Methods for the Treatment of Multiple States in Non-Adiabatic Direct Quantum Dynamics Simulations

by

Kaite Eryn Spinlove

A thesis submitted to

the University of Birmingham

for the examination of

DOCTOR OF PHILOSOPHY

School of Chemistry
College of Engineering and Physical Sciences
University of Birmingham
2017

Abstract

Following the pioneering work of A. Zewail the field of Femtochemistry was opened up to experimental studies. However, the interpretation of the results of these studies is often difficult and hence theoretical and computational methods for the modelling of these processes have been developed. In these dynamical systems highly quantum behaviours are exhibited and hence result in the break down of the Born-Oppenheimer approximation. While classical and semi-classical simulations can be used as approximations to the full quantum behaviours, the Multi-Configurational Time-Dependent Hartree (MCTDH) method was developed as a grid-based full solution to the Time-Dependent Schrödinger equation. A major bottleneck in the use of grid-based methods is in the requirement for pre-computed potential energy surfaces upon which the dynamics simulations can be run. As an alternative formulation of the MCTDH ansatz equation, the variational Multi-Configurational Gaussian (vMCG) method was developed whereby the basis functions of MCTDH are replaced with Gaussian functions. A recent extension of the vMCG method, the Direct-Dynamics-vMCG (DD-vMCG) method, has been developed in which the Gaussian basis-functions, comprising the total wavepacket, are propagated on surfaces that are calculated "on-the-fly," or as required. In the DD-vMCG method the relevant energies, gradients and Hessians are calculated utilising an external quantum chemistry software package, the results of which are stored in a database.

In this study two highly relevant systems have been used, formamide and 2-pyridone, to test the limitations of the DD-vMCG method combining (relatively) high-level quantum chemistry calculations with a large number of excited states, and a challenging number of degrees of freedom . In addition, these systems have not yet been studied using full quantum dynamics simulations.

A more recent development in experimental photochemistry, attosecond laser spectroscopy, utilises ultra-fast laser pulses in order to be able to study the electron dynamics of systems. However, as the laser pulses are short, the band width of the pulses is broad and hence complementary computational studies require a large number of excited states to be represented. Outlined in this thesis is the initial testing, development and implementation of a new Ehrenfest method, whereby an averaged potential surface of a manifold of excited states is calculated, with test dynamics run on the highly conjugated molecule allene. This Ehrenfest-Multi-Configurational Gaussian (eMCG) method presents a potentially important addition to the suite of computational methods with a focus on the reproduction and interpretation of attosecond experimental results.

"Everything is theoretically impossible, until it is done."

Robert A. Heinlein

"Even if we could turn back, we'd probably never end up where we started."

Haruki Murakami

Acknowledgements

The entire process of completing my PhD, as I am sure it is for everyone, has been wrought with the highs and lows of everyday life, the highs and lows of scientific research, the highs and lows of learning in general, all magnified by my desire to not mess it all up! It hasn't been an easy journey, and there is absolutely no way I would have been able to get through it all without the help, support, understanding and most of all the patience of the people around me.

Consequently I have A LOT of thanks to give, so in no particular order:

I would like to thank all of the Worth group members, past and present. I am fairly certain I can remember everyone so, mostly in chronological order, and probably with terrible spelling here goes...

Thank you to all of the Masters students: Sam, Alice, Sam, Andrew, Nick, Tom, Divya, Matt and Semyon.

Thank you very much to my fellow PhD-ers: Marcus and Ceridwen.

Thank you very much to all of our amazing Post-Docs: Gareth, Simon, Chris, Iakov, Daniel, Marianna and Angelo.

Thank you to the visiting students: Daniel, Martin, Johannes, Xavvi and Sandra.

I sincerely hope I haven't missed anyone, though knowing me I have probably missed someone important...

I would like to extend a big thank you to my supervisor Mike Robb, and the team at Imperial (or at least some *were* at Imperial) - Mike, Andrew and Morgane.

I would also like to take the opportunity to thank the EPSRC and Gaussian Inc. for funding my PhD.

My next thanks might seem like a bit of an odd one but... I feel that I have learnt a lot about myself during the last 4 years, and though it is terribly self

serving, I would actually like to thank myself from 4 and a half years ago, who does feel like an entirely different person, for daring to believe that I could actually do this.

I would also like to take this opportunity to thank some very random people, none of whom I have met, and things for some random reasons. So big thanks to: Haruki Murakami for writing beautiful stories, Spotify for providing me with all the music I can listen to and more, Netflix - we all know why, Justin Roiland and Dan Harmon for giving the world Rick Sanchez, and Pagan Min and his hapless army.

Words cannot express how much I appreciate the love and support given to me by my family. They have been there for me through everything, even when I was terrible. I love you Mum, Nanna and my little big Sister Kyra.

Of course, I also could not have gotten through this experience without the patience, understanding, and boundless optimism of my supervisor Graham Worth. I think I owe you a pint!

And finally, thank you Harry.

This thesis and the work described in it are entirely my own, except where I have acknowledged *either* help from a named person *or* a reference is given to a published source or a thesis. Text taken from another source will be enclosed in quotation marks and a reference will be given.

June 11, 2018

List of Publications

1. K. E. Spinlove, M. Vacher, M. Bearpark, M. A. Robb, and G. A. Worth. Using quantum dynamics simulations to follow the competition between charge migration and charge transfer in polyatomic molecules. *Chem. Phys.*, 482:52–63, 2017. DOI: 10.1016/j.chemphys.2016.10.007.

2. G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne. Quantum dynamics simulations using gaussian wavepackets: The vMCG method. *Int. Rev. Phys. Chem.*, 34:269–308, 2015. DOI: 10.1080/0144235X.2015.1051354.

