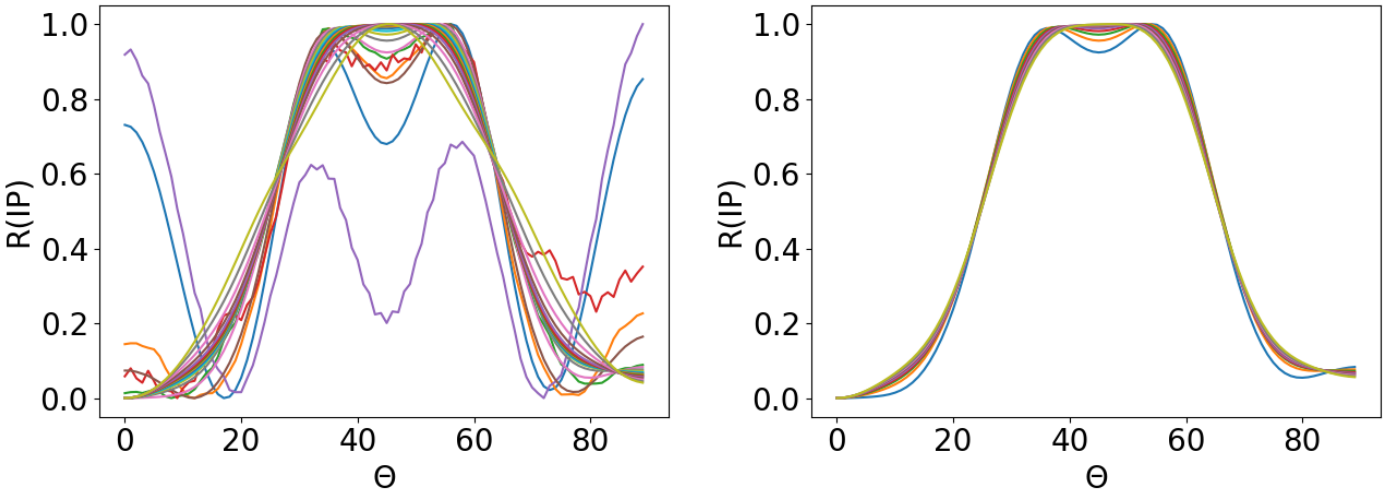
5.7 Height Dependence of the Surface Interactions

A factor so far neglected from the discussion is the effect of the monolayer height on simulations. Given the surface model and subsequent calibration, it is the relative strengths that are given precedence and so incremental changes of these interactions would not alter the simulations. However, that case must still be demonstrated otherwise the simulations would lack a degree of validity.

To analyse the effect of the height in this model, surface grids were calculated for six surface points, at a monolayer height that ranges from 1Å to 20Å at 1Å intervals. In this context these height values are measured from the central metal atom (which lies in the plane $z = 0$), with the phthalocyanine curvature orientated towards the surface, meaning the metal atom is at the highest point on the layer (excluding the oxygen atom in the TiOPc example), as displayed in figure 5.1b. The molecules can reside in either direction and form can intricate multilayer structures as a result [111]. In this work the concave side was favoured as facing towards the surface as it would increase the surface effects making it easier to draw a comparison.

By normalising each set of grids it is easier to make a direct comparison of the cost of each data set. Examining all normalised sets from 1Å to 20Å at one surface point produces the graph in figure 5.8a:

In figure 5.8a it is clear that there are some significant discrepancies between the cost of the surface grids values at a number of height positions. However, it can be shown that these deviations arise from heights values at each extreme i.e.

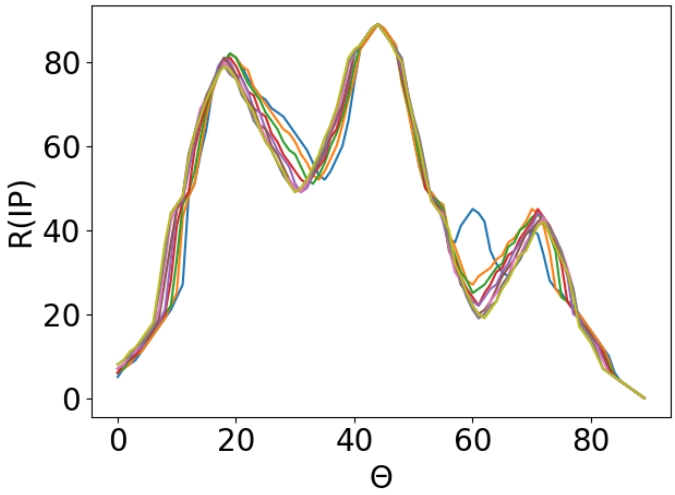


(a) Position five, normalised values, $1 \leq h \leq 20$

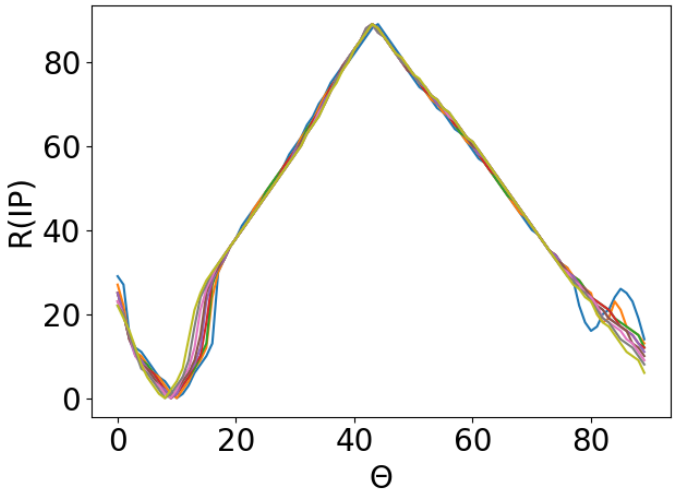
(b) Position five, normalised values, $5 \leq h \leq 15$

Figure 5.8: Depiction of two graphs making a height comparison, a) shows a graph displaying the change in the height of the surface from 1Å to 20Å, in 1Å intervals, at a single point on the surface. Graph b) shows the height data set at the same point as height changes from 5Å to 15Å. Each plot is set to follow a colour scheme that repeats after nine plots, starting from the 1Å line, the colours are as follows: blue, orange, green, red, purple, brown, pink, grey, yellow, light blue. This is the same colour scheme that was used in each figure in this section, starting from the smallest value.

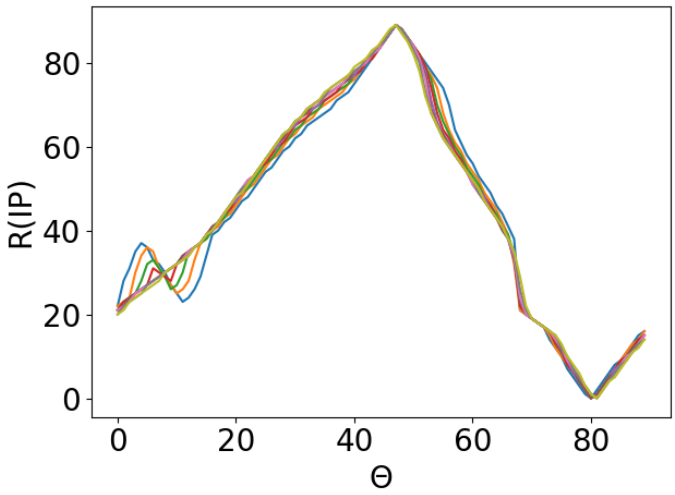
at low values or high values. Plotting normalised interactions between 5Å and 15Å creates the graph in figure 5.8b. This graph shows each curve closely resembling one another, displaying a circumstance in which the difference in cost of the height positions are negligible between a given range. Testing this hypothesis further, each the value of the grid sets, of each position were ranked, so each grid value ranged between 0 and 89. The images in figure 5.9 display the results of performing that calculation. In most cases, the cost at each point on each curve is identical or closely related to each other. Where there exists small variations is at points when the trend of the curve is changing, for monolayers closer to the surface, the stationary points edge towards occurring sooner. However these differences are minimal, with there only being three regions across the six graphs where the difference is more than 2° (in figures (5.9a, 5.9d and 5.9e). The conclusion follows that handling of the surface height between these ranges will only impact the surface calibration and, as such, will not impact on the simulation results.



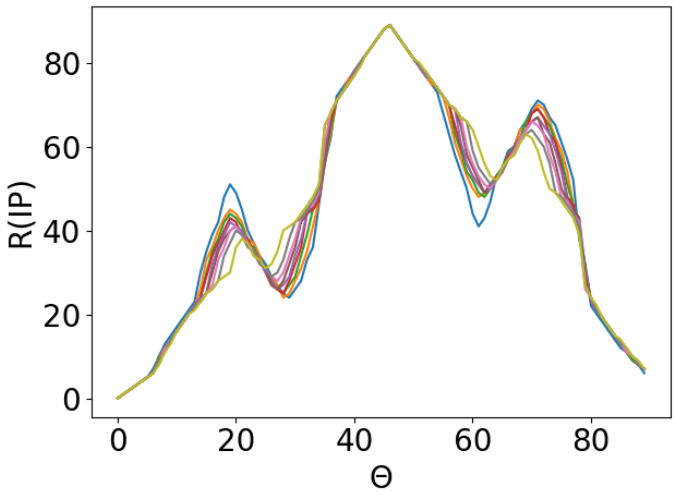
(a) Position one, ranked values



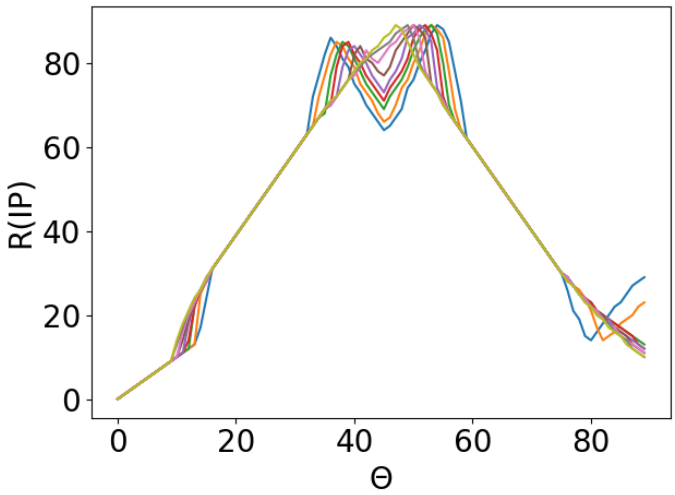
(b) Position two, ranked values



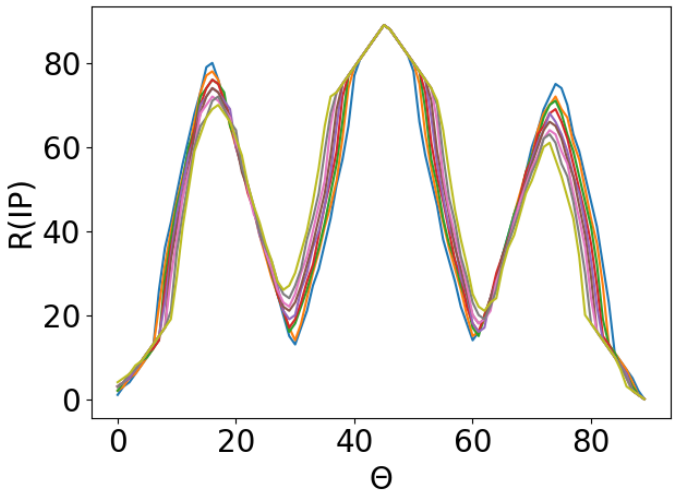
(c) Position three, ranked values



(d) Position four, ranked values



(e) Position five, ranked values



(f) Position six, ranked values

Figure 5.9: Display of six figures comparing the cost function of the molecular rotation at different surface points as the height is raised from 5Å to 15Å.

Chapter 6

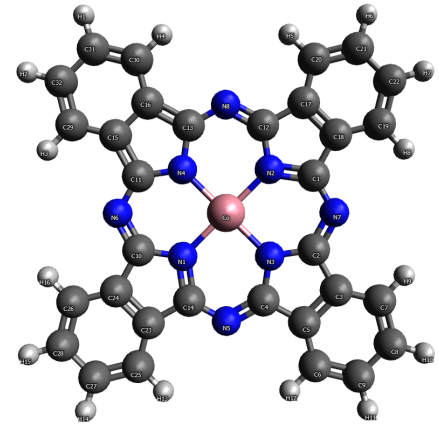
Comparison of Various Metal Phthalocyanines

6.1 Comparison of Different Central Atoms

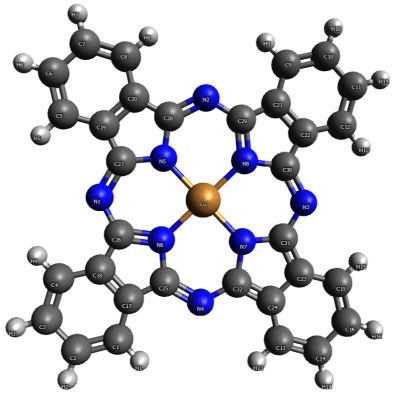
The last chapter demonstrated the effectiveness of a Hückel based approach for monolayer simulations. The rest of this body of work will be focused on following the same EHT and Monte Carlo procedures but applying it to a number of metal phthalocyanine molecules, across a wide variety of initial conditions. The particular phthalocyanine molecules being added to the investigation here are cobalt, copper, tin, and iron phthalocyanine (CoPc, CuPc, SnPc and FePc), alongside TiOPc from the last chapter.

Each variant of phthalocyanine was created by changing the single metallic atom bound into the centre, however, this alteration can have a significant impact on the molecular and electronic structures present in calculations. An assessment of these structural changes is presented in figures 6.1 and 6.2, in which a ball and stick model is presented of both a top down view and a side elevation of each of the newly added MPc molecules. All of the molecular structures were derived from, and thus presented, using the Avagadro chemistry program [101] by applying the same inbuilt functions utilized to obtain the TiOPc atomic structure in the previous chapter.

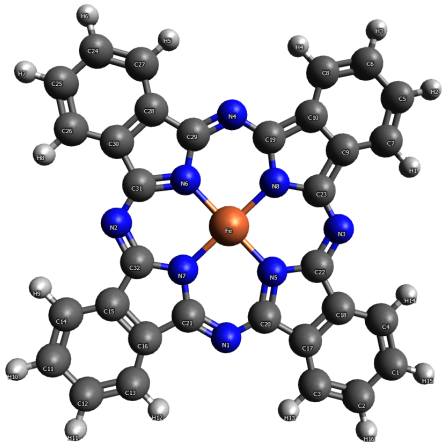
Contrasting the four molecules, it can be determined that of the four, CoPc and FePc are the most similar in their features and have approximately the same electronic distributions. CoPc, CuPc and FePc are all found to be highly planar with CuPc being the most planar. This is evidenced by the maximum difference in the z value between features, which is 0.10Å, 0.11Å and 0.0002Å for CoPc, FePc and CuPc respectively. CuPc is also the smallest of the group, having a maximum span that is less than that of CoPc, FePc and SnPc by 0.16Å, 0.16Å and 0.22Å respectively. The most distinct structure of the four molecules is attributed to SnPc, the dissimilar structure occurs because the tin atom is significantly larger than the transition metals that comprise the rest of the central atoms used in this study. The size of the orbitals centred at the tin atom cause them to sit out the plane, which creates the curvature seen in figure 6.2.



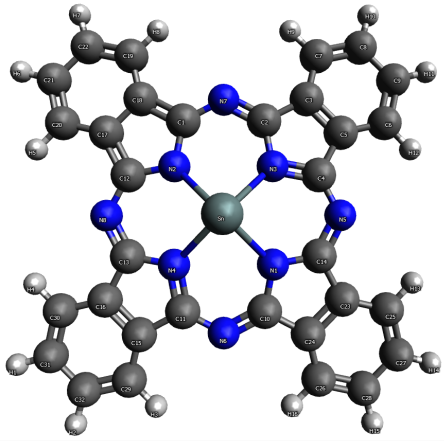
(a) CoPc



(b) CuPc

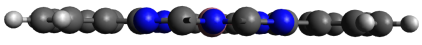


(c) FePc

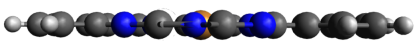


(d) SnPc

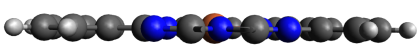
Figure 6.1: Images depecting a ball and stick model for CoPc, CuPc, FePc and SnPc. CuPc and FePc structures were calculated using the avogadro program like TiOPc. CoPc and SnPc were structures calculated though the molecular studio.



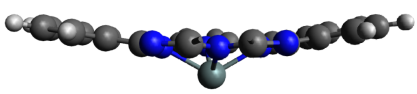
(a) CoPc



(b) CuPc



(c) FePc



(d) SnPc

Figure 6.2: Images showing the same ball and stick model but from a side elevation as to show the planarity of the molecules.

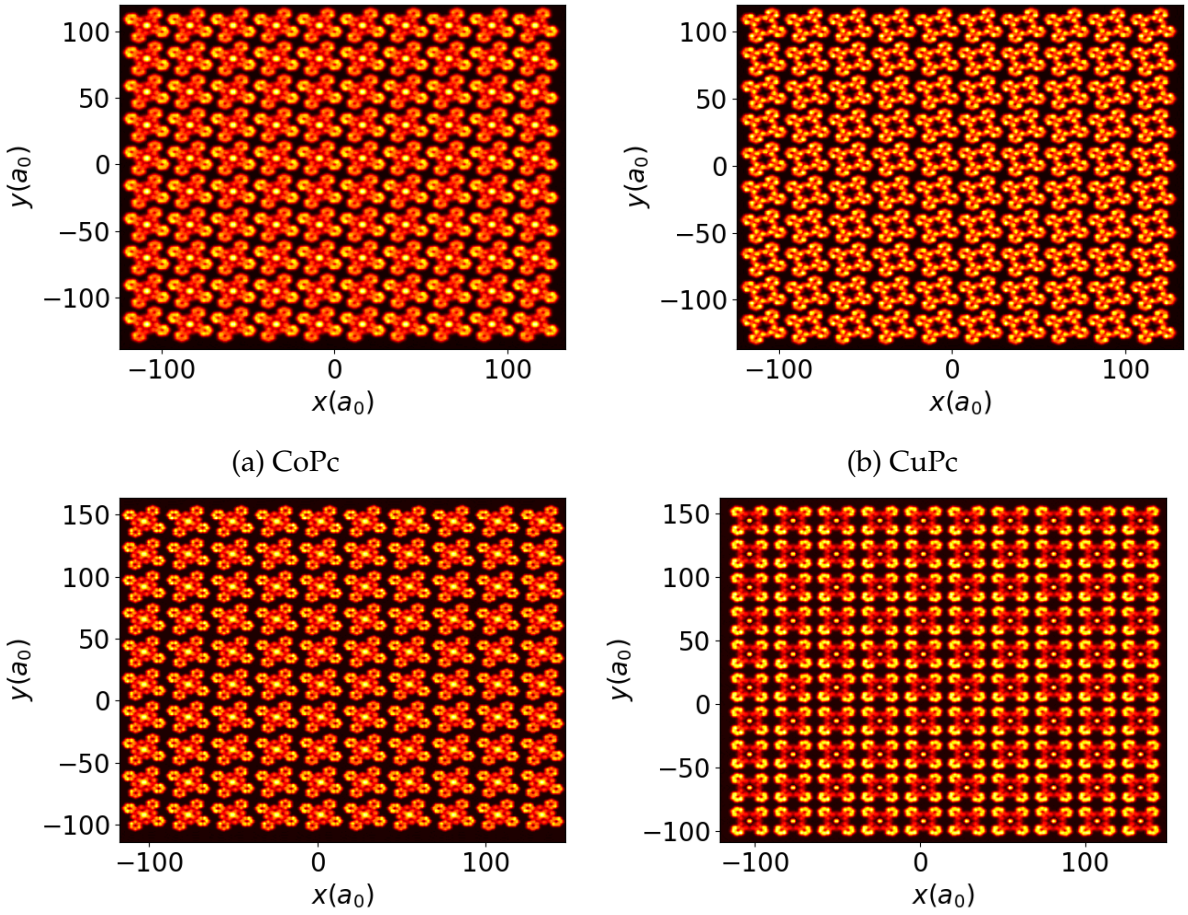
A number of these differences are slight but they are observable and will ultimately impact the minimum energy configuration of each monolayer. To observe how these structural changes affect the monolayer alignment, a simulation was run for each MPc using an identical set of starting parameters to those used in TiOPc simulations conducted in the last chapter (section 5.6). Recalling that these particular conditions are, the molecular spacing $b = 13.9$Å , the surface orientation angle $\phi = 30^\circ$ and the angle between rows $\Gamma = 90^\circ$, as shown in figure 5.6. The set of simulations produced from these initial conditions are displayed with the minimum angle positions in figure 6.3. Each of these images is an STM simulation over the molecular monolayer, these were generated using the HOMO (this was the same level used though out the calculation process) for each molecule in a constant height simulation at $z = 3$Å, these STM simulation parameters are used for each of the simulated images created for this chapter.

The molecular angles in these simulations vary by single degrees and so tend to place themselves around the angle that represents the true minimum of the model, as was observed when simulating TiOPc. Again, for each molecule, ten simulations were run and the average angle was recorded for each, the average of those was taken a better representation of the systems minimum orientation and the subsequent range was taken to be the error.

Table 6.1: Results of ten completed simulations of FePc, CoPc, CuPc, SnPc on a gold surface, so the spanning maximum is the maximum gold value in the SMCM calibration $\times 1000$.