Contents

List of Figures **xiii**

List of Tables **xvii**

Glossary **xxv**

1 Introduction **1**

2 Theory **8**

2.1 Introduction 8

2.2 The Schrödinger Equation 8

2.3 The Hamiltonian 10

2.4 The Born-Oppenheimer Approximation 12

2.5 Beyond the Born-Oppenheimer Approximation 16

3 Methodology **22**

3.1 Introduction 22

3.2 Electronic Structure 22

3.2.1 The Hartree-Fock Self-Consistent Field Method 24

3.2.2 Basis Sets 36

3.2.3 Beyond Hartree-Fock 47

3.3 Nuclear Dynamics 62

3.3.1 Multi-Configuration Time-Dependent Hartree Method . . . 65

4 Formimidic Acid Proton Transfer **85**

4.1 Introduction 85

4.2 Electronic Structure Calculations 89

4.2.1 Ground State Calculations 90

4.2.2 Excited State Calculations 96

4.3 Direct-Dynamics 114

4.3.1 The Direct-Dynamics Protocol 114

4.3.2 State Population Analysis 117

4.3.3 Gross Gaussian Populations 123

4.3.4 Potential Energy Surfaces 128

4.4 Summary and Conclusions 131

5 2-Pyridone/2-Hydroxypyridine Proton Transfer **137**

5.1 Introduction 137

5.2 Electronic Structure Calculations 140

5.2.1 Excited State Calculations 141

5.3 Direct Dynamics 158

5.3.1 State Population Analysis 158

5.3.2 Gross Gaussian Populations 165

5.3.3 Potential Energy Surfaces 169

5.4 Summary and Conclusions 172

6 The Electronic Wavepacket **177**

6.1 Introduction 177

6.2 Theoretical Framework of the Ehrenfest Approach 179

6.2.1 The Ehrenfest Multi Configurational Gaussian Method . . . 183

6.2.2 The Model Hamiltonian 188

6.3 Results 193

6.3.1 Charge Migration versus Charge Transfer 193
6.3.2 Different Dynamics Models: Initial Conditions and Convergence 197
6.3.3 Direct Dynamics Calculations 200
6.3.4 Direct Ehrenfest Dynamics Calculations 208
6.4 Summary and Conclusions 210

7 Conclusions **215**

Bibliography **219**

List of Figures

1.1 The Jablonski Diagram. This shows the possible relaxation pathways following an excitation to an S_1 excited state, where A is absorption, ISC is intersystem crossing, IC is internal conversion, F is fluorescence, P is phosphorescence and R_1 to R_4 are non-radiative transitions between vibrational levels. 2

1.2 Wavepacket Dynamics. Depicted is a photoexcitation process with a wavepacket representing the nuclei moving over two potential energy surfaces related to the ground- and first-excited electronic states. . . 3

2.1 The adiabatic (b) and diabatic (c) surfaces of the butatriene molecule (a) along the C-C symmetric stretching mode (14Ag) and the CH_2 twisting (5Au) modes. 17

2.2 An example of a conical intersection (a) and a seam of intersection (b). a) This conical intersection shows an example of where a wavepacket is excited at the Franck-Condon point to a higher lying state, the wavepacket moves towards the conical intersection (the red and white arrows), the wavepacket passes through the intersection and either continues (white arrow) or bifurcates (blue arrows). b) This seam of intersection shows two possible relaxation pathways (red and blue arrows) through the seam, after an UV excitation from the S_0 to the S_1 states. The branching space is labelled orthogonal to the seam of intersection. 20

4.1 The 6-311+G∗ CCSD planar optimised geometries of a) formamide, b) formimidic acid and c) the transition state. The labels H1, H2, H3 are for clarity in the interpretation of the data in Table 4.3. . . 93

4.2 Character of MO for formamide using a CAS(10,9) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01). 101

4.3 Character of MO for formamide using a CAS(10,8) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01). 102

4.4 Character of MO for formimidic acid using a CAS(10,8) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01). 105

4.5 CAS molecular orbitals of formamide as a result of the CAS(10,8) SA8 calculation. 108

4.6 CAS molecular orbitals of formimidic acid from the CAS(10,8) SA8 calculation. 109

4.7 Diabatic state populations from DD-vMCG simulations of formamide starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. 120

4.8 Diabatic state populations from DD-vMCG simulations of formamide starting with a vertical excitation to various states. Final Results. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (norm^2): Thick black. 121

4.9 Diabatic state populations from DD-vMCG simulations of formimidic-acid starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. . . . 122

4.10 Diabatic state populations from DD-vMCG simulations of formimidic-acid starting with a vertical excitation to various states. Final Results. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (norm^2): Thick black. 123

4.11 The fraction of density going into different product channels from DD-vMCG simulations of formamide with SA8-CAS(10,8)/6-31G$*$ following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3. Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies, IP and OOP signify if the dissociation occurs of in- or out-of-plane. 125

4.12 The fraction of density going into different product channels from DD-vMCG simulations of formimidic acid with SA8-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3 Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies, IP and OOP signify if the dissociation occurs of in- or out-of-plane. 127

4.13 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formamide from DD-vMCG simulations . (a) ν_{11} (N-H_2 symmetric stretch) adiabatic (b) ν_{11} diabatic (c) ν_{12} (N-H_2 antisymmetric stretch) adiabatic (d) ν_{12} diabatic. 130

4.14 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formimidicacid from DD-vMCG simulations . (a) ν_{10} (N-H(2) stretch) adiabatic (b) ν_{10} diabatic (c) ν_{11} (O-H(1) stretch) adiabatic (d) ν_{11} diabatic. 131

5.1 The ground state optimised structures of a) 2-hydroxypyridine (TIM) and b) 2-pyridone (TAM). 137

5.2 The general structures of the lactam series ordered in increasing ring size, from β-lactam (4 membered ring) to ϵ-lactam (7 membered ring). 2-pyridone is a δ-lactam where the ring is conjugated. 138

5.3 The 6-31G$*$ MP2 optimised geometries of a) 2-pyridone, b) 2-hydroxypyridine and c) The transition state. The labels for the carbon atoms and hydrogen atoms are for clarity in the interpretation of the data in Tables 5.4 and 5.6. 142

5.4 The characterisation of the molecular orbitals and excited states for 2-pyridone using CAS(10,8) (left) and CAS(10,9) (right) sized active spaces. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. The key to the colour scheme is on the right hand side. The list of orbitals above the cells list those in the CAS space. The numbers in red are for states with significant oscillator strength (> 0.01). 143

5.5 The characterisation of the molecular orbitals and excited states for 2-hydroxypyridine using CAS(10,8) (left) and CAS(10,9) (right) sized active spaces. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. The key to the colour scheme is on the right hand side. The list of orbitals above the cells list those in the CAS space. The numbers in red are for states with significant oscillator strength (> 0.01). . . 145

5.6 CAS molecular orbitals of 2-pyridone from the CAS(10,8) SA7 calculation. 149

5.7 CAS molecular orbitals of 2-hydroxypyridine from the CAS(10,8) SA7 calculation. 149

5.8 Diabatic state populations from DD-vMCG simulations of 2-pyridone starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red. 159

5.9 Diabatic state populations from DD-vMCG simulations of 2-pyridone starting with a vertical excitation to various states. Final Results. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; Total density (norm^2): Thick black. 161

5.10 Diabatic state populations from DD-vMCG simulations of 2-hydroxypyridine starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red. 163

5.11 Diabatic state populations from DD-vMCG simulations of 2-hydroxypyridine starting with a vertical excitation to various states. Final Results. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; Total density (norm^2): Thick black. . 164

5.12 The fraction of density going into different product channels from DD-vMCG simulations of 2-pyridone with SA7-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3. Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies OOP signifies out-of-plane ring vibration. 166

5.13 The fraction of density going into different product channels from DD-vMCG simulations of 2-hydroxypyridine with SA7-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3 Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies OOP signifies out-of-plane ring vibration. 167

5.14 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-pyridone from DD-vMCG simulations . (a) ν_{30} (N-H stretch) adiabatic (b) ν_{30} diabatic. 170

5.15 Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-hydroxypyridine from DD-vMCG simulations . (a) ν_3 (O-H OOP1) adiabatic (b) ν_3 diabatic (c) ν_4 (O-H OOP2) adiabatic (d) ν_4 diabatic (a) ν_{30} (O-H stretch) adiabatic (b) ν_{30} diabatic. 171

5.16 Cut through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-hydroxypyridine from DD-vMCG simulations. The adiabatic S_0 and S_1 in the coordinates ν_3 (O-H OOP1) and ν_{30} (O-H stretch). 172

6.1 The allene molecule 178

6.2 The (a) diabatic and (b) adiabatic potential energy surfaces for the allene radical cation model Hamiltonian along the $Q_7(B_2)$ vibrational coordinate with a torsion angle kept at 45°. The molecular orbitals shown in the top panel define the diabatic surfaces, while those in the lower show the conjugation obtained at 45° due to the non-adiabatic coupling. 191

6.3 Population dynamics of allene after ionisation on the right hand end of the molecule. In (a) - (d) the molecule had a torsional twist of $\Theta = 75°$ and in (e) - (h) the angle was $\Theta = 45°$. (a), (e) Charge migration with static nuclei. (b), (f) full quantum dynamics of charge migration and charge transfer using the MCTDH method. (c), (g) Full quantum dynamics using the vMCG method with 50 GWPs. The MCTDH result is the black dotted line. (d), (h) Population dynamics with a single GWP, i.e. classical nuclei. 195

6.4 Expectation Values of the main vibrational coordinates (a) Q_3 (b) Q_4 (c) Q_7 from simulations starting with a torsion angle of $\Theta = 45°$. In green are the results from the MCTDH calculation, in red the classical Ehrenfest trajectory from the Franck-Condon point. (d) The reduced density along Q_7 is shown as a function of time. 196

6.5 Population dynamics of allene with a torsion angle of $\Theta = 45°$ after ionisation on the right hand end of the molecule calculated using different levels of theory and including only the main modes, Q_3, Q_4 and Q_7 in the dynamics. (a) The vMCG method with 30 variational GWPs with a width $1/\sqrt{2}$. (b) The clMCG method with 150 classical GWPs with a width $1/\sqrt{2}$ (c) The clMCG method with 200 classical GWPs with a width 0.4. (d) The iMCG method with 200 independent classical GWPs with a width 0.4 (e) The iMCG method with 200 independent classical GWPs with a width 0.4 and initial equal weights (f) The eMCG method with 150 classical GWPs with a width $1/\sqrt{2}$ In the iMCG, clMCG and eMCG calculations the initial GWPs were taken from a Wigner distribution. The full quantum result is shown in all plots as a black dotted line. 198

6.6 Allene doubly degenerate HOMO that are ionised to form the $\tilde{X}$ cation state. (a),(b) Conjugated orbitals with D_{2d} symmetry (c),(d) Localised π orbitals 201

6.7 Diabatic coupling from the SA2-CAS(4,3)/6-31G* potential energy surfaces of allene from DD-vMCG simulations. Cuts along the non-degenerate vibrations with all other coordinates kept at the equilibrium geometry. 204

6.8 Potential energy surface from the SA1-CAS(4,3)/6-31G* potential energy surfaces of allene from DD-vMCG simulations. Cuts along the key non-degenerate vibrations with all other coordinates kept at the equilibrium geometry. 205

6.9 State populations from DD-vMCG simulations of the allene radical cation with SA2-CAS(3,4)/6-31G* potential surfaces. All the population was on the first excited state with the geometry displaced from the Franck-Condon to have a torsion angle of (a) 45° and (b) 75°. 206

6.10 Localised orbital populations from DD-vMCG simulations of the allene radical cation with SA2-CAS(3,4)/6-31G* potential surfaces. The initial wavepacket was localised at one end of the molecule with the geometry displaced from the Franck-Condon to have a torsion angle of either (a),(c),(e) 45° or (b),(d),(e) 75°. (a) and (b) are results from full DD-vMCG calculations. (c) and (d) use a single time-independent wavepacket to simulate charge migration. (e), (f) are DD-vMCG simulations with classical GWPs. 207

6.11 Charge transfer using DD-vMCG with Ehrenfest surfaces at a CAS (4,3)/6-31G* level of theory and an initial electronic wavepacket equally populating the electronic states. The initial geometry is with a 45° torsion angle. (a) Spin densities on the atoms from a simulation with 1 GWP. Atoms 2,3 are the end carbons. Atom 1 is the central carbon. (b) Spin densities on the atoms from a simulation with 12 GWP. (c) State populations from a DD-vMCG calculation with an explicit 2-state representation and 1 GWP. (d) As (c) but showing only the first 12 fs. 209

List of Tables

4.1 A comparison of the energies (in Hartrees) of the optimised structures of formamide showing the level of theory versus the basis set. 91

4.2 A comparison of the CPU times taken (in seconds) for each of the calculations showing theory level versus basis set. Gaussian 09 [46] was used on a Xeon ES-2640 8-core workstation. 92

4.3 The optimised geometries of formamide (FAM), formimidic acid (FIM), and the transition state (TS) between the two isomers. The "CCSD opt" columns are the results of the optimisations as seen in Table 4.1, while the "CCSD pl" columns represent the optimisations with a planar starting geometry. Here, H1 is the proton that transfers from the nitrogen in FAM to the oxygen in FIM, H2 is the proton bonded to the nitrogen in both FAM and FIM, and H3 is the proton bonded to the carbon. These labels can be seen in Figure 4.1. 94

4.4 A comparison of the energies for the formamide and formimidic acid geometries calculated in Table 4.3. The energies of the optimised geometries are in Hartrees. The ΔE values are in eV and are the difference in energy between a) the relevant optimised, non-planar geometry and the planar geometry of formamide, b) the transition state, and c) the acid form. 95