Simulation Number	**FePc**	**CoPc**	**CuPc**	**SnPc**
1	30.96	31.25	27.95	45.18
2	31.23	31.36	27.91	44.66
3	30.73	31.18	27.94	44.92
4	31.11	31.41	27.81	44.35
5	30.77	31.39	27.65	44.36
6	31.2	31.89	27.91	44.33
7	31.25	31.36	28.01	45.03
8	31.01	30.94	27.86	44.47
9	31.22	31.14	27.98	44.39
10	31.08	31.13	27.85	44.69
Mean Angle	31.1 ± 0.3	31.2 ± 0.3	27.9 ± 0.2	44.7 ± 0.4

Each image in figure 6.3 represents a highly ordered array of molecules with the standard deviation between the alignment angles of each simulation being less than one. FePc and CoPc share similar almost identical phases with the minimization angle being 31.1° and 31.2° respectively, as shown in the above simulations. The CuPc aligns at the distinctively lower angle of 27.9 showing a clear difference in how the electron density is distributed in CuPc. The SnPc behaviour was entirely different from the other three molecules, with the average azimuthal angle between molecules been equal to 44.6°. This difference is most likely a result of the significant structural differences discussed above.



(a) CoPc

(b) CuPc

(c) FePc

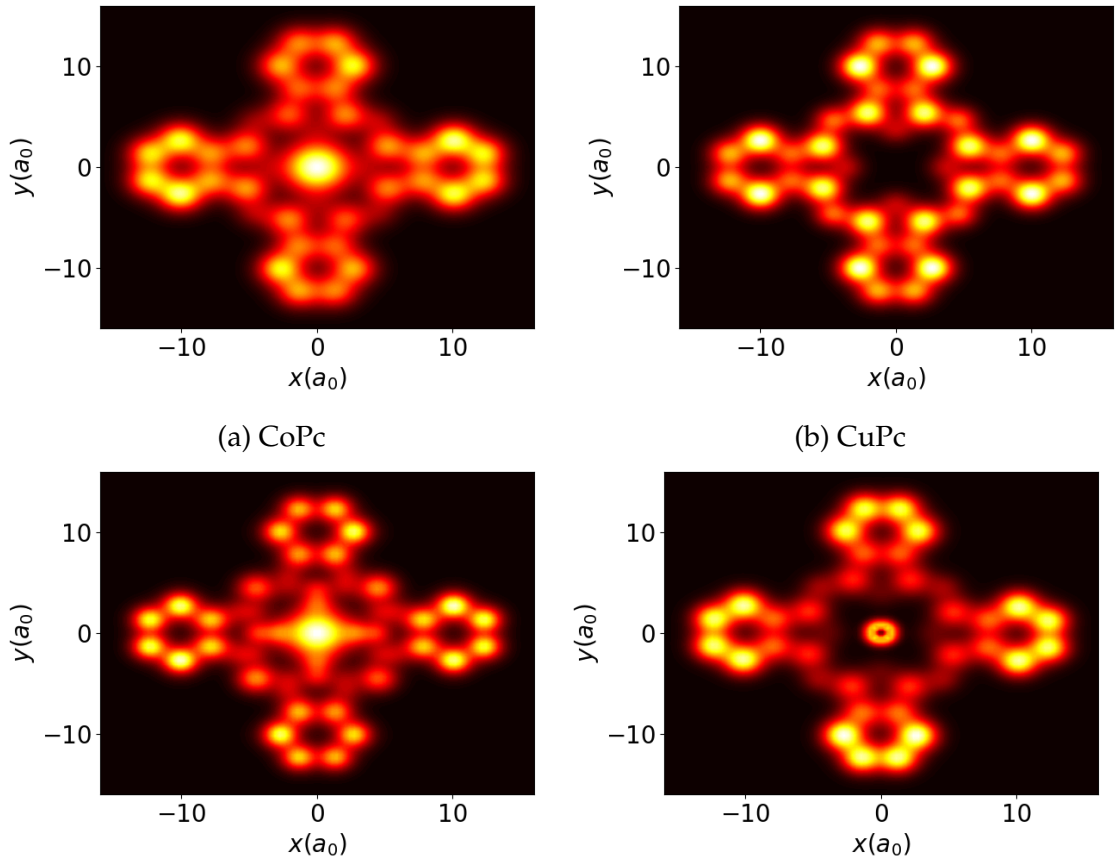
(d) SnPc

Figure 6.3: Figure comparing the alignment of the four new MPc's after inputting the same starting parameters used for the TiOPc simulations presented in the last chapter.

6.1.1 Comparison of the STM simulations of Metal Phthalocyanine to experimental images

In figure 6.3, four STM images are presented of different MPc monolayers. The STM simulation program creates a unique pattern for each MPc molecule, each being dependent on the particular MO used for generating the simulations. Comparing these images, with those described in STM experiments, provides a check on the validity of those particular MO calculations. To make a detailed comparison, an individual molecule occupying the HOMO is simulated for each MPc in figure 6.4.

Both the images of CoPc and FePc are, in these simulations, shown to have a bright spot in the centre which is in accordance with those found in experimental images [7, 13, 14, 15]. The images of FePc and CoPc are also distinctly similar which is again expected from those molecules as iron and cobalt are similar elements. In general, the particular bright spots away from the centre are not specifically obtained in STM images, even for high resolution images, these generally appear as regions of consistent brightness that form a luminous cross in experimental images [13, 112]. From this basis, the CuPc simulation are also highly



(a) CoPc

(b) CuPc

(c) FePc

(d) SnPc

Figure 6.4: Figure comparing the alignment of the four new MPc's after inputting the same starting parameters used for the TiOPc simulations presented in the last chapter.

representative of real STM images, as here they are simulated as having a dark centre and the four symmetrically bright arms that are emblematic of phthalocyanine molecules. Some discrepancy is clearly present for the simulation of SnPc, particular with the high resolution results presented in [113]. They present a detailed picture of a number of orbital structures which are not visible in the simulation displayed in figure 6.4d. In the STM simulations there are bright spots at each end of each isoindole and a noticeable but comparatively dull centre. In real experiments the sections away from the centre would be expected to be more consistently bright, not just bright at each end and the central tin atom would be significantly brighter since it is positioned 2Å above each end, being positioned on the side with the molecule up. A possible explanation could be that the non-planarity of SnPc means that the constant height simulations are not able to detail all of the components of the molecule effectively. It is also a possibility that this mode of calculation does not fully account for the complex sub-molecular interactions that the large tin atom creates. Despite these inaccuracies, in terms of the visual representation, the results of the monolayer simulations are generally consistent with the results of similar simulations that use other MPc molecules. This leads to the conclusion that the monolayer simulation results for SnPc are equally valid as those produced from the other types of MPc.

6.2 Comparison of Intermolecular Spacings

Earlier in the work, it was mentioned that a significant advantage in choosing a Hückel based approach is that its simplicity allows for has a high level of flexibility. In practice, this translates to the idea that certain system parameters can be easily adjusted on an individual basis in order to isolate the variables and study their effects. The following sections of this work will demonstrate this fact by performing multiple simulations across a multitude of variable parameters.

The natural evolution for this research was to next investigate the effects of changing the spacing parameters. With this consideration, the next set of data was obtained by adjusting the intermolecular spacing variable b. To remain consistent, the simulations were performed on a gold surface in the SMCM calibration (equation (5.5)) and the angle between each row of molecules Γ was 90°. The effects were then compared from altering the spacing between molecular centres from 10Å to 20Å in 1Å intervals. At each step the Monte Carlo algorithm was applied and the angular positions with the subsequent minimums measured, producing the results shown in figure 6.5.

Figure 6.5a shows the change in the minimum azimuthal angle for each molecule as a function of the intermolecular spacing. Figure 6.5b maps the standard deviation of the alignment angles to the spacing. The standard deviation is a useful value because it is a measure of how well the program performed in finding a minimum i.e. if the standard deviation is low the molecules are in a highly ordered state and if they are disordered the standard deviation is high. Figures 6.5c and 6.5d are presented to aid visualising the data, each shows an STM simulation of the the ordering of an FePc monolayer, at 12Å and 18Å respectively.

In both sets of data SnPc is an exceptional case. At smaller ranges of the atomic spacing, there was particular difficulty for the program to find a minimum. An indication of how unlikely the routine was to settle on a minimum in those cases is provided by examining the implementation of the calculation. The program was designed so the minimization routine would repeat up to ten cycles and would not return the data unless the standard deviation across the angles was less than two, $\sigma < 2$; if after 10 cycles had been performed, no routine meeting that criteria had been achieved, the values at which the standard deviation was the smallest were taken. To add further context, when the system successfully settled on a minimum, that met the $\sigma < 2$ condition, the success rate was approximately $\approx 85\%$. This leads to the conclusion that in cases where the simulation failed to meet that condition, in any cycle, it is likely that factors inherent in the system prevented it from doing so.

Examining the case of SnPc where the molecules did not fully align, this effect was mostly observed at small distances. Given that spacings of that size between molecules are unlikely to be observed in reality, this suggests that the inability to find a minimum is an artefact of the artificially high interaction at those spacings. It is also interesting that SnPc behaviour becomes similar to the other MPcs at values of b less than 14Åor greater than 16Å, indicating that the example seen in

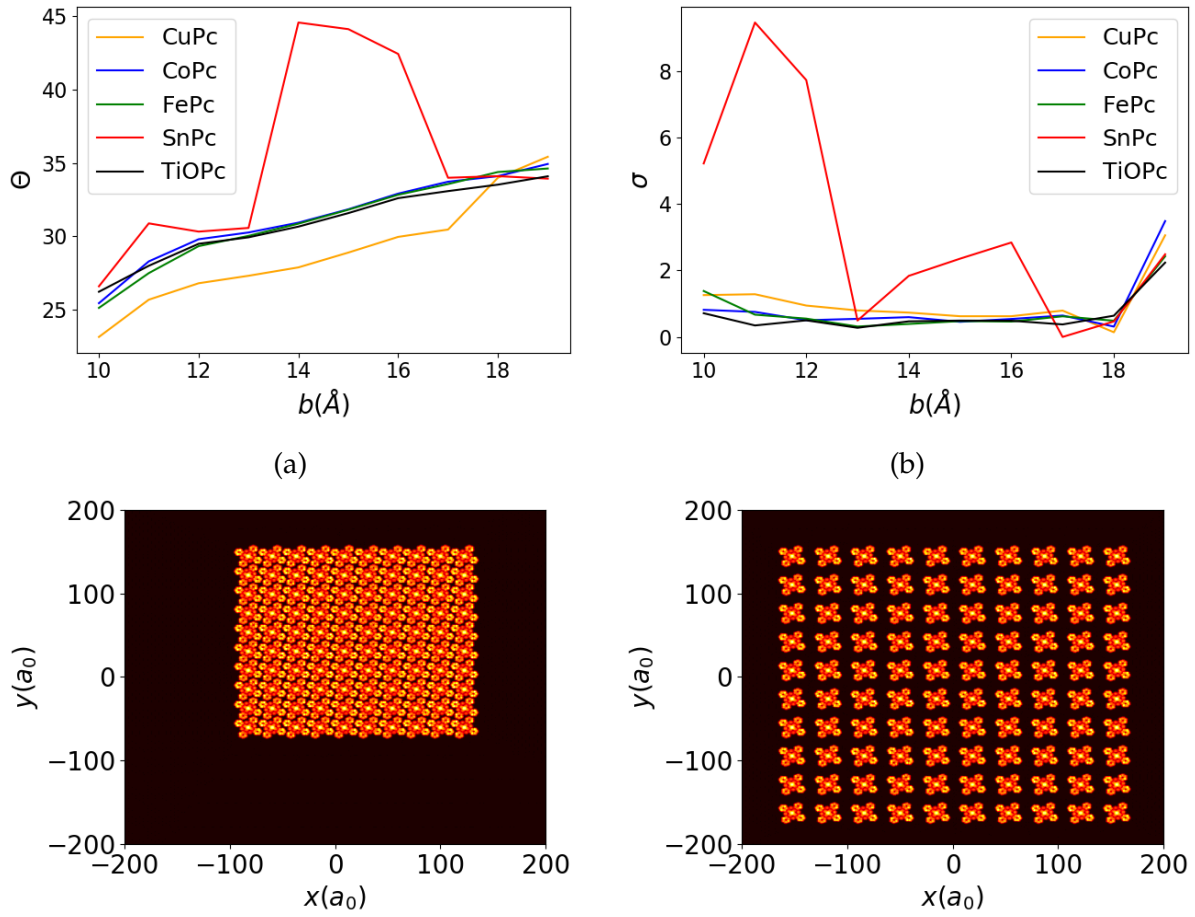


Figure 6.5: Figures comparing data for how changing the intermolecular spacing effects the alignment minimisation for all five MPc molecules. Figure (a) shows how the average angle across the monolayer, Θ, changes as the spacing increases. Figure (b) displays the standard deviation, σ, in the azimuthal angle for the monolayer. The second row presents two STM simulation images of FePc, (c) at $b = 12$Å and (d) at $b = 18$Å, these were found to have alignment angles of 29.3° and 34.4° respectively.

6.3d is, in fact, a special case observed in particular circumstances. The TiOPc, FePc and CoPc molecules all displayed behaviour that was closely aligned with one and another. This was mostly expected from FePc and CoPc as cobalt and iron are elements that exhibit a lot of similar behaviours and so their respective molecules are structurally similar, as described in section 6.1. For TiOPc, this was unexpected as there are some significant structural differences particularly with regards to the planarity of the molecules. TiOPc has a notable concave side that affects the electron distribution. However, in this circumstance the alignment angle is unaffected.

The minimization of CuPc exhibits the same general behaviour as the FePc, CoPc and TiOPc molecules but the orientation is at an angle distinctly lower, by approximately 2°. The steepness increases for CuPc at the last two data points so the alignment angle at 20Å is functionally similar. An interesting effect that

comes out of the simulation is that the standard deviation, for all MPc molecules, sharply increases at the distance of 19Å. This effect is likely a result of the decreasing interaction between neighbouring molecules minimising their incentive to form a rigid structure. It can therefore be theorized this increases the significance of the surface effects to a point that causes molecules to deviate from the optimum molecular alignment positions. Although several intermolecular lattice structures in the above demonstration would not be physically observed, it is clear there is a linear pattern that could apply to realistic structures.

Before examining the results of further simulations an important note must be made on the error of the calculations in this chapter. In the simulations above and the subsequent variations that form the rest of the investigation, only a single routine was performed to find each minimum angular position (so long as the minimum criterion was met). Therefore, the error of each position is a whole degree ($\pm 0.5^{\circ}$) as that is the minimum amount each molecule can be rotated.

6.3 Comparison of Various Super-Lattice Structures

In the previous set of simulations, the data was gathered by altering the distances between molecular centres alone, however, experimental measurements often quote these values in terms of a super-lattice structure ([6, 7, 114, 115] are examples of just a few). A super-lattice being a description of the molecular positions as described by a 2×2 matrix that represents the shifts in terms of the surface lattice spacings:

$$\begin{bmatrix} b1 \\ b2 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{6.1}$$

where i and j are the basis vectors and equal to the atomic spacing of the surface atoms. The terms inside the matrix represent how much the molecule is shifted by in each direction. An example is shown in figure 6.6.

Alterations of these macro-intermolecular lattices require the angles between neighbouring molecules to be more carefully considered, however, introducing this form increases the potential for investigating realistic overlayer structures. This, in turn, allows the study of increasingly realistic systems, as will be demonstrated in the next section that looks to compare different surfaces. This section will, however, focus on the impact of changing the super-lattice in isolation.

Changes to the super-structure variables have considerably more potential in terms of the size of the position space available. To reduce this to a more manageable set of possible variations each variable of the matrix in equation (6.1) was altered individually. The implementation of this process was actualized by starting with a square structure, similar to the images presented in the previous simulations, and adding or subtracting an amount to each term in the super-lattice matrix, equation (6.1). In this formulation, The lattice parameters for a square

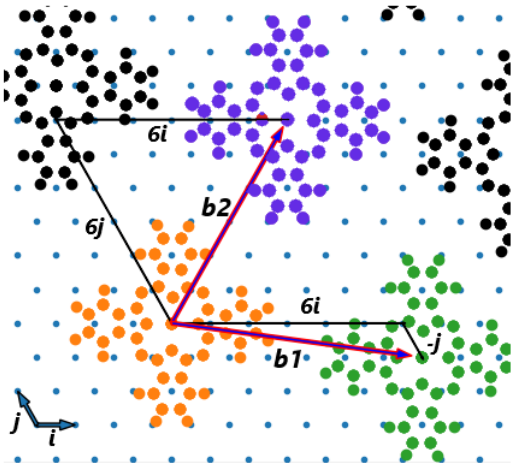


Figure 6.6: Diagram shows an arrangement of pthalcyannine molecules with a super lattice structure corresponding to the matrix values $a = 6$, $b = -1$, $c = 6$, $d = 6$.

lattice with a number of spacings n are given by:

$$\begin{bmatrix} b1 \\ b2 \end{bmatrix} = \begin{bmatrix} n & 0 \\ \frac{n}{\sqrt{3}} & 2\frac{n}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{6.2}$$

Starting from this basis, a change can, for example, be introduced into the square by, subtracting two from the c value:

$$\begin{bmatrix} b1 \\ b2 \end{bmatrix} = \begin{bmatrix} n & 0 \\ \frac{n}{\sqrt{3}} - 2 & 2\frac{n}{\sqrt{3}} \end{bmatrix} \begin{bmatrix} i \\ j \end{bmatrix} \tag{6.3}$$

The result of this change can be seen in figure 6.7, where each row in the new set has shifted in the negative x-direction from the original square. What is now clearly observable is the angular considerations mentioned previously in this section. A gold(111) surface has an atomic spacing equal to 2.88Å, that corresponds to a change in the angle of 21.80° from the 90° when $n = 5$, as displayed figure 6.7.

For the following simulations the base $n = 5$ was used as it corresponds to an intermolecular spacing of 14.4Å on gold, which maintains a close value to the other previous simulations. The graphs in figure 6.8 were found by running the optimization program for each MPc as the c variable ranges from $\frac{n}{\sqrt{3}} - 4$ to $\frac{n}{\sqrt{3}} + 4$ in whole integer steps.

The first important detail in figure 6.8, are the results of figure 6.8b, that show the standard deviation of each simulation is less than or approximately equal to one. This demonstrates that each simulation was stable and produced a highly ordered structure, giving a strong indication that these minimization routines, resulting in a structure that represents a close approximation to the true minimum

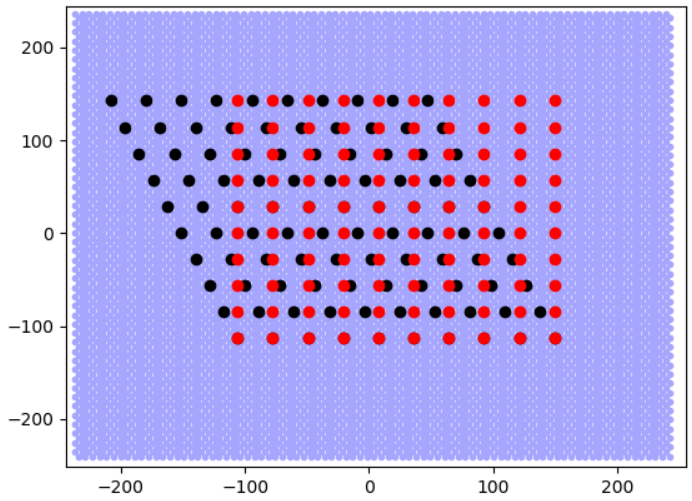


Figure 6.7: Point display comparing the difference between a square super-lattice structure formed using equation (6.2) (red), and one formed using the altered version presented in equation 6.3 (black).

in this system. A unique feature of the graph displayed in figure 6.8a is that the angular measurement records the molecular rotation in the clockwise direction from $c = -4$ to $c = 0$, but from $c > 0$ this measurement is made by observing the rotation in the anti-clockwise direction. The reason for this is two-fold, to improve readability and better represent the symmetry. In the $c > 0$ region, for a value x on the graph, the actually recorded value from the simulation was $90^\circ - x$. However, on the graph this would look like two plateaus with a steep gradient in between, whereas this clearly displays the symmetry for the off center positions that can be seen in figures 6.8c and 6.8d.

Comparing each MPc it can be seen in figure 6.8a that the five molecules exhibit similar behaviour. The angle remained stable for all c values that deviate away from the square, and then peak at the point of becoming square before reducing back to approximately the same angle except rotated from the opposite direction. It can be seen that the molecules TiOPc, CoPc, FePc, and SnPc, in these simulations, all favoured similar positions being within 0.5° of one another. CuPc, whilst still following the same pattern, was at an angle of around 1° less than the other three. This can be simply put down to differences in electron configuration. The similar behaviour of SnPc in this environment is unexpected given the results displayed for SnPc in Section 6.2. A possible explanation for this could be that the effect of the surface has been altered because the change in the method of construction has translated the monolayer compared to previous examples. Whilst unexpected, the similar behaviour of SnPc in this circumstance would further indicates the original example, displayed in figure 6.3d, is the exceptional case for SnPc monolayer alignment. It is interesting that the angle remains stable in positions where $c \neq 0$. A possible reason for this could be that as c increases, the effect of changing the angle of the super-lattice could be offset by the increasing distance, due to the relationship put forward in the previous section.

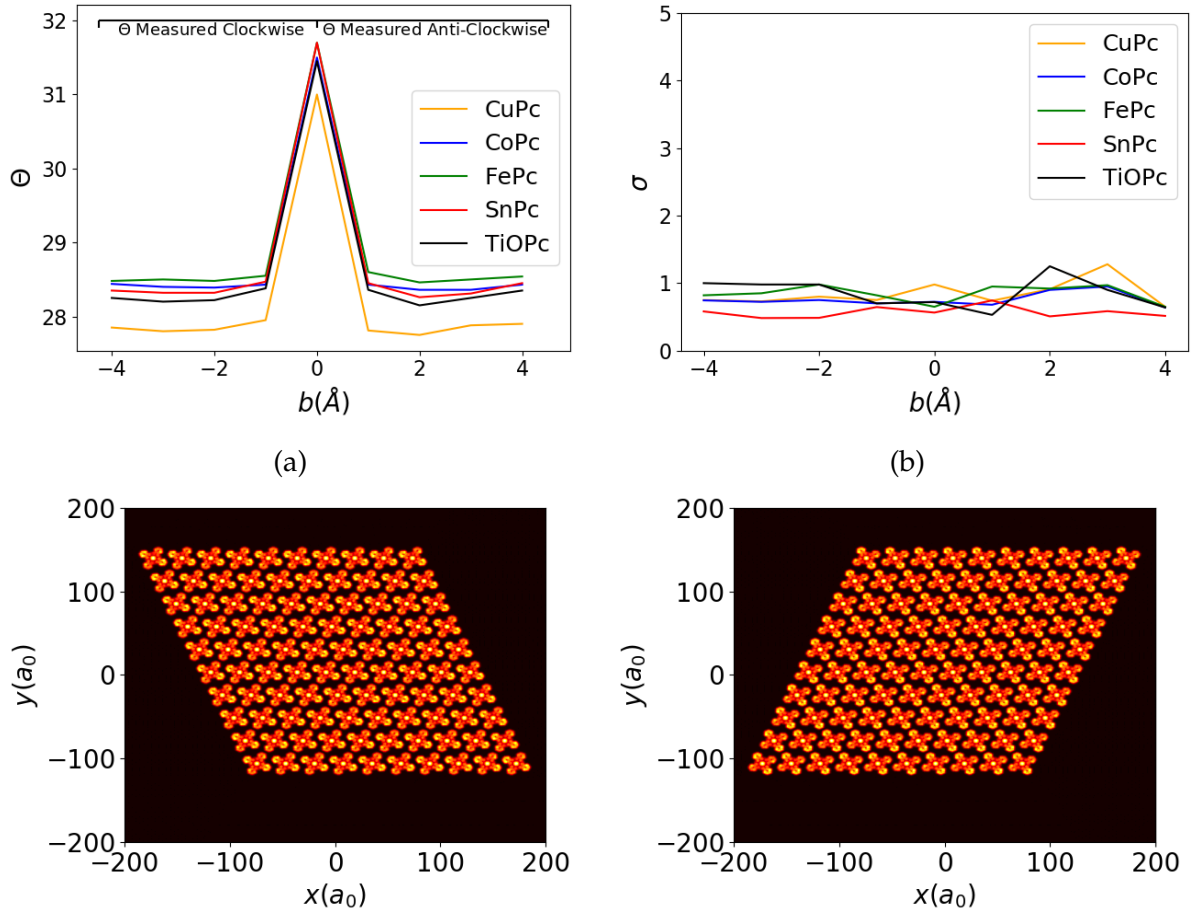


Figure 6.8: Figure displaying plots to show the effect of varying the c variable in the overlayer structure. Graph (a) displays the minimization for each molecule, (b) shows the standard deviation of each points and, (c) and (d) are STM simulations for the results of CoPc at $c = +2$ and $c = -2$ respectively, relative to the square position.

The continuation of this section will examine the results that are obtained by changing the b and d components. To maintain consistency this investigation was performed using exactly the same routine, each grid was set up in a square and a number was added to alter this position. Like the previous example the n value in the square super-lattice equation (equation (6.2)) was $n = 5$ and the constant addition ranged from -4 to $+4$. A simulation was performed for each value to find the minimisation, the results are displayed in figure 6.9.

Before discussing those results a note needs to be added for why adjusting the a value in equation (6.1) has not been similarly recorded. In this context changing the a value is akin to squeezing or stretching the square into a rectangle which in essence only changes the intermolecular spacing and so the effect of that alteration is covered at an earlier point in the chapter.

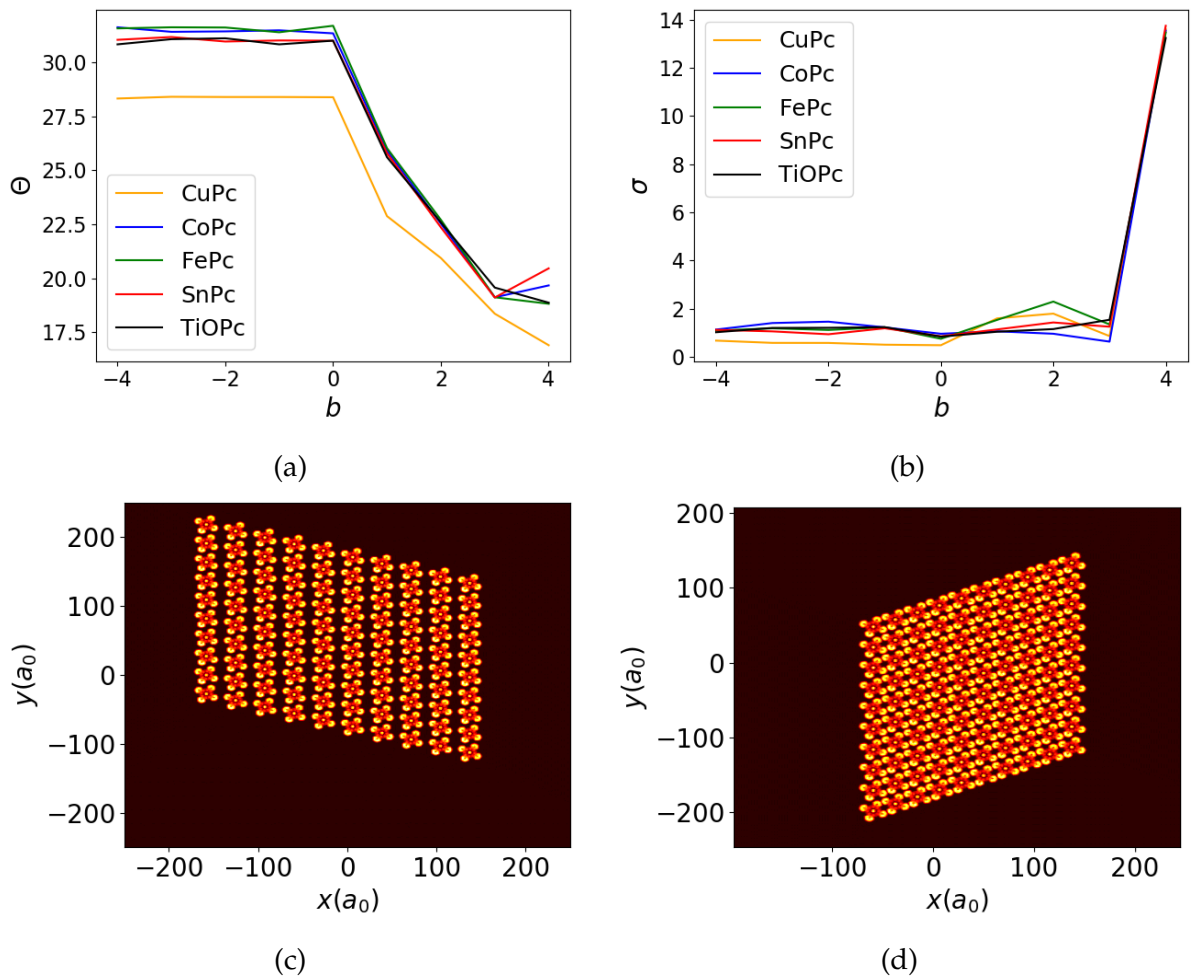
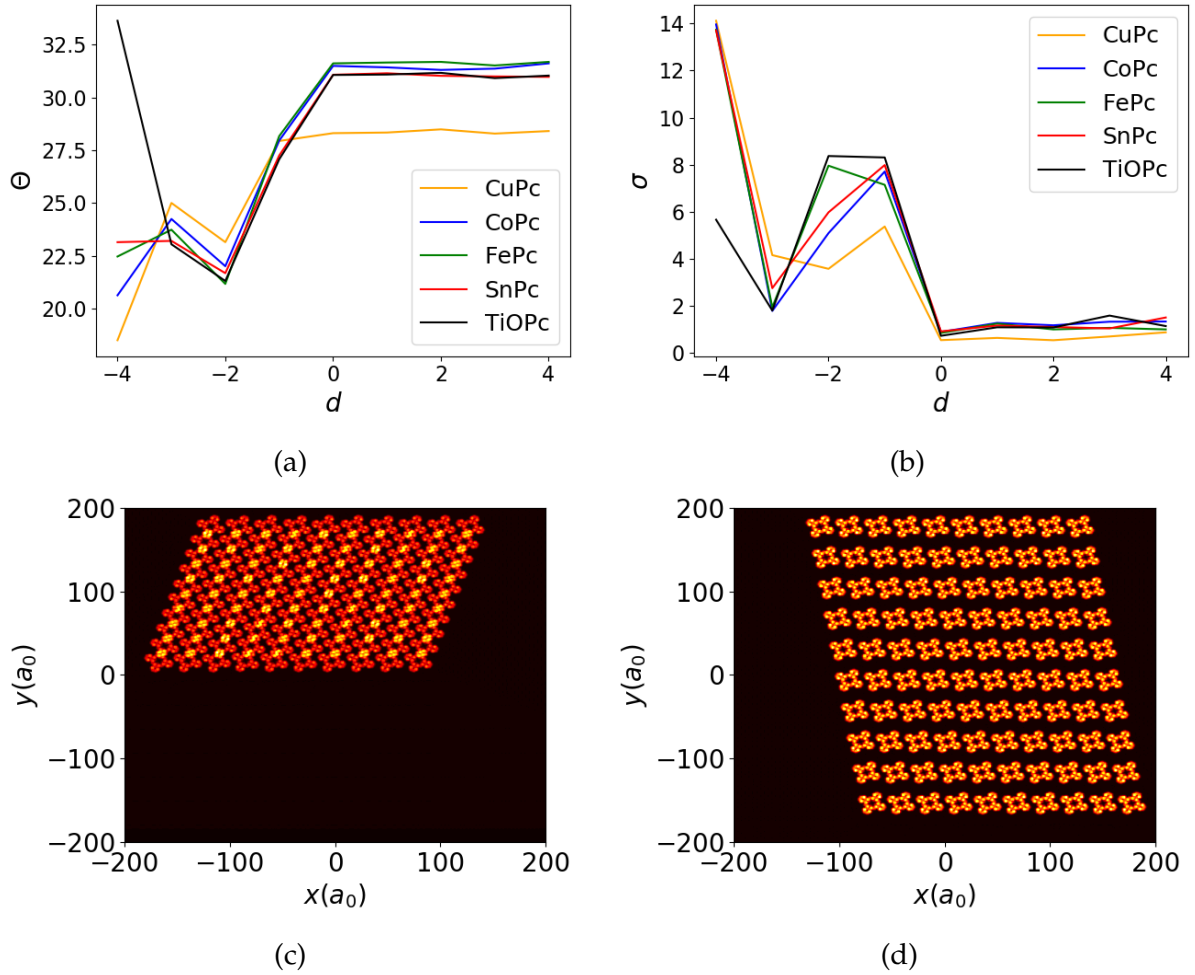


Figure 6.9: Figure displaying plots to show the effect of varying the b variable in the overlayer structure. Graph (a) displays the minimization for each molecule, (b) shows the standard deviation of each points and, (c) and (d) are STM simulations for the results of SnPc at $b = -2$ and $b = +2$ respectively, relative to the square position.

Again, as was the case with adjusting the c component, each molecule follows essentially the same pattern with the CuPc molecule coming to rest at an angle $1 - 2^{\circ}$ lower than the other four molecules. Another point of similarity between the two sets is they maintain a region of stability whist the intermolecular spacing is high. The explanation, however, does differ in this region the decreasing b-value causes an increased intermolecular spacing between horizontally placed adjacent molecules but maintains the spacing between those vertically aligned, as is displayed in figure 6.9c. Therefore, the interaction between those vertically spaced molecules dominates causing the overall minimisation to remain unchanged. Positive increments of b cause the reverse effect in terms of intermolecular spacing, as b increases in the $0 < b < 4$ region, the intermolecular spacing between horizontally placed neighbours becomes increasingly small, as can be seen in figure 6.9d. The consequence of this change causes the alignment angle to decrease until the spacing is reduced to a point that no realistic alignment is possible, at which point the angular positions becomes unstable resulting in the sharp increase in the standard deviation in figure 6.9b.



(a) (b)

(c) (d)

Figure 6.10: Figure displaying plots to show the effect of varying the d variable in the overlayer structure. Graph (a) displays the minimization for each molecule, (b) shows the standard deviation of each points and, (c) and (d) are STM simulations for the results of CuPc at $d = -2$ and $d = +2$ respectively, relative to the square position.

The next and final set graphs for this section show the effect of altering the d-component in the super-lattice, equation (6.1). The pattern shows on the graph in figure 6.10a is approximately a reflection of the plot shown in figure 6.9a, the angular positions being stable in the region $d \geq 0$, and showing a steep linear increase where $d < 0$. Two key distinctions are present in this comparisons to the b variable positions, the instability in the angular positions when d is negative, and the linear decrease in the region from $d = -4$ to $d = -2$. An explanation for both of these features can be found from observing the STM simulation in figure 6.10d, the image displays the monolayer at $d = -2$, in which the routine has found a set of positions in which atomic orbitals in the phthalocyanines overlap. The high interaction is found to be the minimum position because there are no possible orientations, for this monolayer arrangement, that do mix atomic positions. Therefore, at negative values of d, the artificially high interactions affect the ability of the algorithm to effectively minimise.

An interesting conclusion is that changes to individual lattice vectors alone seem to have little overall effect on the angle at which the molecules orientated. An explanation for this is that individual changes to single variables in equation (6.1) maintain the intermolecular spacing between one set of adjacent molecules in the monolayer. Given the exponential decay of these interactions those molecules that are closest to each other dominate the minimization process, reducing the impact of other surrounding molecules. Thus, a conclusion from this process is that the individual spacings of nearest neighbours have by far the greatest impact on the monolayer minimisation.

6.4 Comparison of Surfaces

In the previous chapter, a number of ideas on how to combine the surface and intermolecular interaction values were discussed because this method of simulation requires some form calibration. The choice of this can have a significant impact on the minimization of the system. In the simulations used throughout this chapter, so far, the calibration has been set to a scale using the SMCM calibration. Recalling that this method simply transforms the surface values, x', by a multiplication factor, f, (where $f = \frac{x_{min}}{x'_{min}}$) in order to equalise the minimum of the two sets. Previously expressed was the critique that this has the disadvantage of significantly reducing the impact of the surface on simulations, thus, making comparisons between surfaces difficult. Consequently, before being able to measure the impact of different surfaces a more robust and reasoned calibration of the interactions needs to be determined.

To investigate possible calibrations it is logical to start with examining a case where the interaction has a known outcome, here that outcome is the results of the 2005 paper by Mannsfeld *et al* [6]. The results of their work are detailed in the last chapter (table 5.2); the intention at this current step is to investigate the effect of changing the calibration and use the data from that experiment to make an informed decision on what the calibration should be. To demonstrate how this single set of data could be used to calibrate all TiOPc surface interactions, each surface interaction data set is presented in table 6.2:

Table 6.2: BHM secular equation solutions

Surfaces	**Minimum Value**	**Maximum value**
Gold (Au)	3.9784×10^{-4}	6.3970×10^{-4}
Silver (Ag)	5.6401×10^{-10}	1.22954×10^{-9}
Sodium Chloride (NaCl)	1.1437×10^{-5}	2.2642×10^{-5}

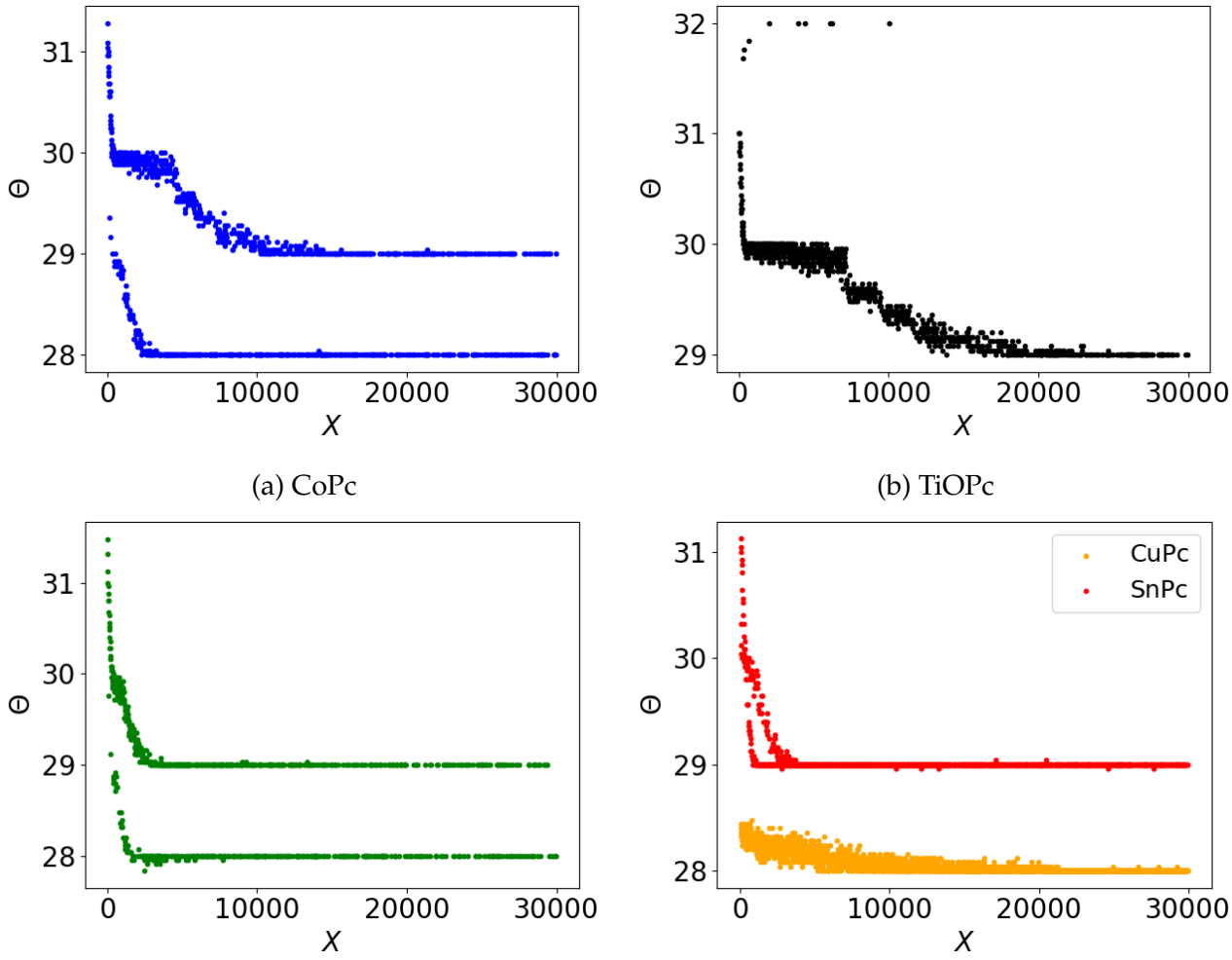
Given the interaction of a TiOPc molecule on a gold surface has an orientation that is known through experimentation, the interaction strength can be altered to reflect that outcome. The other surfaces can then also be adjusted so their values can maintain the same relative strengths to the gold surface. For example, if the

original maximum value for gold was five times larger than that of the NaCl surface, the new maximum value for the NaCl surface would be $NaCl_{max} = \frac{Au_{max}}{5}$, with the rest of the NaCl being scaled accordingly. The other MPc gold surface interactions values can then be calibrated by applying the same theoretical idea that was applied to calibrate the TiOPc molecule on a gold surface. The other MPc surface values can therefore also be adjusted to maintain the relative strength to their respective gold values.

To investigate the effect of changing the strength of the surface values the minimum orientation angle, for each molecule, on a gold surface was recorded at incremental increases to the calibration. These increases were facilitated by scaling up values in the spanning equation from the previous chapter, equation (5.6). The design of this process meant that initially equation (5.6) would mimic the result of the SMCM calibration and over each iteration the program would increase the maximum value of the surface grids, whilst keeping the minimums the same. i.e. the molecular interaction grid and surface interaction grids would have the same minimum though out the calculations. The increase itself was performed by multiplying the maximum value of the surface grids in the SMCM calibration by a factor X, where X was increased from $10 - 30000$ in increments of 10. The Monte Carlo algorithm was performed at each step to record the minimum azimuthal angle at each point. Normally the number of iterations needed for the algorithm to effectively minimise are of the order 10^5, however, the number of simulations that this procedure would perform becomes incredibly cumbersome at that scale. As these results are to be used as a calibration tool, it is not necessary for the full sized simulation to be used. Under this consideration, the decision was taken to reduce the scope of the simulations by reducing the size of the monolayer from a 10×10 set to 5×5 set. Under these conditions. the scale of the required iterations of the Monte Carlo algorithm reduces by an order of magnitude to 10^4. To make the calibration effective the monolayer parameters were set to match the simulations performed in chapter five, (the same that are visually displayed in figure 5.6) so they are identical to those in the experimental data from the previously mentioned Mannsfeld and Fritz paper [6].

Increasing the strength of the surface interaction has the effect of making the system less capable of being able to effectively minimise. Thus, as the calibrated surface values increased, the probability of the standard deviation criteria for minimisation ($\sigma < 2$) being achieved decreased. For the purposes of this work, it is clearer to examine the data points where the program did meet that criterion. Figure 6.11 displays the minimization of each of the five MePc molecules, for each value of X where a minimum value was found. Note the program was designed to run ten simulations before moving onto the next point so at each point where no minimum was reached, the likely hood of that event occurring was small.