4.5 Formamide energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA8-CAS(10,8)/6-31G* calculation 107

4.6 Formimidic acid energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA8-CAS(10,8)/6-31G* calculation 107

4.7 The normal mode frequencies of formamide calculated at different levels of theory at the optimised C_s structure where the numbering of the hydrogen atoms is as in Figure 4.1a), "ip" means in-plane and "oop" means out-of-plane. 110

4.8 The normal mode frequencies of formimidic acid calculated at different levels of theory at the optimised C_s structure where the numbering of the hydrogen atoms is as in Figure 4.1b), "ip" means in-plane and "oop" means out-of-plane. 111

4.9 Normal modes of formamide calculated at the CAS(10,8)/6-31G* SA8 level of theory at the optimised C_s structure. It should be noted that the arrow representing the CH stretching mode (ν_{10}) has been reduced in scale by 40%. 112

4.10 Normal modes of formimidic acid calculated at the CAS(10,8)/6-31G* SA8 level of theory at the optimised C_s structure. It should be noted that the arrow representing the OH stretching mode has been reduced in scale by 40%. 113

5.1 A comparison of the CPU times taken (in seconds) for equivalent optimisations of formamide and 2-pyridone showing selection from levels of theory and basis sets. Gaussian 09 [46] was used on a Xeon ES-2640 8-core workstation. 141

5.2 2-pyridone energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA7-CAS(10,8)/6-31G* calculation 147

5.3 2-hydroxypyridine energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA7-CAS(10,8)/6-31G* calculation 148

5.4 The normal mode frequencies of lactam calculated at different levels of theory at the optimised C_s structure where the numbering of the carbon atoms and hydrogen atoms is as in Figure 5.3a). 151

5.5 Normal modes of lactam calculated at the CAS(10,8)/6-31G* SA7 level of theory at the optimised C_s structure, where "Freq." is an abbreviation of frequency. 152

5.5 Continued 153

5.6 The Normal mode frequencies of lactim calculated at different levels of theory at the optimised C_s structure where the numbering of the carbons and hydrogens is as in Figure 5.3b). 155

5.7 Normal modes of lactim calculated at the CAS(10,8)/6-31G* SA7 level of theory at the optimised C_s structure, where "Freq." is an abbreviation of frequency. 156

5.7 Continued 157

6.1 Definitions and descriptions of the 15 vibrational modes in the allene molecule. Frequencies are for the ground-state calculated at the MP2/cc-pVTZ level. N are the number of harmonic oscillator DVR functions used in the MCTDH calculations, and n the number of multi-dimensional single particle functions with the combination of modes included by the bracket. 192

6.2 Frequencies of the normal modes of the neutral allene molecule at the MP2/6-31G* level of theory. The labels are the Wilson nomenclature used in the model Hamiltonian. The symmetry labels are for the D_{2d} point group to which the molecule belongs. Note the accidental degeneracy of modes ν_2 and ν_5 202

Glossary

ab initio From the latin term meaning "from the beginning". It is derived from "ab" meaning "from" and "initio" being the singular form of "initium" meaning "beginning". In the context of computational Chemistry it is usually in reference to methods by which the Time-Independent Schrödinger equation is solved, without the use of experimental data. 22

AIMS Ab-Initio Multiple Spawning. This is a semi-classical dynamics software, developed by the group of Todd Martinez, used for running on-the-fly trajectory surface hopping calculations with a focus on the use of Graphics Processing Units to speed up calculations [1, 2]. 65

AMBER Assisted Model Building with Energy Refinement. It is a software, originally developed by the group of Peter Kollman, used for running force field based molecular dynamics simulations of biomolecules. [3] It is also the name of the force fields used in the software. 64

ASC Active Space Characterisation (section). 100–102, 105, 108, 109, 150, 152

BOA Born-Oppenheimer approximation. 13, 16, 17, 22, 30, 65

CAP Complex Absorbing Potential. 118, 119, 121, 129, 138, 164, 172

CAS Complete Active Space. An abbreviated name for the CASSCF method, or an alternative name given to the active space (often referred to by redundant acronym syndrome sufferers as CAS space). v–ix, xi, xii, xiv, xv, xvii, xviii,

56, 57, 61, 98–117, 133–138, 140, 146, 147, 149–158, 160–162, 175–181, 184, 209, 213–216, 218, 222, 225

CASSCF Complete Active Space Self-Consistent Field (method). xx, 56–58, 61, 98, 209

CC Coupled Cluster (theory). 61, 63

CCSD Coupled Cluster, Singles and Doubles. v, xvi, 63, 91–97, 111–114, 117, 156, 160

CHARMM Chemistry at HArvard Macromolecular Mechanics. It is a software, originally developed by Martin Karplus, used for running force field based molecular dynamics simulations. [4] It is also the name of the force fields used in the software. 64

CI Configuration interaction. 50, 51, 53–58, 61–63, 67

CISD Configuration Interaction Singles and Doubles. 55

CMD Centroid Molecular Dynamics. This is a method for the semi-classical treatment of molecular systems, developments in the field can be found in References [5–7]. 66

DD-vMCG Direct-Dynamics Variational Multi-Configurational Gaussian (method). An on-the-fly quantum dynamics method based on vMCG. vi–viii, xiv, xv, xxii, 5–7, 78–80, 82, 83, 85, 90, 98, 111, 118, 119, 123–126, 133–138, 140, 145, 174, 180, 183, 185, 186, 191, 212–219, 221–227, 277, 291

DFT Density Functional Theory. An alternative representation to MO theory methods, Density Functional Theory utilises n one-electron Schrödinger like equations, known as the Kohn-Sham equations, which are used to construct external and effective potentials, which act upon the electrons of a system.