Examining the behaviour of each molecule in the calibration routine allows for the observation of the following. TiOPc, FePc, CoPc, and SnPc all start at the same value of around 31°, before increases in the surface potential interaction values result in a general decline in the angle for all molecules. The amount by



(a) CoPc

(b) TiOPc

(c) FePc

(d) SnPc and CuPc

Figure 6.11: Set of graphs showing the effect of increasing the span of the surface interaction by setting the equals minimum and multiplying the maximum value by a factor x, whilst keeping relative strengths between surface positions the same.

which each molecule decreases differs in each graph. TiOPc reduces quickly but then plateau's at the 30° mark before decreasing more gradually to 29°. CoPc follows essentially the same pattern initially but also has a second curve starting at around 29° before quickly reducing to 28°. FePc, on the other hand, has two curves, like CoPc, but does not follow the same plateau. SnPc essentially just reduces to 29° quickly and becomes stable but also has two curves in the initial phases. The CuPc molecules behaved entirely differently from the other four, firstly the minimization angle recorded is lower by about 2°, secondly, although there is a decrease along with the data it's smaller and less steep than the other molecules. This is most likely the result of the CuPc molecule being the most planar of the set, as the concave face each MPc is facing towards the surface (as discussed in section 5.7), the orbitals for the other MPcs are simply closer to the surface, increasing the interaction. This, in turn, means the surface has the least relative impact on the CuPc molecules in these simulations.

The figures display a number of occurrences that might not be expected to be observed in real systems such as the plateaus and the existence of two separate curves in some cases. These phenomena are explained as a result of the decreased scope of the simulations. Reducing the number of molecules in the grid from being a 10×10 set to a smaller 5×5 set has the effect of making the system more susceptible to align to a single value. For the cases where there are two curves, for instance, these occur because the ideal minimum is between those two values and so the simulation alters between minimizing to one or the other. The plateaus are explained following a similar line of reasoning at values where the minimum is close to 30° or 31° all the molecules will align to that single angle. This is in contrast to the case of for the full sized simulations, which aligned around that angle in such a way so the average of those angles can be taken as a better representation of the minimum. An effect of increasing the simulation not represented on the graphs (by choice, as expressed earlier in this section), is the tendency of the probability of a minimum value occurring to become less likely as the surface interaction values increase. For the case of those potentials increasing, the program becomes increasingly likely to find local minima that cause the routine to be unable to minimize any further. Whilst that information isn't expressed in the data sets above it will be used in upcoming calculations and therefore worth explaining.

Focusing on the graph for TiOPc in isolation initially, from figure 6.11b, it can be seen that the overall trend reduces and approaches a value of 29°. For a TiOPc molecule on a gold surface, with those surface parameters, the experimental data from from the Mannsfeld and Fritz paper[6] (recorded in table 5.2) measured the angle to be equal to 28 ± 2. This would imply that the best calibration would be to use a value of X which was at a maximum, as for this range is stable and closer to the experimental data. However, there is a trade off because a higher surface range increases the probability the simulation will not effectively minimise. Finding a point that balances these two aspects is the definition of an effective calibration in this system. Given that each graph follows similar trends for each molecule, it makes sense that the calculation used to calibrate TiOPc can be used to calibrate the other MPcs. Quantitatively deriving that point in a manner that is complete (fully satisfies all aspects of the system), is tricky as an 'acceptable' probability of minimisation is hard to define without some subjective element. With these considerations, the definition for an effective minimum probability was chosen to be the calibration value, X, at which more than 50% of the previous results had effectively minimised (met the standard deviation criteria, $\sigma < 2$). This is effectively describing the best calibration for gold for each molecule as the highest available value at which the system is in favour of producing a global minimum. The results of these calculations are presented in table 6.3:

These values do vary between molecules as they have different levels of interaction with the surface that can impact at different points, however, most criteria needed from the calibration are met. The system produces accurate results for TiOPc on a gold surface, each molecule is treated in the same vein meaning results should (and do) bear consistency with each other. Importantly each surface is distinguished from one another making it possible to compare the interactions.

Calibrations	Au		Ag		NaCl	
	Span Org	Span Cal	Span Org	Span Cal	Span Org	Span Cal
TiOPc	1.608	1.608×17840	1.980	1.980×1	2.180	2.180×512
SnPc	1.113	1.113×27840	1.803	1.803×1	1.604	1.604×725
FePc	1.376	1.376×14440	1.626	1.626×1	1.746	1.746×540
CuPc	1.287	1.287×30000	1.670	1.670×1	1.500	1.500×1112
CoPc	1.364	1.364×17840	2.067	2.067×1	1.692	1.692×703

Table 6.3: Table showing the multiplication factor used to calibrate the surface grid values.

Like with many other effects incorporated into this Hückel based simulation model, these surface interactions are useful but a relative measurement. The gold measures a strongly interacting surface, the interaction is significant but weaker on NaCl a surface and silver is modelled as a more weakly interacting surface. Given the methods of calculation in this work, the true nature of the interaction between the surface and molecular monolayer was never the intended goal. The intention was to provide an effective robust comparison between the MPc molecules on different surfaces, which is adequately provided using table 6.3 in conjunction with equation 5.6. Some notes on the table itself, the CuPc molecule has the smallest interaction value and as a result still maintained more the 50% minimisation at the maximum value of X, 30000 and so this value was taken instead. Similarly, the difference in strengths between gold and silver surfaces was on a scale of 10^5, meaning the relative difference from the gold and silver maximum values were less than one. In these cases, the values were approximated to be equal to one. The justification for these is that they create the opportunity to make a comparison between surfaces and whilst not a perfect representation have a basis in experimental data and produce reasonable simulations.

Before being able to perform these simulations some consideration needs to be given to the choice of super-lattice structures for the monolayer. In this instance, the decision was made to make the comparison with molecular lattice structures that were determined experimentally because it would provide the best way to isolate and study the surface effects. With this consideration, three superlattice structures were chosen and used to make the comparisons. Two from the previously discussed Mannsfield and Fritz paper [6] and another from a study conducted by Bobaru and Salomon *et al* in 2011 [7]. This last paper was an investigation examining the variety of surface structures found when depositing FePc onto silver. In their research three super-lattice structures were observed, each of these depending on how long the evaporation process used to deposit the FePc onto the surface had taken place. The exact details of that work are found in [7], looking into those structures, they have the form:

$$FePc1 = \begin{bmatrix} 5 & -1 \\ 5 & 6 \end{bmatrix}, FePc2 = \begin{bmatrix} 4.8 & -0.4 \\ 2.58 & 5.68 \end{bmatrix}, FePc3 = \begin{bmatrix} 5 & 0 \\ 3 & 6 \end{bmatrix} \tag{6.4}$$

When these are compared to the three super-lattices presented in the Mannsfield and Fritz paper [6]:

$$TiOPc1 = \begin{bmatrix} 5.169 & 4.975 \\ -2.589 & 2.995 \end{bmatrix}, TiOPc2 = \begin{bmatrix} 4.971 & -0.110 \\ 2.997 & 5.781 \end{bmatrix}, TiOPc3 = \begin{bmatrix} 0.110 & 4.971 \\ 5.781 & -2.997 \end{bmatrix} \quad (6.5)$$

it can be seen that of the six overlay structures displayed three describe only minor differences, each being slight variations on the structure:

$$\begin{bmatrix} 5 & 0 \\ 3 & 6 \end{bmatrix}$$

these being the structures labelled FePc2, FePc3, and TiOPc2. Therefore, these three can be condensed and used as one unique structure when utilized for examining the surface interaction. It is also noticed that the other two TiOPc structures, TiOPc2, and TiOPc3, from [6] are 90° rotations of one another and so can also be considered singularly. It is a given that there are more possible structures available, however, since this section is focused on examining the surface it is useful to examine a few given unique structures for which data exists to better isolate the effect of the surface. In summary, the three chosen super-lattices used in the simulations were:

$$TiOPc1 = \begin{bmatrix} 5.169 & 4.975 \\ -2.589 & 2.995 \end{bmatrix}, TiOPc2 = \begin{bmatrix} 4.971 & 0.110 \\ -2.997 & 5.781 \end{bmatrix} FePc1 = \begin{bmatrix} 5 & -1 \\ 5 & 6 \end{bmatrix} \quad (6.6)$$

Of the three FePc1 variations the TiOPc1 structure was chosen because the experimental data on the orientation in the Mannsfeld and Fritz paper [6] is more comprehensive and so makes for a better comparison. Each monolayer is now constructed by generating the molecules centre positions relative to the locations of the surface atoms. In each case, the surface reconstruction was completed using the (111) facet of the chosen material. Reconstructing Au and Ag in this manner reflects the surface constructions found in the experiments performed in [6, 7], both of which were used comparatively to access the validity of the model created for this work. Those choices of reconstruction were therefore uncomplicated to make and require little deliberation, however, using the (111) facet, for NaCl, has two caveats that need exploring to justify this decision. These are that: (i) NaCl(111) surface is rarely used experimentally, limiting the ability for any experimental comparison; (ii) this surface will be entirely made of either Cl^- or Na^+ ions, giving rise to polarising effects not accounted for in the calculations. For completeness, in the calculations, it was the chlorine atoms that were chosen to represent the surface. This choice was made primarily because the nature of that surface reconstruction bears the largest degree of similarity to the Au and Ag surfaces. In essence, the Au, Ag and NaCl (111) surfaces all have highly similar constructions in this model and so can be calibrated and compared more directly, having removed any phenomenon that might arise from changing the layout of the surface atoms.

Taking this set of input information with the data on the calibrations described in the calibration table above. It was possible to run simulations for each molecule, on each surface, in each structure and record the minimization. The data on each minimum is presented in tables 6.4, 6.5 amd 6.6.

SL = TiOPc1	Au		Ag		NaCl	
	angle	std	angle	std	angle	std
TiOPc	60.3	1.31	60.1	2.28	58.2	10.68
SnPc	60.0	1.58	59.8	1.98	57.4	9.62
FePc	59.9	3.24	60.0	0.90	57.6	9.60
CuPc	58.1	6.03	58.1	5.66	56.7	9.81
CoPc	60.2	1.65	60.0	1.85	56.3	10.42

Table 6.4: Table of results displaying the average angle and standard deviation for each MPc on Au, Ag and NaCl in the TiOPc1 overlayer structure.

SL = TiOPc2	Au		Ag		NaCl	
	angle	std	angle	std	angle	std
TiOPc	29.4	1.94	31.1	0.76	17.1	10.61
SnPc	29.3	0.46	30.9	0.76	16.9	9.41
FePc	28.3	0.48	31.5	1.07	16.8	10.33
CuPc	28.3	0.46	28.3	0.69	17.5	10.37
CoPc	28.3	0.46	31.5	0.9	15.7	10.33

Table 6.5: Table of results displaying the average angle and standard deviation for each MPc on Au, Ag and NaCl in the TiOPc2 overlayer structure.

SL = FePc1	Au		Ag		NaCl	
	angle	std	angle	std	angle	std
TiOPc	75.3	2.47	65.4	0.73	48.1	22.12
SnPc	75.3	3.14	65.0	1.21	47.9	23.08
FePc	74.8	3.04	65.3	1.01	46.1	22.39
CuPc	75.8	3.52	65.3	0.64	46.2	21.74
CoPc	75.1	3.26	65.2	0.91	47.9	22.65

Table 6.6: Table of results displaying the average angle and standard deviation for each MPc on Au, Ag and NaCl in the FePc1 overlayer structure.

Staying consistent with all the simulations so far the angle is measured clockwise relative to the horizontal axis. Given the lattice structures themselves are rotated, this creates a question, of what the various structural phases in the above table actually look like. For completeness each of the forty five results in tables 6.4, 6.5 and 6.6 are displayed as an STM simulation in appendix II, here a sampling is presented in figure 6.12. In tables 6.4, 6.5 and 6.6, and subsequently figure 6.12, six identifiable phases occur for various reasons that will be examined. For

convenience when referring to them, these phases are labelled in terms of the superstructure and respective surface creating; TiOPc1-AuAgNaCl, TiOPc2-AuAg, TiOPc2-NaCl, FePc1-Au, FePc1-Ag, and FePc1-NaCl. Starting by referring back to the experimental data the fundamental question to be answered first is how do these compare to the results of those experiments.

The case for TiOPc in TiOPc2-Au phase has already been well established in the last chapter, this section shall consider FePc in the FePc1-Ag phase, as was put forward in the paper [7]. The primary reason why this comparison was not previously considered because in that paper the angular alignment was not measured quantitatively. They did, however, use STM techniques to image the monolayer which provides qualitative information that makes for an effective comparison. In that work, the picture put forward for FePc in the FePc1-Ag setting was of molecules that position themselves so two adjacent FePc molecules would lie with their centres open towards each other, an image that is mirrored in the simulated version displayed in 6.12e. Although this comparison would ideally come available with more quantitative analysis, the result provides further indication that the simulated results are an accurate portrayal of self-assembly behaviour. The case of TiOPc on a gold surface in the TiOPc1 structure has the same requirements that must be fulfilled, to be considered valid, as the TiOPc on gold in the TiOPc2 case, the azimuthal angle θ must be between 26° and 30°, as this is experimentally measured to be $\theta = 28° \pm 2°$ (Recorded in table 5.2). Given that the angle between rows, Γ, in the TiOPc1 overlayer structure is also 90° the calculated angle θ is equivalent to the angle $90° - \theta$, as a result of the four fold symmetry of the system. Therefore, the angle for TiOPc on a gold surface in the TiOPc1 structure that is recorded as 60.4°, in table 6.4, is structurally identical to 29.6° and is, therefore, also within the experimental parameters recorded in [6].

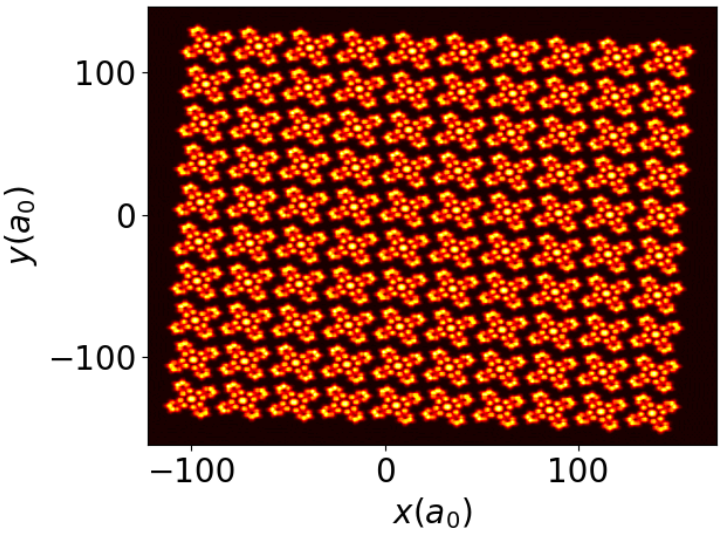
Examining the rest of the results presented in the tables above, it can be seen that in general molecules on the NaCl surface do not align as rigidly as on the other surfaces, as indicated from the higher standard deviation values. The reason for this is a result of the distance relation described earlier in this chapter. Comparing the lattice vectors, i and i, these are significantly larger in the NaCl crystal, the lattice vectors of the (111) surface of each being 2.87Å for Ag, 2.88Å for Au and 3.948Å for NaCl [110, 116]. Therefore, placing the same overlayer structures on an NaCl (111) surface results in a significantly higher molecular spacing in the monolayer, as can be seen in the figures 6.12c and 6.12f. Given that effect, it is clear that the molecular orientation angle in the TiOPc1 lattice structure only displays relatively minor differences due to the change in the central metal atom across all surfaces. TiOPc, SnPc, FePc and CoPc all around 60° and CuPc close to 58°. The lower average value for molecules on the NaCl surface is explained by the increased randomness in the system. This is because the reduced stability of the angle positions causes mean angle approaches 45°. The mode of alignment for the MePc monolayers in the TiOPc1 arrangement display the same fundamental behaviour on NaCl as other surfaces but with less of an incentive to be as rigidly aligned. Given this, some consideration should be given to the plausibility of these structures being observed in real experiments as monolayers of

these spacial dimensions would be hard to replicate. However, the effects themselves are useful here, as they aid to portray the ranges between which surface forces overcome intermolecular forces, creating a potentially interesting marker that could theoretically be analysed experimentally.

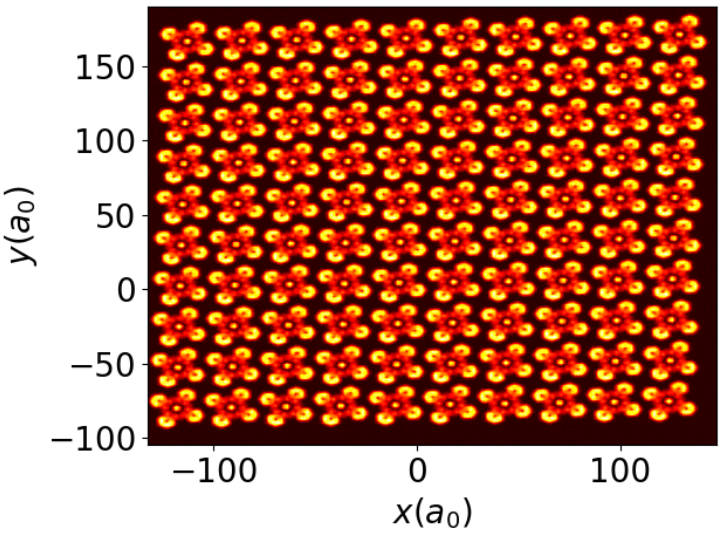
The TiOPc2 structure results in two different modes of alignment; the molecules on the Au and Ag surfaces are considered to display the same form of minimization observed in the previous chapter. The difference seen between gold and silver is the same as the variation that is observed when documenting the calibration results. Therefore, the slight variation in the molecular orientation of the monolayer when comparing a gold and silver surface is concluded to be a consequence of the increased surface interaction. The molecules on the NaCl surface, however, align themselves as to be positioned so that the tip of the benzene rings are facing towards each other, figure 6.12c. This interaction is clearly at a higher energy cost than those that would result from other orientations, it is therefore concluded these are a result of molecules aligning to reduce the surface interaction.

The FePc1 lattice interestingly provides three different modes between the surfaces. The Ag surface alignment, as previously stated, is again similar between the molecules and that of the description found in [7]. The molecules in the FePc1-Au phases align approximately 10° away from those on the Ag surface, this is concluded to be an effect of the strongly interacting surface that creates some deviation away from the optimum molecular position. The FePc1-NaCl results for this surface are essentially random, the distances in this super-lattice have approached a point where there just isn't enough of an incentive for neighbouring molecules to align.

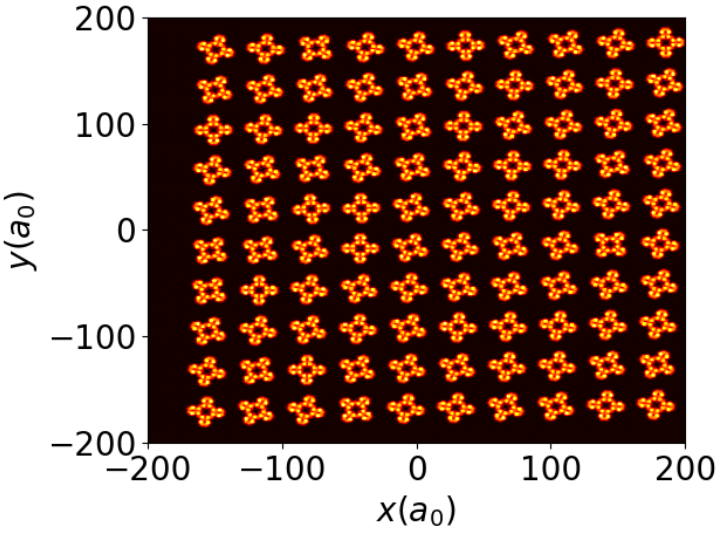
The variety of phases found in these simulations helps demonstrate the effectiveness of this method for providing usable outcomes. The flexibility of these methods allow for a multitude of parameters to be altered and compared whilst still being reliable and robust enough to be validated by experimental data where it is available. Overall this would not be achievable without relying on the semi-empirical Hückel methods on which these calculations are based.



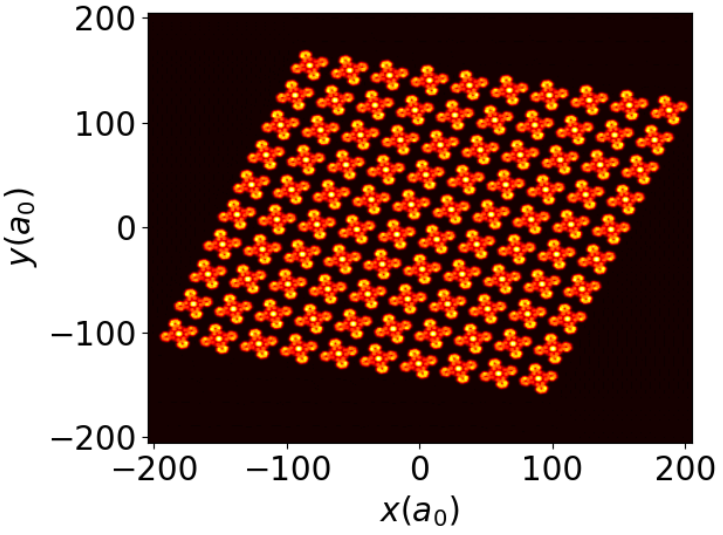
(a) TiOPc in TiOPc1-AuAgNaCl



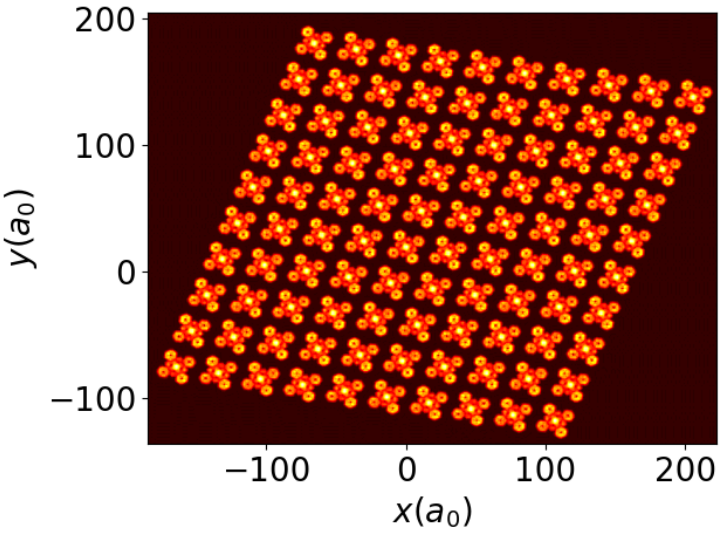
(b) SnPc in TiOPc2-AuAg



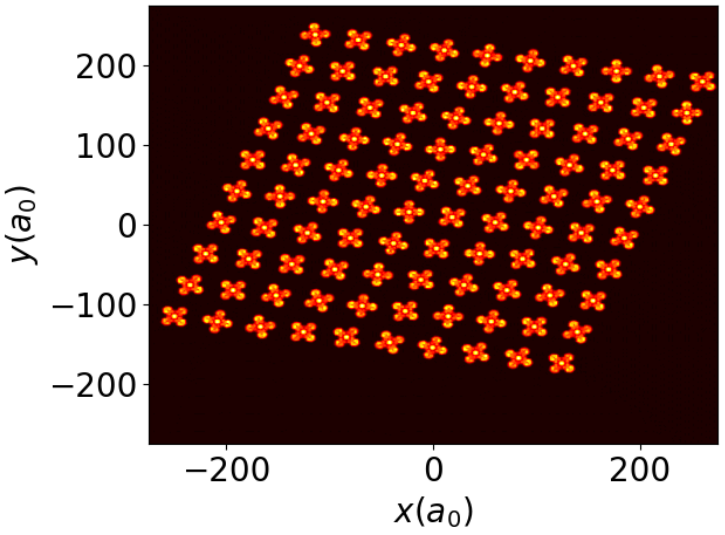
(c) CuPc in TiOPc2-NaCl



(d) CoPc in FePc-Au



(e) FePc in FePc-Ag



(f) CoPc in FePc-NaCl

Figure 6.12: STM simulations of the six identifiable phases in the set of surface comparisons.