In 1998, Walter Kohn was awarded the Nobel Prize in Chemistry for the development of Density Functional Theory. xxiv, 23, 24, 49, 263

DOF Degree(s) of freedom. 67–70, 72, 74

DVR Discrete variable representation. xix, 68, 199

Elk Test An Elk Test in the automobile industry, also known as the Moose Test or more formally the Evasive Maneuver test, is defined on wikipedia as a test "performed to determine how well a certain vehicle evades a suddenly appearing obstacle." [8] In the context of this work, the choice of a previously unstudied systems containing 6 and 12 atoms, requiring an unknown number of excited states with an unknown number of non-adiabatic features presents an extreme form of an Elk Test to the DD-vMCG method. 6

eMCG Ehrenfest Multi-Configurational Gaussian (method). In this method the Gaussians follow classical trajectories. xiii, 6, 195, 196, 205, 206, 208, 217, 221

ESIPT Excited-State Intramolecular Proton Transfer. 144

FAM Formamide (or methanamide). xvi, 86–90, 94, 95, 97, 99, 100, 105, 108–112, 117, 118, 122, 128, 129, 131, 132, 145, 146, 150, 164, 257, 271

fermion A fermion is any subatomic particle that is characterised by Fermi-Dirac statistics, and by definition obeys the Pauli exclusion principle. Consequently it is cyclic to use "fermion" as a term in the definition of the Pauli exclusion principle. 27

FIM Formimidic acid (or formamidic acid). xvi, 90, 94, 95, 97, 99, 108–111, 113, 117, 118, 122, 128, 129, 132, 145, 146, 150, 164, 258, 271

FK-QCW Feynman-Kleinert Quasi-Classical Wigner. This is a method for the semi-classical treatment of molecular systems, developments in the field can be found in References [9, 10]. 66

GGP Gross Gaussian Population. 128, 130

G-MCTDH Gaussian - Multi-Configuration Time Dependant Hartree (method) where some of the SPFs of the MCTDH method are replaced by Gaussian functions. 71

GROMACS GROningen MAchine for Chemical Simulations. It is a software, originally developed by the Department of Biophysical Chemistry at the University of Groningen, used for running force field based molecular dynamics simulations of nucleic acids, proteins and lipids. [11] It is an open-source software based on the GROMOS force fields. xxiii, 64

GROMOS GROningen MOlecular Simulation. Is the name of the force fields used in the GROMACS software package. GROMOS is also the name for a molecular dynamics simulation software, [12] which itself also uses the GROMOS force fields [13]. xxiii

GTO Gaussian type orbital(s). 43–45

GWP Gaussian wavepacket. vi, vii, xii–xv, 5, 71, 72, 78–80, 120–125, 127, 128, 138, 139, 164, 166, 168–171, 181, 186, 194, 196, 202–208, 210, 215–223, 288

Hadean The Hadean Eon is the geologic eon spanning from the formation of Earth, around 4.6 billion years ago, to the Archean Eon, 4 billion years ago. 89

HCN Hydrogen cyanide (or hydrocyanic acid). It is the most abundant 3-atom organic compound in the universe. 87, 88

iMCG Independent (G)aussian Multi-Configurational Gaussian (method). In this method the Gaussians are no longer coupled, hence "independent" Gaussians. xiii, 196, 203, 206, 208, 221

LCAO Linear combination of atomic orbitals, sometimes extended to LCAO-MO (see MO glossary entry below). 5, 37, 38, 40–42, 45, 55

LHA Local harmonic approximation. 75, 79–82, 130, 209

MCE Multi-Configurational Ehrenfest (method). 186, 195

MCSCF Multiconfiguration Self-Consistent Field (theory). 56

MCTDH Multi-Configuration Time Dependant Hartree (method). xii, xiii, xix, xxiii, xxvii, 5, 67–71, 73, 118, 186, 191, 199, 201, 202, 204, 205, 220–222

MO Molecular-Orbital (theory). An alternative representation to DFT methods. More information on Molecular Orbital theory is found in Chapter 3. v, vi, xxi, 23, 24, 38, 41, 80, 103, 104, 107, 199, 201

MP Møller-Plesset (perturbation theory). ix, xix, 60, 61, 91–93, 111–114, 117, 146, 147, 156, 157, 160, 161, 199, 209, 211

Newton-X Newtonian DynamiX. This is a semi-classical dynamics software, developed by Mario Barbatti and collaborators, used for running excited state molecular dynamics calculations, as well as simulations of absorption and emission spectra [14, 15]. 65

on-the-fly "As needed." Usually referring to a dynamics method that does not require precomputed potential energy surfaces and hence the surfaces are calculated as needed. xxi, 79, 191

OTS Office of Technical Services. 88

PES Potential energy surface. 13, 80, 81

PIMD Path-Integral Molecular Dynamics. 65

RAS An abbreviated name for the RASSCF method, or, if accompanied by the numbers 1, 2, or 3, indicating the region in the active space. Amusingly RAS users are often sufferers of redundant acronym syndrome - which is also abbreviated to RAS. 58

RASSCF Restricted Active Space Self-Consistent Field (method). xxv, 58

RNA World Hypothesis The RNA (ribonucleic acid) World Hypothesis, originally proposed separately by C. R. Woese [16], L. E. Orgel [17] and F. H. C. Crick [18] in the 1960s, is an origin of life theory postulating that primitive life depended exclusively on RNA for the storage of genetic information and catalysis of chemical reactions. This theory was supported by the discovery, in the 1980s, of ribozymes [19, 20] (RNA enzymes) and the later discovery that peptide chain assemblage in ribosomes is not catalysed by proteins but by RNA itself [21, 22]. 89

RPMD Ring Polymer Molecular Dynamics. This is a method for the semi-classical treatment of molecular systems, developments in the field can be found in References [23–25]. 66

RS Rayleigh-Schrödinger (perturbation theory). 58

SCF Self-consistent field (method). 37, 41, 45, 55, 56

SgrA Sagittarius A - Observed terrestrially in the Sagittarius constellation, SgrA is located towards the centre of the Milky Way, around 8 kiloparsecs from Earth. It is comprised of three regions including the proposed location (SgrA*) for the galactic central supermassive black hole. 89

SgrB2 Sagittarius B2 - Observed terrestrially in the Sagittarius constellation, SgrB2 is the largest giant molecular complex (star forming cloud) in the contiguous region of the galactic centre. 89

SHARC Surface Hopping including ARbitrary Couplings. This is a semi-classical dynamics software, developed by the group of Leticia González, used for the representation of non-adiabatic effects as a result of photophysical processes [26–28]. 65

SPF Single-particle function(s). xxiii, xxvii, 67–71

STO Slater type orbital(s). 43–46

TAM 2-pyridone, or lactam. viii, 142–147, 153–155, 160, 171, 172

TDSE Time Dependent Schrödinger Equation. 4, 5, 9, 10, 22, 64, 66, 67, 187, 193, 194

TIM 2-hydroxypyridine, or lactim. viii, 142–147, 152, 154, 155, 160, 172, 173

TISE Time Independent Schrödinger Equation. 4, 5, 9, 10, 22, 53, 59, 64

VB Valence Bond (theory). 23, 24

veMCG Variational Ehrenfest Multi-Configurational Gaussian (method). In this method the Gaussians follow fully quantum trajectories. 195

vMCG Variational Multi-Configurational Gaussian (method) where all SPFs of the MCTDH method are replaced with Gaussian functions. xii, xiii, xxi, 5, 6, 71, 72, 77, 78, 191–193, 196, 202, 204–207, 220, 221, 225

W33A Protostellar object W33A is found in the giant molecular complex (star forming cloud) W33, located at $l \simeq 12.8^o$ (galactic longitude), approximately 2.4 kiloparsecs from Earth [29]. 89

XSC (E)Xcited State Characterisation (section). 100, 101, 105, 106, 108, 109, 150, 152

Chapter 1

Introduction

In the study of fundamental Chemistry the calculation of ground state, static molecules can yield useful information such as stable structures (energetic minima), the location of transition states, and the calculation of vibrational frequencies. However, in order to elucidate the experimentally observed behaviours of a molecule, the dynamical behaviour also needs to be determined. In addition, in some systems the excited states of the system should be considered. This is especially important in the interpretation of photophysical experimental results, when, in general, multiple competing pathways are available for the relaxation of a molecule following a photoexcitation. A familiar summary of these relaxation pathways is given by the Jablonski diagram, Figure 1.1.

This diagram depicts states and the pathways connecting transitions between them. These transitions are classified as either radiative, if a photon of light is emitted along with the transition, or non-radiative if it is not. Fluorescence and phosphorescence are radiative transitions which occur between states of the same spin, or different spins, respectively. Internal conversion and inter-system crossing are the non-radiative transitions that occur between states of, again, states of the same spin and different spin, respectively.

The Jablonski diagram is, however, an essentially static picture with no relationship to nuclear motion. For this, chemistry requires the use of potential

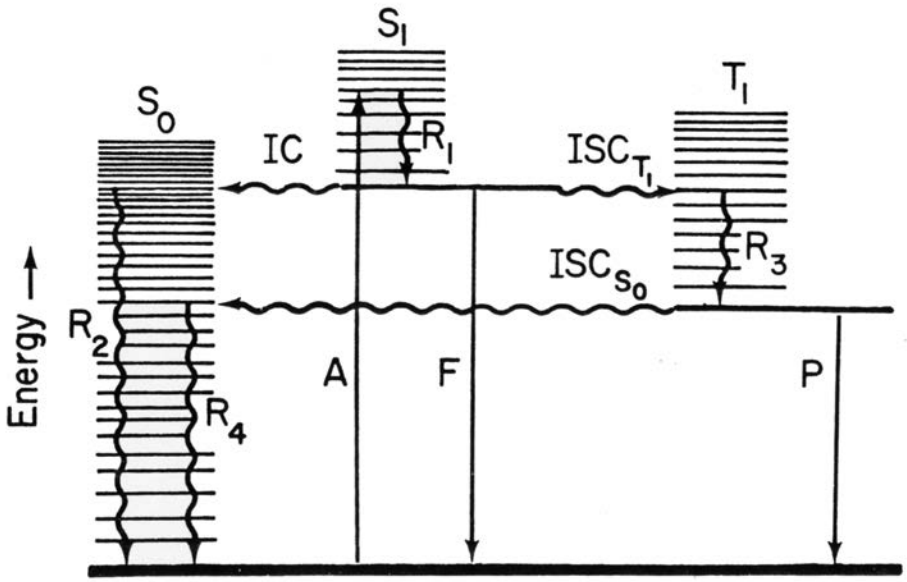


Fig. 1.1: The Jablonski Diagram. This shows the possible relaxation pathways following an excitation to an S_1 excited state, where A is absorption, ISC is intersystem crossing, IC is internal conversion, F is fluorescence, P is phosphorescence and R_1 to R_4 are non-radiative transitions between vibrational levels.

energy surfaces, such as those shown in Figure 1.2. In this, a wavepacket representing the molecular wavefunction is depicted moving over such surfaces, which give the electronic energy as a function of molecular geometry.

Spectroscopic methods are the best way to study the dynamics of a photo-excited molecule, giving direct information on the states accessible to excitation and the timescales of processes. In the 1960s and 70s techniques were developed for the observation of nanosecond (10^{-9}) and picosecond (10^{-12}) behaviours, such as nonradiative decay, which represented a major leap forward from the previous nonradiative pathways which had been inferred from radiative experiments. [30]

It was the work of A. H. Zewail, who pioneered the use of ultrafast laser pulses to study femtosecond (10^{-15}) timescale molecular processes [31,32], which enabled research to access directly the timescales of molecular motion. This used a pair of laser pulses in what is known as a pump-probe configuration to selectively excite and examine a set of molecules at known times. For this he was awarded the 1999 Nobel Prize in Chemistry in the new field of experimental "femtochemistry". [33]

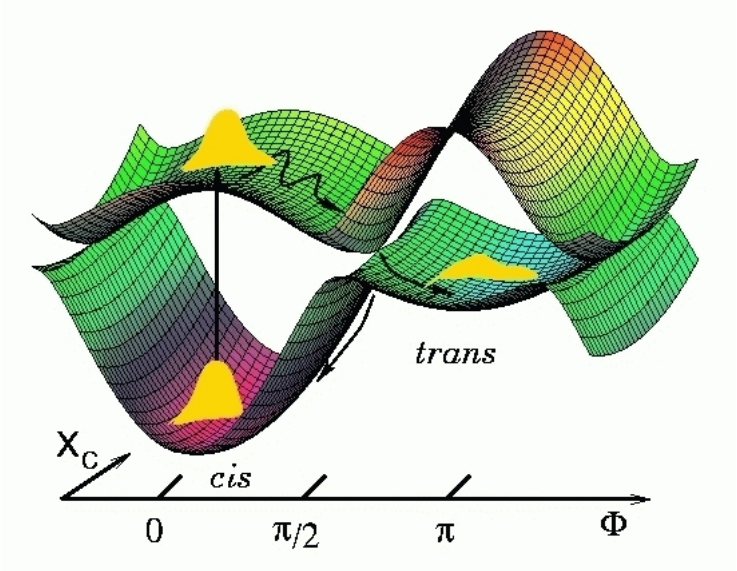


Fig. 1.2: Wavepacket Dynamics. Depicted is a photoexcitation process with a wavepacket representing the nuclei moving over two potential energy surfaces related to the ground- and first-excited electronic states.