Chapter 7

Conclusion

The first objective of this thesis was to demonstrate the effectiveness of an adapted Hückel method technique for studying the self-assembly of organic molecules, building on the work previously completed by J. Leaf *et al* in 2016 [5]. The second was to compare different potential monolayers of phthalocyanine molecules by using a Monte Carlo metropolis algorithm to deduce information about their respective alignments of different surfaces. In most regards, these outcomes were broadly achieved. The predicted alignments from the simulations closely resembled the experimental data where it was available, in [6] and [7], and the results of the simulations in section 6.4 displayed a comprehensive investigation of monolayer alignments even making novel predictions about various structures. This demonstrated the model to be an effective and useful tool, despite, having a relatively crude set up.

Even with the success in terms of predicting results, there is considerable space left for improvements in this system. There are the previously mentioned ideas around improving the MO calculations, yet these would mainly be of interest as a comparative measure because ideally the method for the electron density approximations would remain as simple as possible. Most notably this system could be broadened and improved by considering two details, bonding interactions and surface orbitals. This model assumes all interaction between electronic orbitals will be minimized and so the molecules will assemble in order to reduce the overlap. However, this detail is not always true, as is indicated from the overlayer structures presented in [7]. In that paper, two of super-lattices are commensurate i.e. all the matrix values in the super-lattice formula, equation (6.1), are integers so the centres of the molecules lie exactly on the surface points, which in turn implies some form of bonding. This effect also can apply to the molecular interactions themselves and not just between the molecules and the surface. For instance, the TiOPc molecules have a well known attractive dipole-dipole interaction [117], that is again not included in this model. Generally, these calculations assume the orientation is found from reducing the electron overlap, however, molecular systems can, and often do, produces attractive forces and more complex interactions that are not formulated in this method. Some measures that would work to include such features would greatly improve the diversity of molecular systems available to be studied.

The methods used for understanding the electronic features of the molecules were tried and tested ideas with many years worth investigative studies behind them ([118, 119]) just to name a few). This is not the case with how the surface model was constructed. The thought process behind the idea of building the surface model solely out of p_z orbitals was to see what results could be achieved with the simplest approach and then building the surface model further if it was found to be unsuitable. As the results show the methodology had plenty of merits, nonetheless this aspect of the calculation was constructed in a mostly ad hoc and somewhat primitive manner. Thus, there is potential for a more robust form of calculation that allows for a more detailed analysis of the surface interactions.

The development of both those ideas would have been at the forefront of the work in building the program, and are both mentioned in appendix I that describes issues with improving the ability of the software to make more fundamental predictions. In the simulations performed in the last chapter, the overlayer lattices were taken from experimental data so the alignment could be tested. It would have been ideal be able to start with just the surface and the molecular information and then be able to predict the structure of the super-lattices as well as the angular alignment. The ideas for how this was attempted and why it was unsuccessful are described in appendix II. Nonetheless, improving both the surface model and making the interaction model more accurate are areas that would have certainly been investigated as a way of fixing some issues with deriving the overlay lattice structures. Other scenarios of interest that this model could also be applied to include other organic molecules and possibly even multi-layer structures. Adapting the system to study other molecules such as porphyrins would be a relatively straight forward process and certainly an area that would have been adapted in a continuation of this study. Investigating multilayer arrangements is a much more speculative idea, however, the suggestion is based on precedent, as the paper which this study is based on [5] was successfully adapted to suit such a task, therefore those ideas could be theoretically adapted to attempt multilayer simulations for these systems.

Whilst much of the discussion in chapter three centred around potential modelling choices that were not used for this work, among those ideas are techniques that are used to perform similar simulations. In particular classical force field calculations or a periodic DFT model could (and have [115, 120]) also been used to adequately simulate the monolayers studied in chapters five and six. The mention of these two, in particular, is because either one could have been performed with similar computational effort to the Hückel method applied in this work. There are particular advantages and disadvantages to both of these choices, as there are for a Hückel based model (see chapter 3.2), however, the purposes of this work was to show that a Hückel method approach is a valid and potentially powerful tool that can compete with those ideas. Whilst other modelling choices could have been made for this work the Hückel method techniques applied to generate these simulations are demonstrated to be a legitimate idea for modelling on this scale.

A question that surfaces when considering the success of the model is what does this say about the physics of these materials. To answer that question it is necessary to understand what exactly the model is based on and what it is evaluating. The Hückel based simulation techniques, applied in this work, make a number of assumptions to be able to generate an estimation of the electron density in an orbital model of molecules; that is then used to predict alignment structures by minimising this property. From this basis, it is concluded that, firstly, in the set of circumstances presented those assumptions are reasonable because if they were unsound the density estimations would have been invalid and produced poor results. Secondly, the interaction energy of these systems is directly related to electron overlap because it is by minimizing this property that the materials demonstrate realistic observations. These lead to the conclusions that despite the relatively simplistic approach this model can be adapted and used successfully for identifying and predicting real-world physical processes.

Appendix A

Structure of the Simulation Software

Throughout this thesis there are multiple references to the software or the program, these are generic terms used to describe the collection of functions created for performing calculations. For completeness this section will describe an overview of how these functions are connected to one another by giving a brief description of the program's structure.

The first point of note when considering the code that was written for this work, is that it does not, in its current form, have any kind of front-end development. This means that different routines, such as the grid-file calculations, have to be compiled and run separately from the simulation algorithms. It also means that changes to a number of the simulation and calculation parameters had to be made inside the code itself which is generally less efficient than having these variables available to be chosen from an interface. If this collection of routines was to be made into a fully cohesive piece of software, creating a front-end would be necessary. However, it was the calculations themselves that were the most important aspect of this work and as such given precedence.

The program consists of four routines: the EHT calculation, the intermolecular grid-file calculation, the surface-molecule grid-file calculation and Monte Carlo simulation algorithm routine. The functions that produce the STM simulations are called from a separate file but are not run independently instead they are used inclusively in the EHT routine and the Monte Carlo simulation functions. Whilst each routine is compiled and run independently their results are utilized in the next set of functions and so there is a continuous structure to the program that can be identified. A flow diagram of how these routines relate to one another is displayed in figure 7.1.

Each routine took a number of variables from the so called info files as inputs. These files contained all the information that exclusively related to each unique molecule i.e. one info file is named CoPc-info and contains the inter-atomic spacings, the GTO coefficients and VSIP values for CoPc and other variables related the the number of each type of orbital. All of these are essential for the EHT routine, however, only the GTO coefficients and the inter-atomic spacings are carried through into the other parts of the program, as these are necessary for the LCAO construction of the molecule.

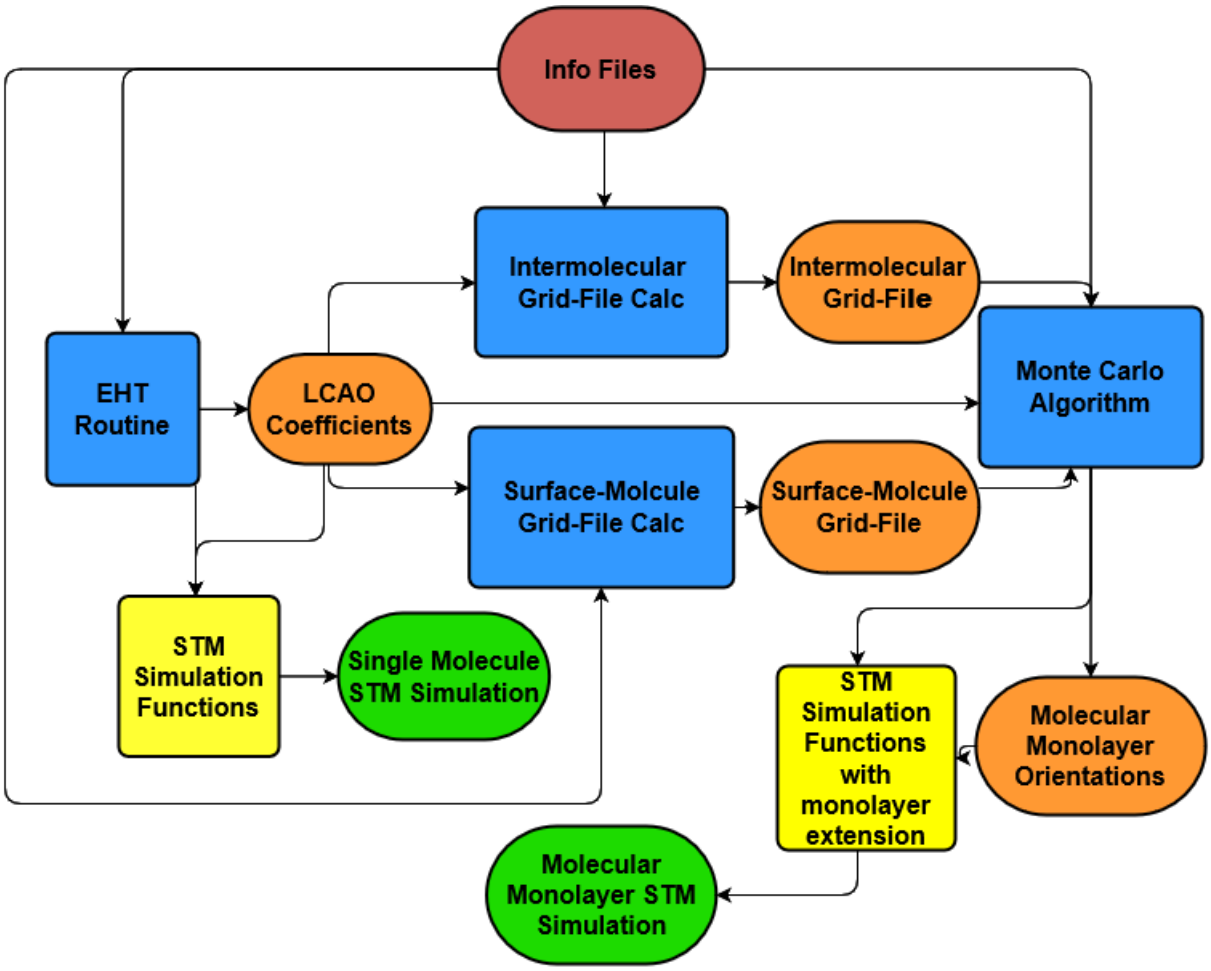


Figure 7.1: Flow diagram showing the structure of program used in this thesis. The oval workflow symbols represent data, with the colour coding used to denote whether its used as input or output: red for solely input information, amber for data that is both outputted by the program and then used as input later on, and green represents the information that is outputted by the program only. The rectangular workflow symbols represent routines, with the blue rectangles portraying the routines that compiled independently, and the yellow rectangles represent the set of functions that called as part of a larger routine.

The EHT routine consists of three main functions, the overlap calculation, the Hamiltonian matrix calculation and the eigenfunction calculation, as described in chapters 3 and 4. In terms of the structure of the program the eigenvectors and their respective eigenvalues are stored into a variable that is saved and loaded into following routines, the STM simulation functions can be called along with these results in order to check the EHT routine. The grid-file calculations both load the same inputs from outside the program, the info-files and the EHT calculations results are needed for creating the molecular orbitals of the molecule being studied. Both routines also include a set of functions to determine, and then iterate through, all calculable positions that were deemed necessary for the monolayer simulations before saving them in a file that can be utilized in later parts of the program. Within the surface-molecule grid-file routine is included a set of calculations that are necessary to construct the surface which adds an extra

degree of complexity to that routine, as discussed in chapter 5. All of these results are included into the Monte Carlo simulation algorithm, the grid-files make up the data by which the simulation is founded and the EHT routines are necessary for as an input for the STM functions which are called from the simulation routine. The final output being the STM simulations which incorporate the results of every routine in the program. The conceptual creation of this algorithm is described in chapter 3, practically there were a number of functions that were included to make the system efficient and effective, such as a system of tracking all the variables that are affected over the simulation, a method for locating the correct value from the grid-files and an implementation of the surface calibrations discussed in chapter 5.

In general each routine was well constructed for performing its given role, however, the program would need more work for each set of functions to be properly integrated. The creation of this code was intended to demonstrate the validity of a Hückel method approach, not to create a software package that is easily distributed and reusable, although that would be a more ideal end product. Thus, further iterations of the program would need to be created before that would be possible.

Appendix B

Results of Extended Hückel Calculations on Benzene

In chapter four an extended Hückel theory calculation was performed on benzene and the detail of those calculations was simply too large to be portrayed in that section. However, given the detailed discussion of those results, it was considered important that a version of them was available. A partial version of the overlap matrix, Hamiltonian, eigenvalues and eigenvectors are presented on pages 102, 103 and 104 respectively.

The overlap, Hamiltonian and eigenvectors are represented in a matrix of dimensions 30×18 because the full 30×30 matrix was too large to practically display. However, the matrices provide enough detail to access that EHT calculations have been correctly implemented and that the requisite assumptions about benzene, with regards to the BHM, are applicable.

For the overlap and Hamiltonian matrices, each element represents an interaction between two orbitals, with diagonal elements being the positions where a particular orbital interacts with itself. There are thirty orbitals in total and six of each type of valence orbital, with this consideration, the overlap and Hamiltonian have the following construction. The first six rows represent the interactions of the hydrogen 1s orbitals, the next set of six are the carbon 2s orbitals, the rest of the eighteen rows, in order, are the $2p_x$, $2p_y$ and $2p_z$ orbitals respectively. The columns have an identical set up and so the matrix is symmetric, with term accounted for twice i.e. matrix element $(13, 20)$ is identical to $(20, 13)$. The overlap and Hamiltonian matrices presented here are each missing the first twelve coulombs in order to comfortably display the important detail. From the overlap matrix, it can be seen that the diagonal elements are equal to one, this provides a good check that the functions performing that calculation are working correctly, as the overlap of an orbital with itself should be identical to the normalisation. As was discussed in chapter four, it can also be seen that there is no mixing between the p_z orbitals and any other orbital. This is displayed in the last six rows (and columns), which are equal to zero, everywhere, except for the last block of terms that represent the overlap between the p_z orbitals. The resulting Hamiltonian maintains many of those features, the lack of mixing between p_z orbitals and the diagonal elements now represented by the relevant VSIP value for a given orbital. The last block, representing the interaction between the p_z orbitals, also demonstrates the symmetric property of the matrix which again portrays that the functions used for this calculation are implemented correctly.

The results of diagonalizing that Hamiltonian are presented on page 100, each eigenvalue is presented in a 30×1 on the left hand side of the page whilst the corresponding eigenvectors are lined against each eigenvalue in an adjacent 30×18 matrix. These are a partial representation of the eigenvectors, missing the first twelve coulombs that correspond to the coefficients of carbon $s2$ and hydrogen $s1$ orbitals for that given MO. The important details of those matrices are that: each of BHM MOs are present, with some slight deviations that account for the inclusion of the overlap matrix; and the HOMO-LUMO gap lies between the same states as it did for the BHM. The results of the BHM and EHT are therefore consistent, leading to the conclusion that the EHT program is correctly implemented and can be applied to more complex molecules.

$$
\begin{bmatrix}
-0.464 & -0.099 & -0.099 & -0.012 & -0.012 & -0.005 & 0. & -0.067 & 0.067 & -0.004 & 0.004 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0.008 & -0.232 & -0.002 & -0.108 & -0.002 & -0.01 & 0.12 & 0.402 & 0.012 & 0.053 & 0.004 & 0.008 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.008 & -0.002 & -0.232 & -0.002 & -0.108 & -0.01 & -0.12 & -0.012 & -0.402 & -0.004 & -0.053 & -0.008 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.01 & 0.108 & 0.002 & 0.232 & 0.002 & -0.008 & 0.008 & 0.053 & 0.004 & 0.402 & 0.012 & 0.12 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.01 & 0.002 & 0.108 & 0.002 & 0.232 & -0.008 & -0.008 & -0.004 & -0.053 & -0.012 & -0.402 & -0.12 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.005 & 0.012 & 0.012 & 0.099 & 0.099 & 0.464 & 0. & -0.004 & 0.004 & -0.067 & 0.067 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
-0. & -0.193 & -0.193 & -0.077 & -0.077 & -0.044 & 0. & -0.334 & 0.334 & -0.044 & 0.044 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0.193 & -0. & 0. & -0.385 & -0.022 & -0.077 & 0.334 & -0. & 0.089 & -0. & 0.038 & 0.044 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.193 & 0. & -0. & -0.022 & -0.385 & -0.077 & -0.334 & -0.089 & 0. & -0.038 & 0. & -0.044 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.077 & 0.385 & 0.022 & 0. & -0. & -0.193 & 0.044 & -0. & 0.038 & -0. & 0.089 & 0.334 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.077 & 0.022 & 0.385 & -0. & 0. & -0.193 & -0.044 & -0.038 & 0. & -0.089 & 0. & -0.334 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.044 & 0.077 & 0.077 & 0.193 & 0.193 & 0. & 0. & -0.044 & 0.044 & -0.334 & 0.334 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
1. & 0.076 & 0.076 & -0.079 & -0.079 & -0.059 & 0. & -0.236 & 0.236 & -0.06 & 0.06 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0.076 & 1. & 0.025 & -0.332 & -0.007 & -0.079 & -0.236 & -0. & 0. & -0. & 0.03 & 0.06 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.076 & 0.025 & 1. & -0.007 & -0.332 & -0.079 & 0.236 & -0. & 0. & -0.03 & 0. & -0.06 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.079 & -0.332 & -0.007 & 1. & 0.025 & 0.076 & -0.06 & 0. & -0.03 & 0. & -0. & 0.236 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.079 & -0.007 & -0.332 & 0.025 & 1. & 0.076 & 0.06 & 0.03 & -0. & 0. & -0. & -0.236 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.059 & -0.079 & -0.079 & 0.076 & 0.076 & 1. & 0. & 0.06 & -0.06 & 0.236 & -0.236 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & -0.236 & 0.236 & -0.06 & 0.06 & 0. & 1. & -0.196 & -0.196 & -0.01 & -0.01 & 0.01 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.236 & -0. & -0. & 0. & 0.03 & 0.06 & -0.196 & 1. & -0.113 & 0.212 & -0.042 & -0.01 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.236 & 0. & 0. & -0.03 & -0. & -0.06 & -0.196 & -0.113 & 1. & -0.042 & 0.212 & -0.01 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.06 & -0. & -0.03 & 0. & 0. & 0.236 & -0.01 & 0.212 & -0.042 & 1. & -0.113 & -0.196 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.06 & 0.03 & 0. & -0. & -0. & -0.236 & -0.01 & -0.042 & 0.212 & -0.113 & 1. & -0.196 & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & 0.06 & -0.06 & 0.236 & -0.236 & 0. & 0.01 & -0.01 & -0.01 & -0.196 & -0.196 & 1. & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 1. & 0.212 & 0.212 & 0.025 & 0.025 & 0.01 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.212 & 1. & 0.025 & 0.212 & 0.01 & 0.025 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.212 & 0.025 & 1. & 0.01 & 0.212 & 0.025 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.025 & 0.212 & 0.01 & 1. & 0.025 & 0.212 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.025 & 0.01 & 0.212 & 0.025 & 1. & 0.212 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.01 & 0.025 & 0.025 & 0.212 & 0.212 & 1.
\end{bmatrix}
$$