Femtosecond spectroscopy has been applied in the study of a range of molecular behaviours. For example, in a classic set of experiments Zewail and co-workers observed a bond vibrating in I_2 [34], a bond breaking in I–CN [35] and the NaI molecule undergoing dynamics in which both bond-breaking and vibration are taking place simultaneously [36]. The latter is particularly interesting as an example of what is termed "non-adiabatic" photochemistry where more than one potential energy surface is required to describe the dynamics correctly.

Femtochemistry has been used to great success particularly in the field of biochemistry. Here, these pump-probe experiments have been used, for example, to observe the excited state dynamics of DNA [37] and the cis-trans photoisomerization of retinal [38], hence furthering the understanding of biological processes and aiding in the prediction of behaviours. They have also lead to the deeper understanding of light harvesting processes, such as photosynthesis and the behaviour of the Green Fluorescent Protein [39, 40], and hence lead to the development of synthetic light harvesting processes used in research developing dye sensitised solar

cells and imaging techniques. Due to the nature of femtosecond laser experiments, techniques involving laser control of biological processes, essentially light induced switches, have been developed. [30]

A more recent development, which develops from the femtochemistry techniques, is the field of attosecond spectroscopy. [41, 42] In these state-of-the-art experiments the dynamics of the electrons of a system can be observed allowing not only for the study of charge transfer, as in the femtosecond experiments, but also charge migrations whereby the nuclear dynamics of a system are induced by the movement of electrons.

Although these spectroscopic techniques have led to a greater understanding of these light/laser induced processes, the results of these experiments are often difficult to interpret. It is the complementary role of computer simulations which are necessary in the understanding of the fundamental reasons as to how and why the processes occur.

In order to carry out these dynamical calculations it is hence necessary to solve the time-dependent Schrödinger equation (TDSE). However, due to the inherent complex nature of this wave equation, the exact solutions to the TDSE is only possible for the simplest of molecules. Consequently a variety of methods exist utilising varying degrees of approximation to the TDSE, ranging from classical to semi-classical, through to "full" quantum methods. A brief summary of these classical and semi-classical methods is given in the introduction to nuclear dynamics in Chapter 3.3.

In this thesis an overview of the underlying principles to all of the methods is described in Chapter 2, including derivations of both the time-independent Schrödinger equation (TISE) and the TDSE, manipulations of the Hamiltonian, the Born-Oppenheimer approximation, which allows the definition of potential energy surfaces, and the description of systems in which the Born-Oppenheimer

approximation breaks down.

In Chapter 3.2 methods for solving the TISE are discussed, including a comprehensive overview of the ramifications of the Pauli exclusion principle in the treatment of electrons in Hartree-Fock theory, the use of the linear combination of atomic orbitals (LCAO) in the development of basis sets, and methods by which electron correlation can be accounted. In Chapter 3.3, after a brief introduction which discusses the existing methods by which the TDSE can be approximately solved, the multi-configurational time-dependent Hartree (MCTDH) quantum dynamics method is introduced, accompanied by the ansatz equations which form the fundamental framework for the method. This is followed by a detailed description of the variational multi-configurational Gaussian (vMCG) method, which uses a Gaussian Wavepacket (GWP) basis to describe the nuclear wavefunction and its relation to the grid-based MCTDH ideology. This leads naturally on to the description of the direct dynamics vMCG DD-vMCG method. In this, the potential surfaces are generated on-the-fly using quantum chemistry programs. Included is a thorough description of the key features of the method: the use of a database to store results, a Hessian updating scheme and the propagation diabatisation process which allows the non-adiabatic coupling to be correctly included. This work has been presented in a recent review "Quantum dynamics simulations using Gaussian wavepackets" [43].

The DD-vMCG method is relatively new and its application, thus far, has mainly been involving tests by comparison to existing systems available to the MCTDH and vMCG methods, as well as some interesting studies into small molecular systems such as ozone. The aim of this series of studies is essentially, what the automobile industry and the MCTDH developers refer to as, an extreme form of "Elk Test", a test designed to see when a vehicle (or program) falls over when swerving around a large animal (or molecule). In Chapter 4, a study of for-

mamide, and its acid analogue formimidic acid, is presented exploring what is required from external electronic structure calculations, such as basis set, level of theory and number of excited states, in order to carry out a meaningful direct dynamics calculation, as well as the results and discussion of a set of calculations using the DD-vMCG method. In Chapter 5, the system size is increased from formamide's 6 atoms to 12 atoms, in a study of 2-pyridone, and its acid analogue 2-hydroxypyridine, which have the same active sites as formamide, in the previous study. As with the previous chapter, a study of 2-pyridone, and its acid analogue, is presented exploring the requirements from the external electronic structure program, as well as the results and discussion of a set of calculations using the DD-vMCG method.

In Chapter 6 a new method for the calculation of Ehrenfest trajectories is presented. The motivation of this method is to model and understand the results of the burgeoning field of attosecond spectroscopy. Traditional quantum dynamics studies, including the DD-vMCG method, require a set of potential surfaces and couplings for all states of interest. This becomes a challenge when more than two or three states are included. As a result of the bandwidth of attosecond lasers many states are accessed and the existing methods are no longer applicable. In Ehrenfest dynamics, the electronic information from all states is collapsed down and represented as a single potential energy surface over which the electrons move, potentially allowing a complete description of these experiments. The theoretical details of this method are described followed by the results of model calculations using the novel Ehrenfest multiconfigurational Gaussian (eMCG) method in comparison to equivalent calculations using the vMCG method. The results presented here follow those published in the paper "Using quantum dynamics simulations to follow the competition between charge migration and charge transfer in polyatomic systems." [44] This is followed by results from direct dynamics calculations

on electronic wavepackets using DD-vMCG and, as an outlook, the first results obtained after linking DD-vMCG in the Quantics program [45] to Ehrenfest potential surfaces provided by a development version of the Gaussian quantum chemistry program [46].

Chapter 2

Theory

2.1 Introduction

If one's aim is to conduct computational or, more fundamentally, theoretical modelling and hence study of a physical system one must first outline the key mathematical principles that form the foundations of the work. This chapter will outline these key principles. The methods by which these principles are followed and approximated will be outlined in the Methodology section (Chapter 3).

In this chapter the Schrödinger equation, in various forms, will be presented with an accompanying description of the construction of the Hamiltonian, dependent on the system of interest. The principles of the Born-Oppenheimer approximation will be discussed followed by demonstrations and implications of its use. Finally, systems for which the Born-Oppenheimer approximation breaks down will be introduced and the implications of the choice of potential representation of the system will be briefly discussed.