$$
\begin{bmatrix}
0.373 & 0.08 & 0.08 & 0.009 & 0.009 & 0.004 & -0. & 0.054 & -0.054 & 0.004 & -0.004 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.007 & 0.186 & 0.002 & 0.087 & 0.002 & 0.008 & -0.096 & -0.323 & -0.01 & -0.042 & -0.003 & -0.006 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.007 & 0.002 & 0.186 & 0.002 & 0.087 & 0.008 & 0.096 & 0.01 & 0.323 & 0.003 & 0.042 & 0.006 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.008 & -0.087 & -0.002 & -0.186 & -0.002 & 0.007 & -0.006 & -0.042 & -0.003 & -0.323 & -0.01 & -0.096 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.008 & -0.002 & -0.087 & -0.002 & -0.186 & 0.007 & 0.006 & 0.003 & 0.042 & 0.01 & 0.323 & 0.096 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.004 & -0.009 & -0.009 & -0.08 & -0.08 & -0.373 & -0. & 0.004 & -0.004 & 0.054 & -0.054 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
0. & 0.203 & 0.203 & 0.081 & 0.081 & 0.046 & -0. & 0.352 & -0.352 & 0.047 & -0.047 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.203 & 0. & -0. & 0.406 & 0.023 & 0.081 & -0.352 & 0. & -0.093 & 0. & -0.04 & -0.047 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.203 & -0. & 0. & 0.023 & 0.406 & 0.081 & 0.352 & 0.093 & -0. & 0.04 & -0. & 0.047 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.081 & -0.406 & -0.023 & -0. & 0. & 0.203 & -0.047 & 0. & -0.04 & 0. & -0.093 & -0.352 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.081 & -0.023 & -0.406 & 0. & -0. & 0.203 & 0.047 & 0.04 & -0. & 0.093 & -0. & 0.352 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.046 & -0.081 & -0.081 & -0.203 & -0.203 & -0. & -0. & 0.047 & -0.047 & 0.352 & -0.352 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.419 & -0.056 & -0.056 & 0.058 & 0.058 & 0.043 & -0. & 0.173 & -0.173 & 0.044 & -0.044 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.056 & -0.419 & -0.018 & 0.244 & 0.005 & 0.058 & 0.173 & 0. & -0. & 0. & -0.022 & -0.044 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.056 & -0.018 & -0.419 & 0.005 & 0.244 & 0.058 & -0.173 & 0. & -0. & 0.022 & -0. & 0.044 & -0. & -0. & -0. & -0. & -0. & -0. \\
0.058 & 0.244 & 0.005 & -0.419 & -0.018 & -0.056 & 0.044 & -0. & 0.022 & -0. & 0. & -0.173 & -0. & -0. & -0. & -0. & -0. & -0. \\
0.058 & 0.005 & 0.244 & -0.018 & -0.419 & -0.056 & -0.044 & -0.022 & 0. & -0. & 0. & 0.173 & -0. & -0. & -0. & -0. & -0. & -0. \\
0.043 & 0.058 & 0.058 & -0.056 & -0.056 & -0.419 & -0. & -0.044 & 0.044 & -0.173 & 0.173 & -0. & -0. & -0. & -0. & -0. & -0. & -0. \\
-0. & 0.173 & -0.173 & 0.044 & -0.044 & -0. & -0.419 & 0.144 & 0.144 & 0.007 & 0.007 & -0.008 & -0. & -0. & -0. & -0. & -0. & -0. \\
0.173 & 0. & 0. & -0. & -0.022 & -0.044 & 0.144 & -0.419 & 0.083 & -0.156 & 0.031 & 0.007 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.173 & -0. & -0. & 0.022 & 0. & 0.044 & 0.144 & 0.083 & -0.419 & 0.031 & -0.156 & 0.007 & -0. & -0. & -0. & -0. & -0. & -0. \\
0.044 & 0. & 0.022 & -0. & -0. & -0.173 & 0.007 & -0.156 & 0.031 & -0.419 & 0.083 & 0.144 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0.044 & -0.022 & -0. & 0. & 0. & 0.173 & 0.007 & 0.031 & -0.156 & 0.083 & -0.419 & 0.144 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0. & -0.044 & 0.044 & -0.173 & 0.173 & -0. & -0.008 & 0.007 & 0.007 & 0.144 & 0.144 & -0.419 & -0. & -0. & -0. & -0. & -0. & -0. \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.419 & -0.156 & -0.156 & -0.018 & -0.018 & -0.008 \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.156 & -0.419 & -0.018 & -0.156 & -0.008 & -0.018 \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.156 & -0.018 & -0.419 & -0.008 & -0.156 & -0.018 \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.018 & -0.156 & -0.008 & -0.419 & -0.018 & -0.156 \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.018 & -0.008 & -0.156 & -0.018 & -0.419 & -0.156 \\
-0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0. & -0.008 & -0.018 & -0.018 & -0.156 & -0.156 & -0.419
\end{bmatrix}
$$

APPENDIX B

$$
\begin{bmatrix} 1.483 \\ 1.029 \\ 1.029 \\ 0.683 \\ 0.683 \\ 0.363 \\ 0.347 \\ 0.247 \\ 0.247 \\ 0.22 \\ 0.058 \\ 0.058 \\ -0.222 \\ -0.327 \\ -0.327 \\ -0.466 \\ -0.466 \\ -0.496 \\ -0.496 \\ -0.522 \\ -0.527 \\ -0.548 \\ -0.548 \\ -0.603 \\ -0.633 \\ -0.736 \\ -0.736 \\ -0.939 \\ -0.939 \\ -1.066 \end{bmatrix}
\begin{bmatrix}
0. & 0. & 0. & -0. & -0. & -0. & 0. & -0. & 0. & -0. & 0. & -0. & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.01 & 0.05 & 0.02 & 0.02 & 0.05 & -0.01 & 0.03 & 0.03 & -0.03 & -0.03 & 0.03 & 0.03 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.01 & 0.01 & -0.05 & -0.05 & 0.01 & 0.01 & 0.06 & -0.01 & 0.02 & 0.02 & -0.01 & 0.06 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.1 & -0.13 & -0.09 & 0.09 & 0.13 & 0.1 & 0.03 & -0.11 & 0.12 & -0.12 & 0.11 & -0.03 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.02 & -0.1 & 0.14 & -0.14 & 0.1 & -0.02 & 0.17 & 0.05 & 0.01 & -0.01 & -0.05 & -0.17 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.16 & 0.08 & 0.08 & 0.08 & 0.08 & -0.16 & -0. & -0.14 & 0.14 & 0.14 & -0.14 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.23 & -0.11 & -0.11 & 0.11 & 0.11 & 0.23 & 0. & 0.2 & -0.2 & 0.2 & -0.2 & -0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0.1 & -0.2 & 0.2 & 0.2 & -0.2 & 0.1 & -0.12 & 0.17 & 0.29 & 0.29 & 0.17 & -0.12 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.35 & 0.05 & -0.06 & -0.06 & 0.05 & 0.35 & 0.03 & -0.27 & 0.14 & 0.14 & -0.27 & 0.03 & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & 0.29 & -0.29 & -0.29 & 0.29 & 0. & -0.33 & 0.17 & 0.17 & 0.17 & 0.17 & -0.33 & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & -0.34 & -0.34 & -0.34 & -0.34 & -0.34 & -0.34 \\
-0.26 & 0.34 & 0.33 & -0.33 & -0.34 & 0.26 & -0.03 & 0.07 & -0.02 & 0.02 & -0.07 & 0.03 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.02 & 0.02 & -0.07 & 0.07 & -0.02 & -0.02 & -0.36 & 0.28 & 0.29 & -0.29 & -0.28 & 0.36 & 0. & 0. & 0. & 0. & 0. & 0. \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.53 & 0.27 & 0.27 & -0.27 & -0.27 & -0.53 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & -0.46 & 0.46 & -0.46 & 0.46 & 0. \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.66 & -0.3 & -0.35 & -0.35 & -0.3 & 0.66 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.03 & -0.58 & 0.55 & 0.55 & -0.58 & 0.03 \\
0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0. & 0.52 & -0.52 & -0.52 & 0.52 & 0.52 & -0.52 \\
-0.01 & -0. & -0.01 & -0.01 & -0. & -0.01 & -0.51 & -0.52 & -0.52 & -0.52 & -0.52 & -0.51 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.52 & 0.51 & 0.51 & 0.51 & 0.51 & 0.52 & -0.01 & -0.01 & -0. & -0. & -0.01 & -0.01 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.32 & 0.16 & 0.16 & -0.16 & -0.16 & -0.32 & -0. & -0.28 & 0.28 & -0.28 & 0.28 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\
0.04 & -0.4 & 0.39 & -0.39 & 0.4 & -0.04 & 0.14 & 0.57 & 0.53 & -0.53 & -0.57 & -0.14 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.78 & -0.07 & -0.11 & 0.11 & 0.07 & -0.78 & -0.01 & 0.37 & -0.43 & 0.43 & -0.37 & 0.01 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0. & 0.66 & -0.66 & 0.66 & -0.66 & -0. & 0.76 & 0.38 & 0.38 & -0.38 & -0.38 & -0.76 & 0. & 0. & 0. & 0. & 0. & 0. \\
0.55 & -0.28 & -0.28 & -0.28 & -0.28 & 0.55 & 0. & 0.48 & -0.48 & -0.48 & 0.48 & -0. & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.15 & 0.65 & 0.15 & 0.15 & 0.65 & -0.15 & 0.46 & 0.34 & -0.29 & -0.29 & 0.34 & 0.46 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.12 & -0. & 0.63 & 0.63 & -0. & -0.12 & -0.58 & 0.22 & -0.28 & -0.28 & 0.22 & -0.58 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.09 & -0.38 & 0.13 & -0.13 & 0.38 & 0.09 & 0.41 & -0.09 & 0.15 & -0.15 & 0.09 & -0.41 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.18 & -0.14 & -0.39 & 0.39 & 0.14 & 0.18 & -0.2 & -0.27 & 0.24 & -0.24 & 0.27 & 0.2 & 0. & 0. & 0. & 0. & 0. & 0. \\
-0.39 & 0.2 & 0.2 & 0.2 & 0.2 & -0.39 & -0. & -0.34 & 0.34 & 0.34 & -0.34 & -0. & 0. & 0. & 0. & 0. & 0. & 0.
\end{bmatrix}
$$

Appendix C

Calculation of Molecular Super-Lattice Sructures

For all the simulations discussed in the main body of the thesis the information about the molecular layout of the monolayers were either been adopted from experiential results or isolated and iterated over a given range. A feature that was explored towards the end of work was to attempt to use the information about the surface a molecular interactions to predict the overlayer structure. The idea was to initialise a set of molecules alongside a surface of N atoms and adapt the Monte Carlo algorithms to attempt to determine a minimum structure as well as their orientation. These attempts had a limited success. To build the program the simulation was scaled and started by allowing the centres to move a few places in any direction across a series of surface points similar to those used for the grid calculations. This was done whilst also keeping them bounded in between a small section in between a set of atoms, as can be seen in figure 7.1.
.

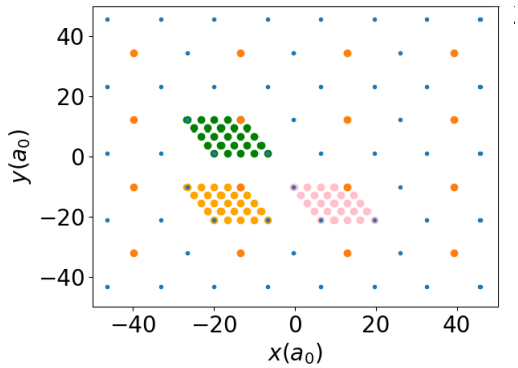


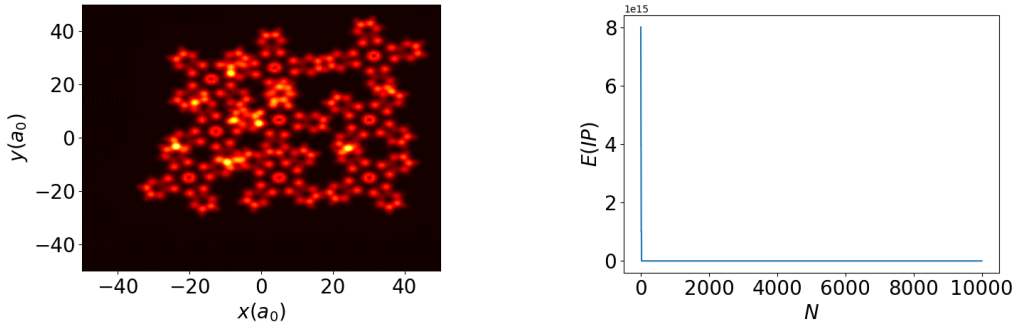
Figure 7.2: Image showing the pattern of available positions for moleculels in the superlattice calculation simulation.

This stage was successful in the context that the system did successfully minimize for the 5×5 set of molecules in which it was tested. The problem is that for these simulations parameters the molecules just maximise their distances away from each other (for periodic boundary conditions that just means being occupying the same surface point) and does not generate realistic super-lattices. This mode, however, was only designed to test the new functions in which the molecular centres were no longer fixed, actually generating an overlayer calculation requires more mobility in the molecules. By allowing them to move a limited

distance about a given region still requires defining a kind of overlay structure to begin with.

The real challenge these simulations were intended to test and solve was that as the molecules reposition themselves the distances and angles between them change meaning they have to be re-evaluated at each iteration. Given the number of iterations involved is of the order 10^6 it is imperative that these are made to be fast and efficient. This was the success of those simulations, that those functions worked and worked well. The problems arrived when trying to move onto the next stage in actually predicting the super-lattice structures, the idea that was being tested was to start with all the molecules bundled tightly together in the centre of the surface and then observe the outcomes when allowing them to spread out. There were two issues in these simulations one conceptual and one practical. Practically speaking it simply did not minimise, the simulation would easily become trapped in what were clearly unfavourable positions but the interactions would not reduce any further, as can be seen in figure 7.3.

The reasons for this behaviour were unclear but a potential solution was thought



(a) (b)

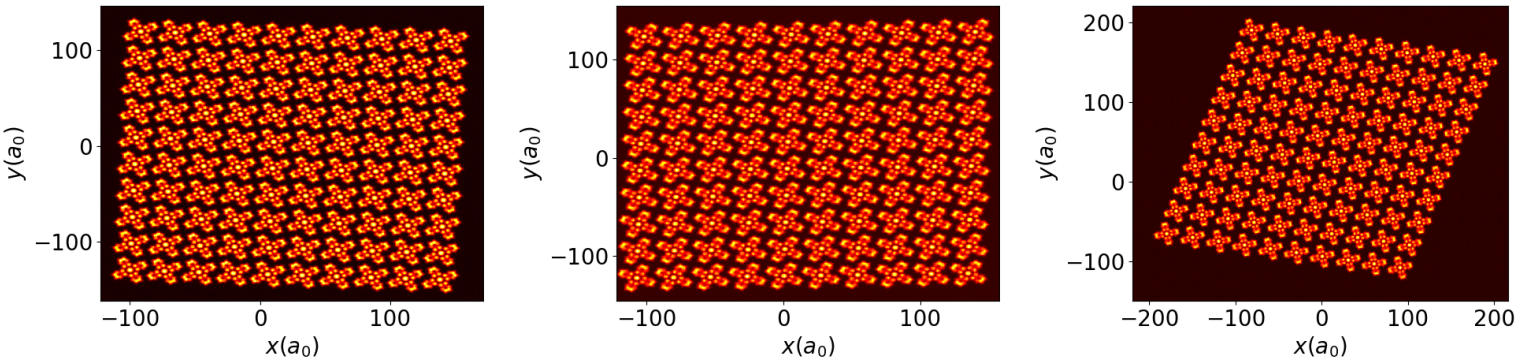
Figure 7.3: Representation of the minimization process, with (a) displaying the final position of nine TiOPc molecules after 5×10^4 iterations. Whist (b) depicts a graph demonstrating the difficulty with the minimisation process, in which the interaction potential would fail reduce any further after only fifty iterations.

to be available from changing the way they moved around the surface. The other conceptual issue is that this still had a slightly limited scope in that even bunched together in the centre they still started in a grid formation. Theoretically, you would expect that they would spread out to fill the surface, a better more appropriate technique would to attempt to mimic the process of evaporating them onto the surface as was performed experimentally in [7]. In this scenario, the surface starts empty and molecules fill the space and start to minimize. That last idea for modelling the phthalocyanine surfaces has a lot of complexity to it and the idea being tested at the time was a stepping stone to such increased complexity. It is believed that these ideas would have been well within the realm of possibility and a natural area in which the work would progress.

Appendix D

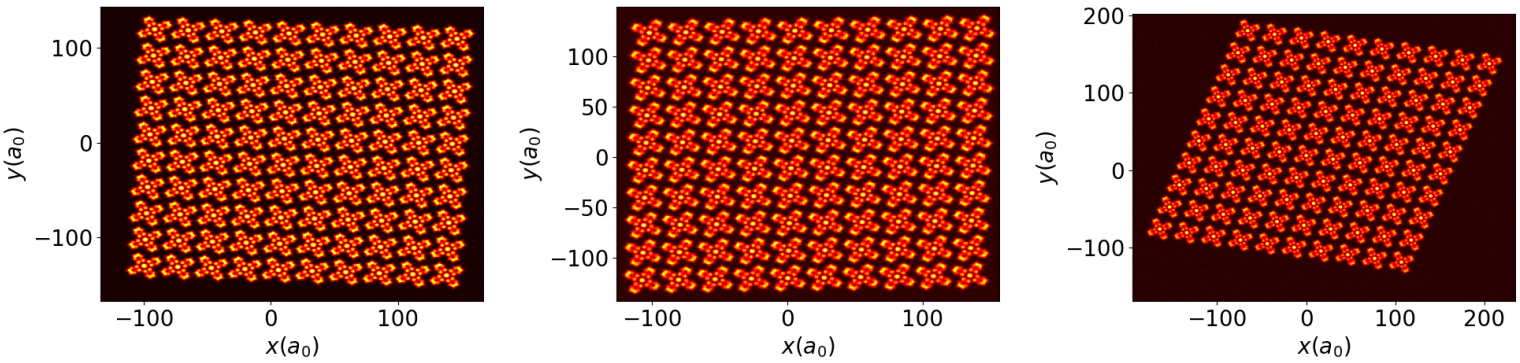
STM simulation images in the surface comparison set

In chapter six there were many simulations performed, generating a large number of STM images. For completeness, each of the forty-five simulations in the last section of that chapter are displayed here, as these will provide clarity for those results and present a comparison. These start on the following page and displayed as follows. From left to right each trio of images presents a monolayer in the super-lattice structure TiOPc1, TiOPc2 then FePc1 respectively. Each page contains three rows of images in which each row is a different surface, from top to bottom each row on a page displays simulations on a gold surface, on a silver surface, then on an NaCl surface. Finally, each page contains all of the surface comparision simulation images for a different molecule, in order each page has the simulations for TiOPc, SnPc, FePc, CuPc and CoPc respectively.



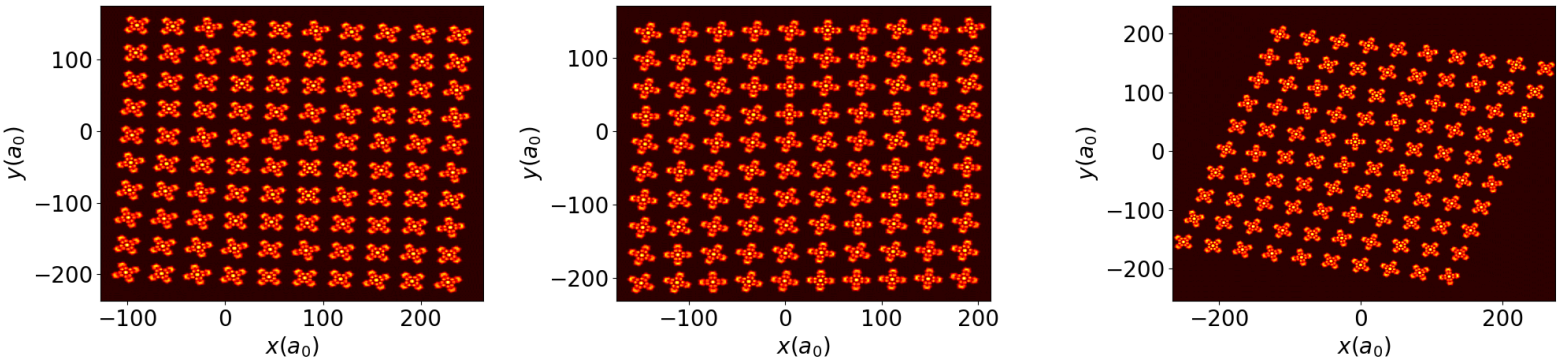
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.4: TiOPc simulations on a gold surface, in each superlattice structure



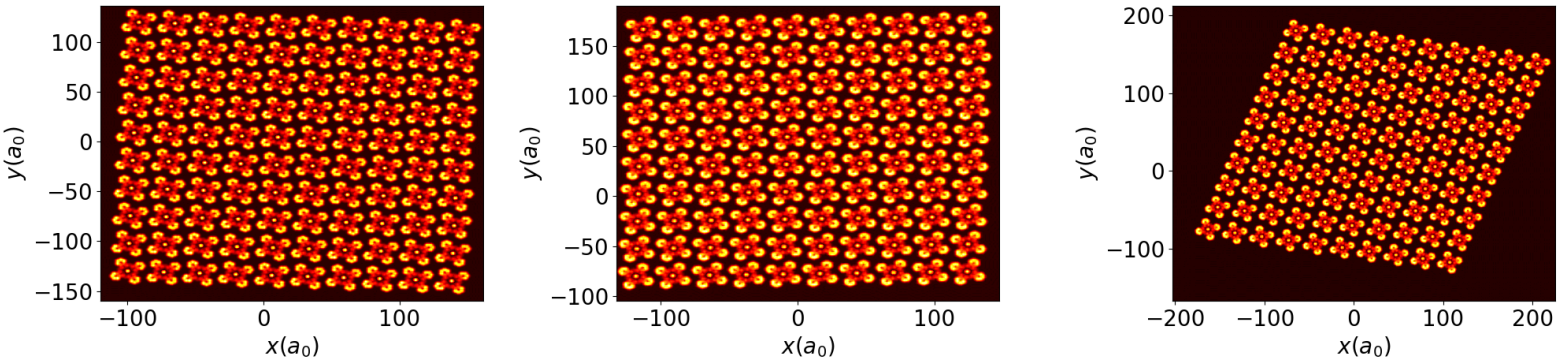
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.5: TiOPc simulations on a silver surface, in each superlattice structure



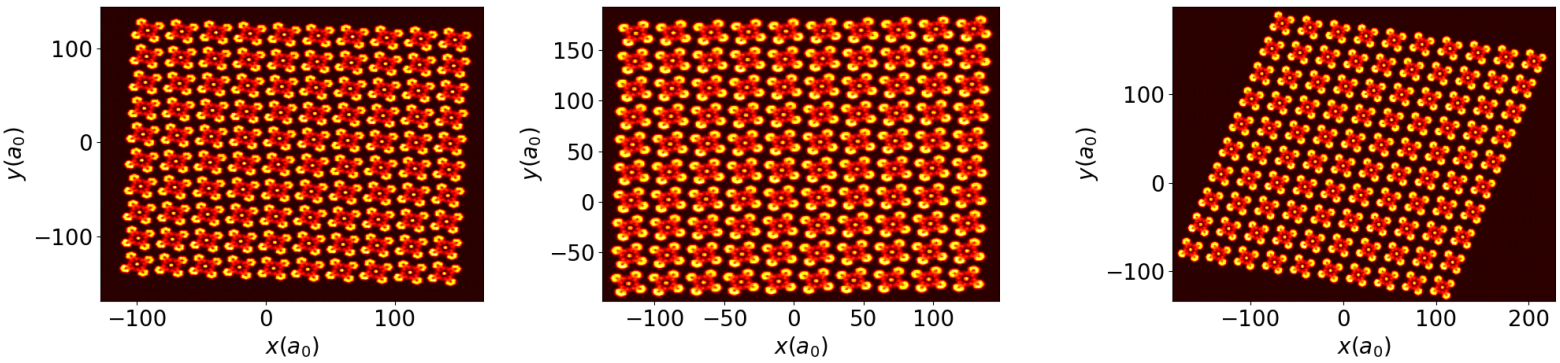
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.6: TiOPc simulations on a sodium chloride surface, in each superlattice structure



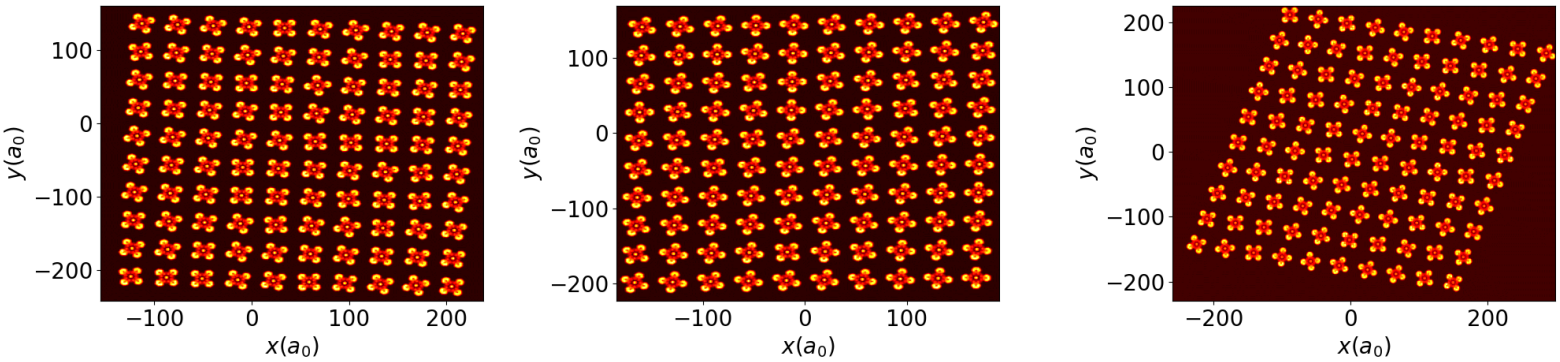
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.7: SnPc simulations on a gold surface, in each superlattice structure



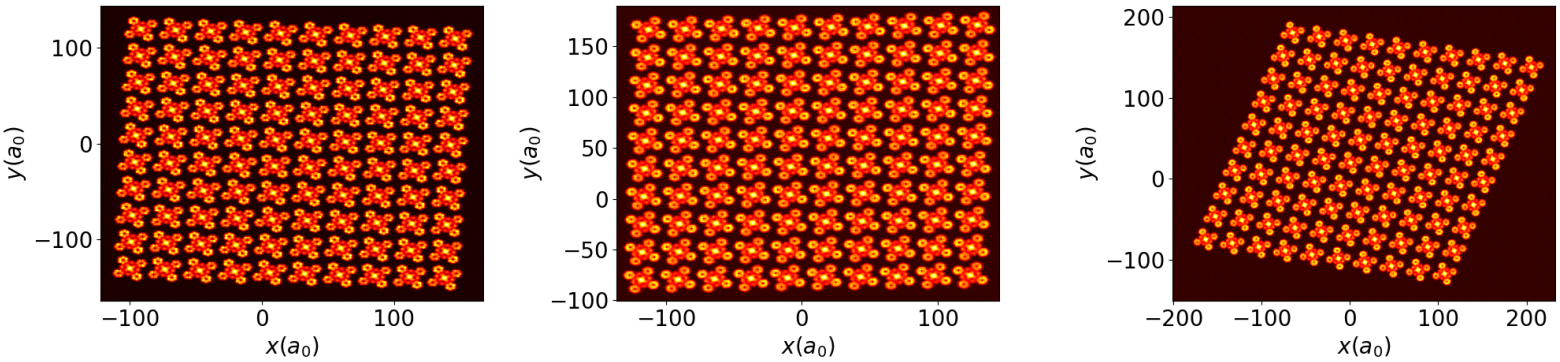
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.8: SnPc simulations on a silver surface, in each superlattice structure



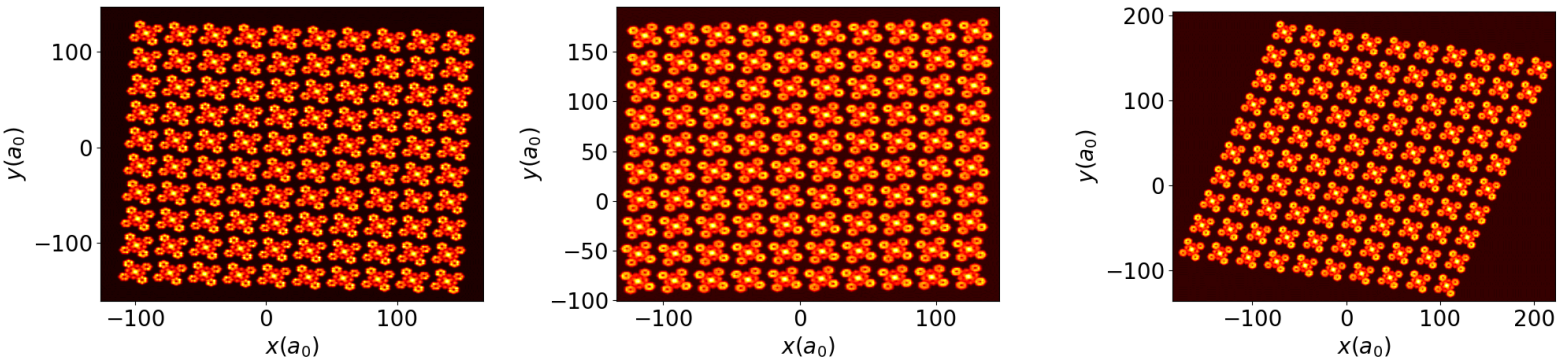
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.9: SnPc simulations on a sodium chloride surface, in each superlattice structure



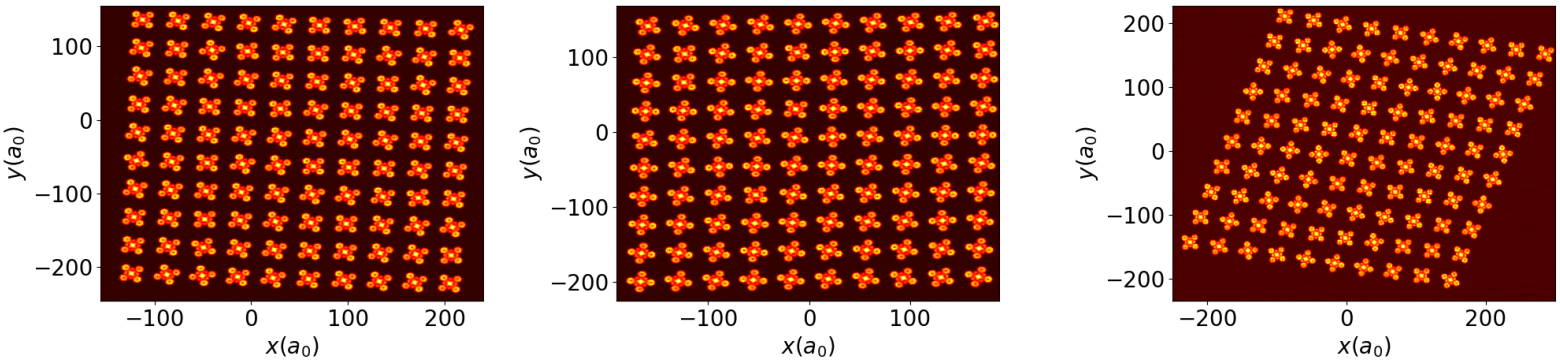
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.10: FePc simulations on a gold surface, in each superlattice structure



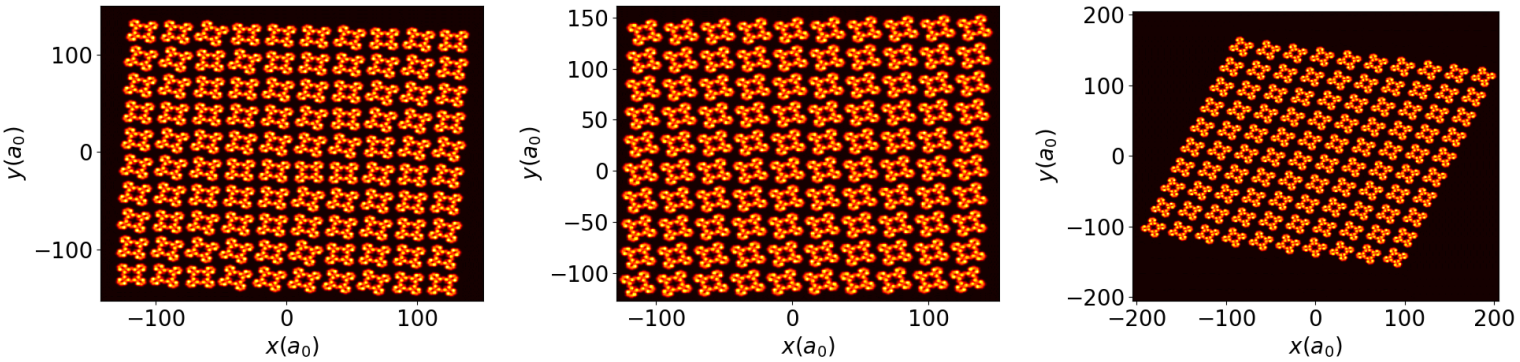
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.11: FePc simulations on a silver surface, in each superlattice structure



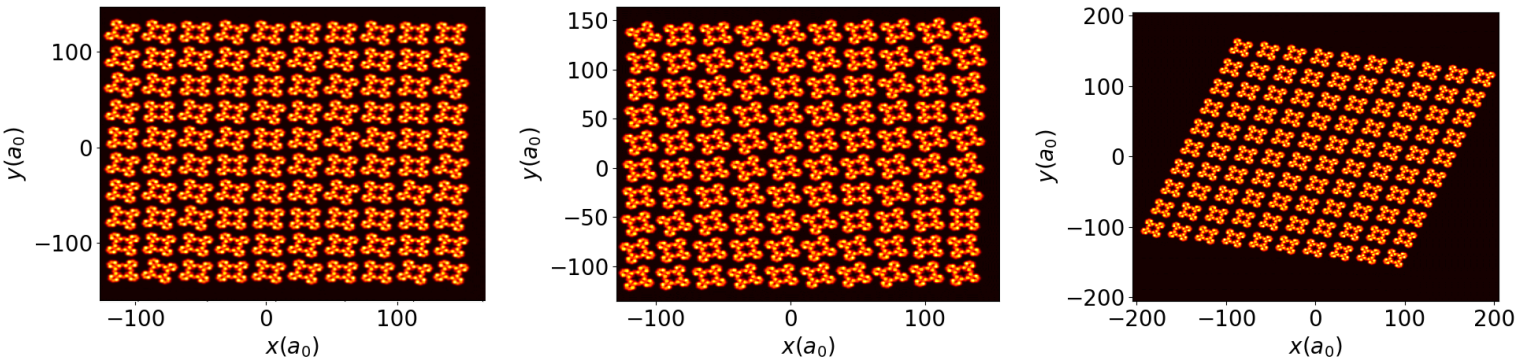
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.12: FePc simulations on a sodium chloride surface, in each superlattice structure



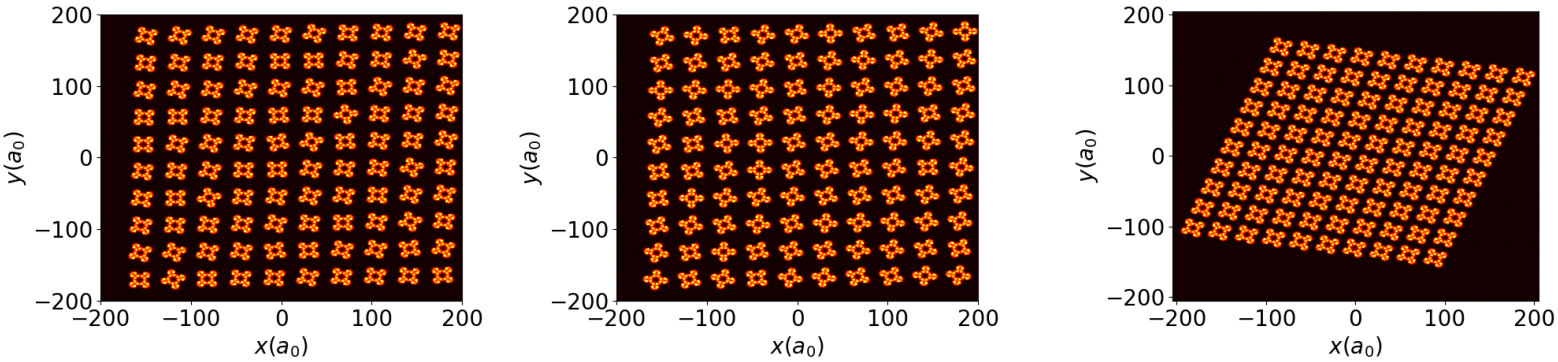
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.13: CuPc simulations on a gold surface, in each superlattice structure



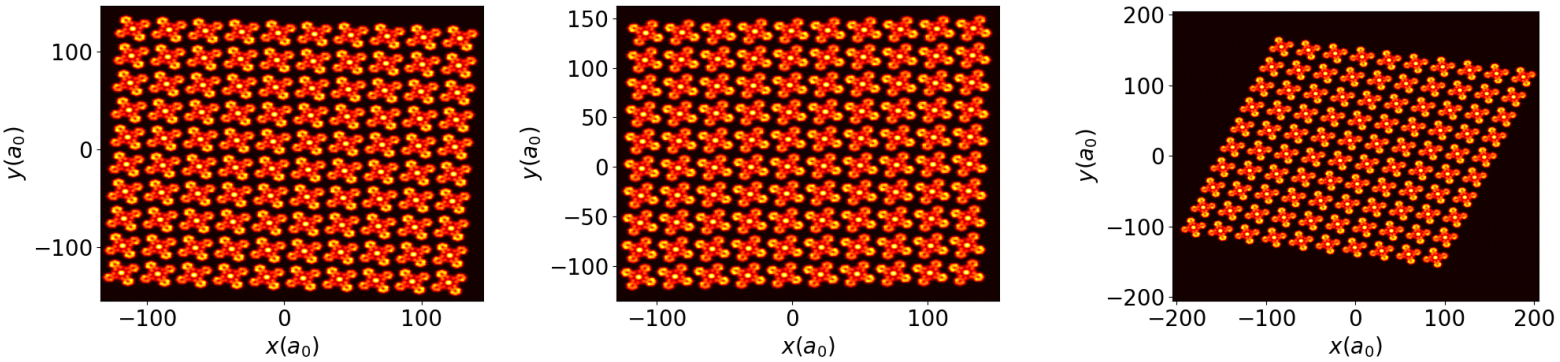
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.14: CuPc simulations on a silver surface, in each superlattice structure



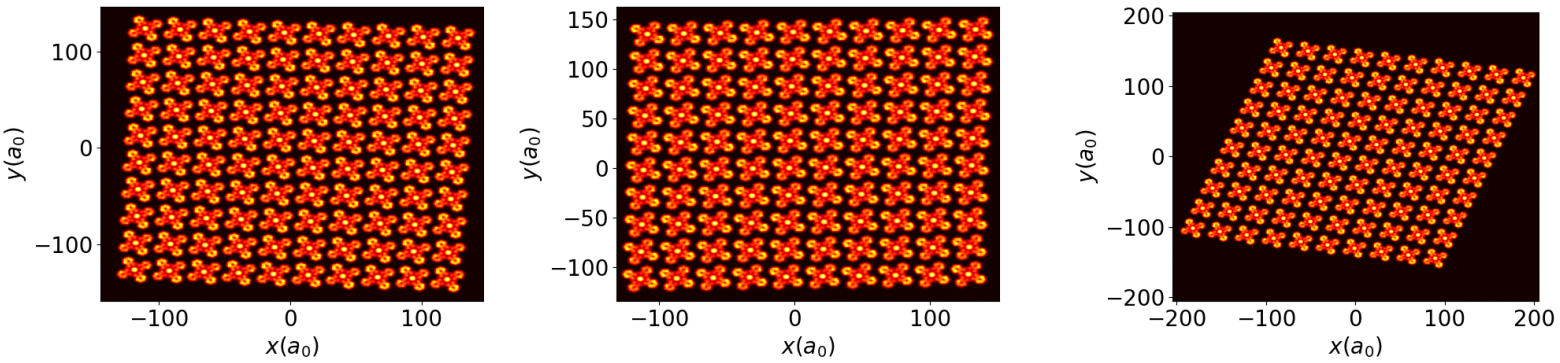
(a) (b) (c)

Figure 7.15: CuPc simulations on a sodium chloride surface, in each superlattice structure



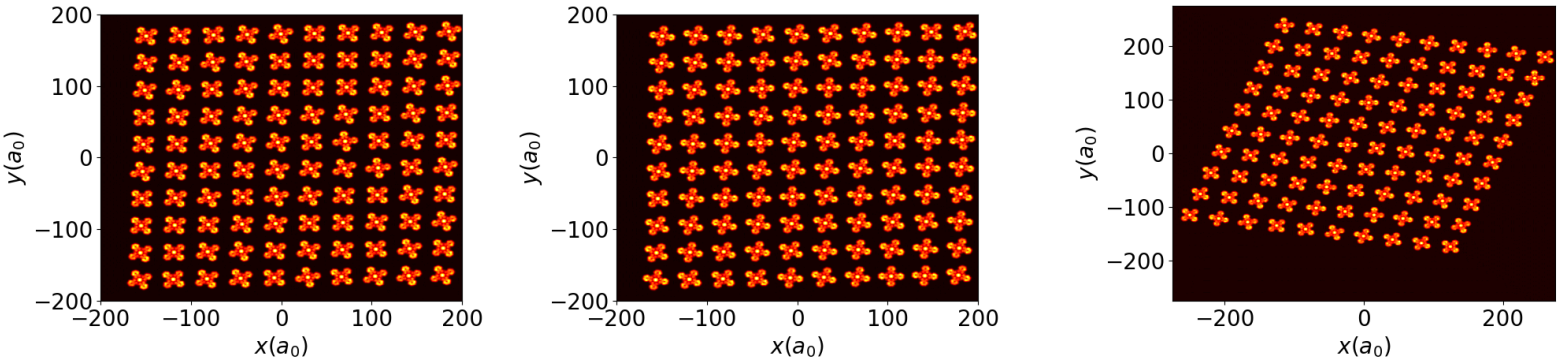
(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.16: CoPc simulations on a gold surface, in each superlattice structure



(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.17: CoPc simulations on a silver surface, in each superlattice structure



(a) TiOPc1 (b) TiOPc2 (c) FePc1

Figure 7.18: CoPc simulations on a sodium chloride surface, in each superlattice structure

Bibliography

[1] J.B. Adams. Bonding Energy Models. *Encyclopedia of Materials: Science and Technology*, pages 763–767, 2001.

[2] James Foresman and Aeleen Frisch. *Exploring Chemistry With Electronic Structure Methods, 3rd edition*. Guassian inc, 2007.

[3] F Jensen. *Introduction to Computational Chemistry*. john wiley and sons ltd, 2007.

[4] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry*. Macmillan publishing co, 1982.

[5] J Leaf, A Stannard, S P Jarvis, P Moriarty, and J L and Dunn. A Combined Monte Carlo and Hückel Theory Simulation of Orientational Ordering in C 60 Assemblies. *The Journal of Physical Chemistry C*, 120(15):8139–8147, 2016.

[6] S C B Mannsfeld and T Fritz. Understanding organic–inorganic heteroepitaxial growth of molecules on crystalline substrates: Experiment and theory. *Physical Review B*, 71(23):235405, 2005.

[7] S C Bobaru, E Salomon, J.-M. Layet, and T Angot. Structural Properties of Iron Phtalocyanines on Ag(111): From the Submonolayer to Monolayer Range. *The Journal of Physical Chemistry C*, 115(13):5875–5879, 2011.

[8] K Yates. *Huckel Molecular Orbial theory*. acedemic press inc, 1978.

[9] James JP Stewart. Semiempirical molecular orbital methods. *Reviews in computational chemistry*, 1:45–81, 1990.

[10] Robert G Parr. *Density Functional Theory of Atoms and Molecules*. Springer Netherlands, 1980.

[11] P H Lippel, R J Wilson, M D Miller, Ch. Wöll, and S Chiang. High-Resolution Imaging of Copper-Phthalocyanine by Scanning-Tunneling Microscopy. *Physical Review Letters*, 62(2):171–174, 1989.

[12] C Ludwig. Epitaxy and scanning tunneling microscopy image contrast of copper–phthalocyanine on graphite and MoS2. *Journal of Vacuum Science and Technology B*, 12(3):1963, 1994.

[13] Xing Lu, K. W. Hipps, X. D. Wang, and Ursula Mazur. Scanning tunneling microscopy of metal phthalocyanines: d7 and d9 cases. *Journal of the American Chemical Society*, 118(30):7197–7202, 1996.

[14] J D Baran, J A Larsson, R A J Woolley, Yan Cong, P J Moriarty, A A Cafolla, K Schulte, and V R Dhanak. Theoretical and experimental comparison of SnPc, PbPc, and CoPc adsorption on Ag(111). *Physical Review B*, 81(7):075413, 2010.

[15] Zhihai Cheng, Shixuan Du, Wei Guo, Li Gao, Zhitao Deng, Nan Jiang, Haiming Guo, Hao Tang, and H J. Gao. Direct imaging of molecular orbitals of metal phthalocyanines on metal surfaces with an O2-functionalized tip of a scanning tunneling microscope. 4(6):523–530, 2011.

[16] G Binnig, H Rohrer, Ch. Gerber, and E Weibel. Surface Studies by Scanning Tunneling Microscopy. *Physical Review Letters*, 49(1):57–61, 1982.

[17] Chunli Bai. *Scanning tunneling microscopy and its application*, volume 32. Springer Science & Business Media, 2000.

[18] Ran Marom, Chen Levi, Tal Weiss, Salman Rosenwaks, Yehuda Zeiri, Ronnie Kosloff, and Ilana Bar. Quantum Tunneling of Hydrogen Atom in Dissociation of Photoexcited Methylamine †. *The Journal of Physical Chemistry A*, 114(36):9623–9627, 2010.

[19] C A Bertulani. Tunneling of Atoms, Nuclei and Molecules. *Few Body Systems*, 56(11-12):727–736, 2015.

[20] Eugen Merzbacher. The Early History of Quantum Tunneling. *Physics Today*, 55(8):44–49, 2002.

[21] J. J. Sakurai and Eugene D. Commins. Modern quantum mechanics, revised edition. *American Journal of Physics*, 63(1):93–95, 1995.

[22] Frank Trixler. Quantum Tunnelling to the Origin and Evolution of Life. *Current Organic Chemistry*, 17(16):1758–1770, 2013.

[23] Yufeng Chen, Wei Xu, and Jinlin Huang. A simple new technique for preparing STM tips. *Journal of Physics E: Scientific Instruments*, 22(7):455–457, 1989.

[24] L A Nagahara, T Thundat, and S M and Lindsay. Preparation and characterization of STM tips for electrochemical studies. *Review of Scientific Instruments*, 60(10):3128–3130, 1989.

[25] Th. Laloyaux, A. A. Lucas, J.-P. Vigneron, Ph. Lambin, and H. Morawitz. Lateral resolution of the scanning tunnelling microscope. *Journal of Microscopy*, 152(1):53–63, 1988.

[26] Lewys Jones, Shuqiu Wang, Xiao Hu, Shams ur Rahman, and Martin R Castell. Maximising the resolving power of the scanning tunneling microscope. *Advanced Structural and Chemical Imaging*, 4(1):7, 2018.

[27] G. Binnig and D. P. E. Smith. Single-tube three-dimensional scanner for scanning tunneling microscopy. *Review of Scientific Instruments*, 57(8):1688–1689, 1986.

[28] Chris Toumey. 35 Atoms That Changed the Nanoworld. *Nature Nanotechnology*, 5(4):239–241, 2010.

[29] H J Lee and W Ho. Single-Bond Formation and Characterization with a Scanning Tunneling Microscope. *Science*, 286(5445):1719 LP – 1722, 1999.

[30] A. Kubetzka, M. Bode, O. Pietzsch, and R. Wiesendanger. Spin-polarized scanning tunneling microscopy with antiferromagnetic probe tips. *Physical Review Letters*, 88:057201, 2002.

[31] Nuclear quantum effects of hydrogen bonds probed by tip-enhanced inelastic electron tunneling. *Science*, 352(6283):321–325, 2016.

[32] Electronic properties of transition metal phthalocyanines: The impact of the central metal atom (d5-d10). *Organic Electronics*, 11(9):1483–1488, 2010.

[33] Neil B McKeown et al. *Phthalocyanine materials: synthesis, structure and function*. Number 6. Cambridge University Press, 1998.

[34] Peter Gregory. Industrial applications of phthalocyanines. *Journal of Porphyrins and Phthalocyanines*, 04(04):432–437, 2000.

[35] Marcel Bouvet. Phthalocyanine-based field-effect transistors as gas sensors. *Analytical and Bioanalytical Chemistry*, 384(2):366–373, 2006.

[36] Harshil N Raval, D S Sutar, and V and Ramgopal Rao. Copper(II) phthalocyanine based organic electronic devices for ionizing radiation dosimetry applications. *Organic Electronics*, 14(5):1281–1290, 2013.

[37] C Barraud, K Bouzehouane, C Deranlot, D J Kim, R Rakshit, S Shi, J Arabski, M Bowen, E Beaurepaire, S Boukari, F Petroff, P Seneor, and R and Mattana. Phthalocyanine based molecular spintronic devices. *Dalton Transactions*, 45(42):16694–16699, 2016.

[38] A Rosa and E J Baerends. Metal-Macrocycle Interaction in Phthalocyanines: Density Functional Calculations of Ground and Excited States. *Inorganic Chemistry*, 33(3):584–595, 1994.

[39] Meng-Sheng Liao, John D. Watts, and Ming-Ju Huang. Dft study of unligated and ligated manganeseii porphyrins and phthalocyanines. *Inorganic Chemistry*, 44(6):1941–1949, 2005. PMID: 15762720.

[40] Philip A Reynolds and Brian N Figgis. Metal phthalocyanine ground states: covalence and ab initio calculation of spin and charge densities. *Inorganic Chemistry*, 30(10):2294–2300, 1991.

[41] Arnold M Schaffer, Martin Gouterman, and Ernest R Davidson. Porphyrins XXVIII. Extended Hückel calculations on metal phthalocyanines and tetrazaporphins. *Theoretica chimica acta*, 30(1):9–30, 1973.

[42] Michael G. Walter, Alexander B. Rudine, and Carl C. Wamser. Porphyrins and phthalocyanines in solar photovoltaic cells. *Journal of Porphyrins and Phthalocyanines*, 14(9):759–792, 2010.

[43] Yuexing Zhang, Xue Cai, Yongzhong Bian, and Jianzhuang Jiang. Organic Semiconductors of Phthalocyanine Compounds for Field Effect Transistors (FETs). In *Structure and Bonding*, volume 135, pages 275–321. 2010.

[44] Werner Kutzelnigg. What i like about Hückel theory. *Journal of Computational Chemistry*, 28(1):25–34, 2007.

[45] Eldon G Emberly and George and Kirczenow. Molecular spintronics: spin-dependent electron transport in molecular wires. *Chemical Physics*, 281(2-3):311–324, 2002.

[46] J. Bardeen. Tunnelling from a many-particle point of view. *Physical Review Letters*, 6(2):57–59, 1961.

[47] J Tersoff and D R Hamann. Theory of the scanning tunneling microscope. *Physical Review B*, 31(2):805–813, 1985.

[48] J Chen. *Introduction to Scanning Tunneling Microscopy*. Oxford University Press, 1993.

[49] U. Eichmann, M. Dörr, H. Maeda, W. Becker, and W. Sandner. Collective multielectron tunneling ionization in strong fields. *Physical Review Letters*, 84:3550–3553, 2000.

[50] Alex D. Gottlieb and Lisa Wesoloski. Bardeen's tunnelling theory as applied to scanning tunnelling microscopy: A technical guide to the traditional interpretation. *Nanotechnology*, 17(8):R57–R65, 2006.

[51] J. M. Zhang and Y. Liu. Fermi's golden rule: Its derivation and breakdown by an ideal model. *European Journal of Physics*, 37(6):65406, 2016.

[52] Werner A. Hofer, Adam S. Foster, and Alexander L. Shluger. Theories of scanning probe microscopes at the atomic scale. *Reviews of Modern Physics*, 75(4):1287–1331, 2003.

[53] Axel D Becke. A new mixing of Hartree–Fock and local density-functional theories. *The Journal of Chemical Physics*, 98(2):1372–1377, 1993.

[54] D. R. Hartree. The Wave Mechanics of an Atom with a Non-Coulomb Central Field Part II Some Results and Discussion. *Mathematical Proceedings of the Cambridge Philosophical Society*, 24(1):111–132, 1928.

[55] S Bludman and P B Daitch. Validity of the Born-Oppenheimer Approximation. *Physical Review*, 95(3):823–830, 1954.

[56] Jan Linderberg. Perturbation treatment of Hartree-Fock equations. *Physical Review*, 121(3):816–819, 1961.

[57] Solution of the Hartree–Fock equations by a pseudospectral method: Application to diatomic molecules. *The Journal of Chemical Physics*, 85(3):1462–1468, 1986.

[58] C. David Sherrill and Henry F. Schaefer. The Configuration Interaction Method: Advances in Highly Correlated Approaches. In Per-Olov Löwdin, John R Sabin, Michael C Zerner, and Erkki B T Advances in Quantum Chemistry Brändas, editors, *Advances in Quantum Chemistry*, volume 34, pages 143–269. Academic Press, 1999.

[59] Rodney J. Bartlett and Monika Musiał. Coupled-cluster theory in quantum chemistry. *Reviews of Modern Physics*, 79(1):291–352, 2007.

[60] Chr. Møller and M S Plesset. Note on an Approximation Treatment for Many-Electron Systems. *Physical Review*, 46(7):618–622, 1934.

[61] P Hohenberg and W Kohn. Inhomogeneous Electron Gas. *Physical Review*, 136(3B):B864–B871, 1964.

[62] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Physical Review Letters*, 100(14):146401, 2008.

[63] Fenggong Wang, Cristiana Di Valentin, and Gianfranco Pacchioni. Electronic and structural properties of WO3: A systematic hybrid DFT study. *Journal of Physical Chemistry C*, 115(16):8345–8353, 2011.

[64] Stefan Grimme. Semiempirical hybrid density functional with perturbative second-order correlation. *Journal of Chemical Physics*, 124(3):34108, 2006.

[65] Haoyu S. Yu, Xiao He, and Donald G. Truhlar. Mn15-l: A new local exchange-correlation functional for kohn–sham density functional theory with broad accuracy for atoms, molecules, and solids. *Journal of Chemical Theory and Computation*, 12(3):1280–1293, 2016.

[66] Michael J.G. Peach, Trygve Helgaker, Paweł Sałek, Thomas W. Keal, Ola B. Lutnæs, David J. Tozer, and Nicholas C. Handy. Assessment of a Coulomb-attenuated exchange-correlation energy functional. *Physical Chemistry Chemical Physics*, 8(5):558–562, 2006.

[67] W Kohn and L J Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Physical Review*, 140(4A):A1133–A1138, 1965.

[68] Stephan Mohr, Laura E. Ratcliff, Luigi Genovese, Damien Caliste, Paul Boulanger, Stefan Goedecker, and Thierry Deutsch. Accurate and efficient linear scaling DFT calculations with universal applicability. *Physical Chemistry Chemical Physics*, 17(47):31360–31370, 2015.

[69] L J Schaad and B A Hess. Huckel theory and aromatically. *Journal of Chemical Education*, 51(10):640–643, 1974.

[70] D. Kienle, J. I. Cerda, and A. W. Ghosh. Extended Hückel theory for band structure, chemistry, and transport. I. Carbon nanotubes. *Journal of Applied Physics*, 100(4):43714, 2006.

[71] F. Zahid, M. Paulsson, E. Polizzi, A. W. Ghosh, L. Siddiqui, and S. Datta. A self-consistent transport model for molecular conduction based on extended Hückel theory with full three-dimensional electrostatics. *Journal of Chemical Physics*, 123(6):64707, 2005.

[72] Hassan Raza and Edwin C. Kan. An extended Hückel theory based atomistic model for graphene nanoelectronics. *Journal of Computational Electronics*, 7(3):372–375, 2008.

[73] Roald Hoffmann. An Extended Hückel Theory. I. Hydrocarbons. *The Journal of Chemical Physics*, 39(6):1397–1412, 1963.

[74] Barry L Kalman. Self-consistent extended Hückel theory. I. *The Journal of Chemical Physics*, 59(9):5184–5188, 1973.

[75] Alexey V. Akimov and Oleg V. Prezhdo. Analysis of self-consistent extended Hückel theory (SC-EHT): a new look at the old method. *Journal of Mathematical Chemistry*, 53(2):528–550, 2014.

[76] D. N. Futaba, J. P. Landry, A. Loui, and S. Chiang. Experimental and theoretical STM imaging of xylene isomers on Pd(111). *Physical Review B - Condensed Matter and Materials Physics*, 65(4):1–7, 2002.

[77] Nt-STM: A step forward in Scanning Tunneling Microscopy (STM) simulations. *Computer Physics Communications*, 183(6):1246–1249, 2012.

[78] Soumen Saha, Ram Kinkar Roy, and Paul W. Ayers. Are the Hirshfeld and Mulliken population analysis schemes consistent with chemical intuition? *International Journal of Quantum Chemistry*, 109(9):1790–1806, 2009.

[79] Wilfried J Mortier, Swapan K Ghosh, and S Shankar. Electronegativity-equalization method for the calculation of atomic charges in molecules. *Journal of the American Chemical Society*, 108(15):4315–4320, 1986.

[80] Boris K Novosadov. Hydrogen-like atomic orbitals: Addition and expansion theorems, integrals. *International Journal of Quantum Chemistry*, 24(1):1–18, 1983.

[81] Martin Kaupp. The role of radial nodes of atomic orbitals for chemical bonding and the periodic table. *Journal of Computational Chemistry*, 28(1):320–325, 2007.

[82] Israfil I. Guseinov and Bahtiyar A. Mamedov. On the accurate evaluation of overlap integrals over slater type orbitals using analytical and recurrence relations. *Zeitschrift fur Naturforschung - Section A Journal of Physical Sciences*, 62(9):467–470, 2007.

[83] Michał Lesiuk and Robert Moszynski. Reexamination of the calculation of two-center, two-electron integrals over Slater-type orbitals. I. Coulomb and hybrid integrals. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 90(6), 2014.

[84] S Huzinaga. Gaussian-type functions for polyatomic systems. i. *The Journal of Chemical Physics*, 42(4):1293–1302, 1965.

[85] Hiroshi Taketa, Sigeru Huzinaga, and Kiyosi O-ohata. Gaussian-Expansion Methods for Molecular Integrals. *Journal of the Physical Society of Japan*, 21(11):2313–2324, 1966.

[86] Emili Besalú and Ramon Carbó-Dorca. The general Gaussian product theorem. *Journal of Mathematical Chemistry*, 49(8):1769–1784, 2011.

[87] J M H-P Minhhuy Hô. Evaluation of gaussian molecular integrals. *The Mathematica Journal*, 16:14–3, 2012.

[88] S Kirkpatrick, C D Gelatt, and M P and Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, 1983.

[89] John Geweke. Chapter 15 monte carlo simulation and numerical integration. volume 1 of *Handbook of Computational Economics*, pages 731 – 800. Elsevier, 1996.

[90] A. Hacura, M. Jadamus-Hacura, and A. Kocot. Risk analysis in investment appraisal based on the Monte Carlo simulation technique. *European Physical Journal B*, 20(4):551–553, 2001.

[91] Kerson Huang. *Introduction to statistical physics*. Chapman and Hall/CRC, 2009.

[92] Nenad Trinajstić, Milan Randić, Douglas J Klein, Darko Babić, and Zlatko Mihalić. On mathematical properties of buckminsterfullerenef. *Croatica Chemica Acta*, 68(1):241–267, 1995.

[93] Benjamin P. Pritchard, Doaa Altarawy, and Theresa L. Windus. A new basis set exchange: An open, up-to-date resource for the molecular sciences community. Manuscript in preparation.

[94] David Feller. The role of databases in support of computational chemistry calculations. *J. Comput. Chem.*, 17, 1996.

[95] Karen L. Schuchardt, Brett T. Didier, Todd Elsethagen, Lisong Sun, Vidhya Gurumoorthi, Jared Chase, Jun Li, and Theresa L. Windus. Basis set exchange: A community database for computational sciences. *J. Chem. Inf. Model.*, 47, 2007.

[96] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-consistent molecular-orbital methods. i. use of gaussian expansions of slater-type atomic orbitals. *J. Chem. Phys.*, 51, 1969.

[97] H. A. Skinner and H. O. Pritchard. The measure of electronegativity. *Trans. Faraday Soc.*, 49:1254–1262, 1953.

[98] Y. J. Dappe, M. Andersen, R. Balog, L. Hornekær, and X. Bouju. Adsorption and STM imaging of polycyclic aromatic hydrocarbons on graphene. *Physical Review B - Condensed Matter and Materials Physics*, 91(4), 2015.

[99] H A Yoon, M Salmeron, and G A Somorjai. Scanning tunneling microscopy (STM) study of benzene and its coadsorption with carbon monoxide on Rh(111). *Surface Science*, 373(2-3):300–306, 1997.

[100] Li-Jun Wan and Kingo Itaya. In SituScanning Tunneling Microscopy of Benzene, Naphthalene, and Anthracene Adsorbed on Cu(111) in Solution. *Langmuir*, 13(26):7173–7179, 1997.

[101] Marcus D. Hanwell, Donald E. Curtis, David C. Lonie, Tim Vandermeerschd, Eva Zurek, and Geoffrey R. Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *Journal of Cheminformatics*, 4(8):17, 2012.

[102] Thomas A Halgren. Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. *Journal of Computational Chemistry*, 17(5-6):490–519, apr 1996.

[103] H.B. Gray. *Electrons and Chemical Bonding*. W. A. Benjamin, 1965.

[104] T Kroll, V Yu. Aristov, and Molodtsova.

[105] Gopalan Rajaraman, Andrea Caneschi, Dante Gatteschi, and Federico Totti. A periodic mixed gaussians-plane waves DFT study on simple thiols on Au(111): Adsorbate species, surface reconstruction, and thiols functionalization. *Physical Chemistry Chemical Physics*, 13(9):3886–3895, 2011.

[106] Binghai Yan, Benjamin Stadtmüller, Norman Haag, Sebastian Jakobs, Johannes Seidel, Dominik Jungkenn, Stefan Mathias, Mirko Cinchetti, Martin Aeschlimann, and Claudia Felser. Topological states on the gold surface. *Nature Communications*, 6(1):10167, 2015.

[107] W. Chen, V. Madhavan, T. Jamneala, and M. F. Crommie. Scanning tunneling microscopy observation of an electronic superlattice at the surface of clean gold. *Physical Review Letters*, 80(7):1469–1472, 1998.

[108] J. Grant Hill. Gaussian basis sets for molecular applications. *International Journal of Quantum Chemistry*, 113(1):21–34, 2013.

[109] Pashaev Geydar, Arzuman Gasanov, and Ali Mahmood. The Study of Gold Nanoparticles in basis of Slater Functions. *Journal Nanotechnology and Advanced Materials*, 2:35–41, 2014.

[110] Wheeler P Davey. Precision Measurements of the Lattice Constants of Twelve Common Metals. *Physical Review*, 25(6):753–761, 1925.

[111] Joseph E. Norton and Jean Luc Bŕdas. Theoretical characterization of titanyl phthalocyanine as a p -type organic semiconductor: Short intermolecular π -π ild large electronic couplings and hole transport bandwidths. *Journal of Chemical Physics*, 128(3):34701, 2008.

[112] P. H. Lippel, R. J. Wilson, M. D. Miller, Ch. Wöll, and S. Chiang. High-resolution imaging of copper-phthalocyanine by scanning-tunneling microscopy. *Phys. Rev. Lett.*, 62:171–174, Jan 1989.

[113] Marius Toader and Michael and Hietschold. SnPc on Ag(111): A Scanning Tunneling Microscopy Study at the Submolecular Level. 115(25):12494–12500, 2011.

[114] Tobias Schmitt, Pascal Ferstl, Lutz Hammer, M. Alexander Schneider, and Josef Redinger. Adsorption and Intermolecular Interaction of Cobalt Phthalocyanine on CoO(111) Ultrathin Films: An STM and DFT Study. *Journal of Physical Chemistry C*, 121(5):2889–2895, 2017.

[115] M. Scheffler, L. Smykalla, D. Baumann, R. Schlegel, T. Hänke, M. Toader, B. Büchner, M. Hietschold, and C. Hess. Structural study of monolayer cobalt phthalocyanine adsorbed on graphite. *Surface Science*, 608:55–60, 2013.

[116] J E Nickels, M A Fineman, and W E Wallace. X-Ray Diffraction Studies of Sodium Chloride–Sodium Bromide Solid Solutions. *The Journal of Physical and Colloid Chemistry*, 53(5):625–628, 1949.

[117] H. Fukagawa, H. Yamane, S. Kera, K. K. Okudaira, and N. Ueno. Experimental estimation of the electric dipole moment and polarizability of titanyl phthalocyanine using ultraviolet photoelectron spectroscopy. *Physical Review B - Condensed Matter and Materials Physics*, 73(4), 2006.

[118] Marc Heinrich Prosenc, Christoph Janiak, and Hans Herbert Brintzinger. Agostic assistance to olefin insertion in alkylzirconocene cations: a molecular orbital study by the extended hueckel method. *Organometallics*, 11(12):4036–4041, 1992.

[119] Mattias Hjort and Sven Stafström. Modeling vacancies in graphite via the hückel method. *Phys. Rev. B*, 61:14089–14094, 2000.

[120] Jittima Meeprasert, Anchalee Junkaew, Nawee Kungwan, Bavornpon Jansang, and Supawadee Namuangruk. A Cr-phthalocyanine monolayer as a potential catalyst for NO reduction investigated by DFT calculations. *RSC Advances*, 6(25):20500–20506, 2016.