Before continuing it should be stated that all theories described herein are based in nonrelativistic quantum mechanics.

2.2 The Schrödinger Equation

In 1926, Erwin Schrödinger's paper "The undulatory theory of the mechanics of atoms and molecules" [47] formed the basis for his development of a wave equation,

a type of linear partial differential equation, which mathematically describes the changes of a system in which quantum effects are significant. His formula for the description of a physical system in terms of a wavefunction, now known as the Schrödinger equation, gives rise to both a time dependent and a time independent form. The general form of the time independent Schrödinger equation (TISE)

$$\hat{H}\psi = E\psi \tag{2.1}$$

describes a system where the Hamiltonian operator $\hat{H}$ acts upon the wavefunction, ψ. This results in a wavefunction, ψ, which is proportional to the initial wavefunction with a proportionality constant E.

The general form of the time dependent Schrödinger equation (TDSE)

$$i\hbar\frac{\partial}{\partial t}\Psi(\boldsymbol{r},t) = \hat{H}\Psi(\boldsymbol{r},t) \tag{2.2}$$

describes a quantum wavefunction, Ψ, evolving in time where $i{=}\sqrt{-1}$, $\hbar$ is the reduced Planck constant ($\hbar{=}\frac{h}{2\pi}$), $\frac{\partial}{\partial t}$ is the partial derivative with respect to time, $\boldsymbol{r}$ and t are the position vector and time respectively, and $\hat{H}$ is the Hamiltonian.

In order to demonstrate the mathematical relationship between these two equations a standard *separation of variables* technique from the theory of differential equations, may be used. [48] Assuming that the potential of the system depends only on the spacial variables ($\boldsymbol{r}$), $\Psi(\boldsymbol{r},t)$ can be written as a product of functions in spacial and temporal (t) variables.

$$\Psi(\boldsymbol{r},t) = \psi(\boldsymbol{r})T(t) \tag{2.3}$$

If the partial derivative, with respect to time, is then taken

$$\frac{\partial\Psi}{\partial t} = \psi\frac{dT}{dt} \tag{2.4}$$

where T has the form

$$T = \exp(\frac{-iE}{\hbar}t) \tag{2.5}$$

Equation 2.3 can therefore be rewritten as

$$\frac{\partial \Psi}{\partial t} = \frac{-iE}{\hbar}\exp(\frac{-iE}{\hbar}t)\psi \tag{2.6}$$

By substituting this back into the TDSE (Equation 2.2) gives

$$\begin{aligned}\hat{H}\Psi = \hat{H}(\psi T) = i\hbar\frac{\partial}{\partial t}(\psi T) &= i\hbar\psi(\frac{-iE}{\hbar}T) \\ &= E\psi T\end{aligned} \tag{2.7}$$

Dividing by T hence gives the TISE (Equation 2.1). Consequently, it can be said that the TISE is an eigenvalue equation which describes the stationary states of, or definite energy solutions to, the TDSE. Contextually these may also be referred to as energy eigenstates or, with certain caveats discussed later (Chapter 3), as orbitals.

2.3 The Hamiltonian

The Hamiltonian, $\hat{H}$, as used in quantum mechanics is based on Hamilton's mathematical formalism of classical mechanics. In the most basic physical (classical) representation of the Hamiltonian, a single particle in one dimension, H is written

$$H = T + V \tag{2.8}$$

in which T is the kinetic energy of the system, expressed as

$$T = \frac{p^2}{2m} \tag{2.9}$$

where m is the mass of the particle and p is the momentum $m\nu$. The potential energy of the system, V, is expressed as

$$V = V(q) \tag{2.10}$$

resulting in a Hamiltonian expressed as

$$H = \frac{p^2}{2m} + V(q) \tag{2.11}$$

It can be seen here that T is a function of the momentum, p, alone, while V is a function only of the spacial coordinate, q. Since T and V contain time as an implicit variable their expressions are described as scleronomous.

When the Hamiltonian is used to describe a single particle in one dimension in terms of the Schrödinger equation, i.e. Hamiltonian quantum mechanics, potential and kinetic energies are now operators. Given that the momentum now also becomes an operator, $\hat{p}$, defined as

$$\hat{p} = -i\hbar\frac{\partial}{\partial x} \tag{2.12}$$

equations 2.8 and 2.11 can be rewritten as

$$\begin{aligned} \hat{H} &= \hat{T} + \hat{V} \\ &= -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x,t) \end{aligned} \tag{2.13}$$

In this case, the potential operator, V, has the time constraint explicitly defined and should now be referred to as rheonomous.

Extending the Schrödinger Hamiltonian further into 3 dimensions, including N non-interacting particles, by utilising the Laplacian operator,

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \tag{2.14}$$

and expanding the expression for $\hat{T}$ and $\hat{V}$ to

$$\hat{T} = \sum_{i=1}^{N} -\frac{\hbar^2}{2m_i}\nabla_i^2 \tag{2.15}$$

$$\hat{V} = \sum_{i=1}^{N} V(\boldsymbol{r}_i, t) \tag{2.16}$$

it can hence be written

$$\begin{aligned} \hat{H} &= \sum_{i=1}^{N} \left(-\frac{\hbar^2}{2m_i} \nabla_i^2 + V_i \right) \\ &= \sum_{i=1}^{N} \hat{h}_i \end{aligned} \tag{2.17}$$

where the Hamiltonian is the sum of the Hamiltonians for each particle. The reasons for explicitly defining this last step will become apparent in Chapter 3.2.1.

Although this is a special case, as there are usually forces interacting with and between the particles, it can be seen here that the Schrödinger Hamiltonian can be separated and reformulated depending on said conditions.

2.4 The Born-Oppenheimer Approximation

In order to solve the Schrödinger equation for a molecular system it is intuitive, for a chemist, to first express the Hamiltonian in separate terms of T and V for the electronic and nuclear parts,

$$\hat{H} = \hat{T}_e + \hat{T}_N + \hat{V}_e + \hat{V}_N + \hat{V}_{eN} \tag{2.18}$$

In this expression the subscript e denotes the electronic and N the nuclear parts, and $\hat{V}_{eN}$ is the potential coupling term between the electrons and nuclei. This can be expanded for a full-molecular system, in mathematical notation, to

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_k \frac{\hbar^2}{2m_k} \nabla_k^2 + \sum_{i<j} \frac{e^2}{r_{ij}} + \sum_{k<l} \frac{e^2 Z_k Z_l}{r_{kl}} - \sum_{i,k} \frac{e^2 Z_k}{r_{ik}} \tag{2.19}$$

where i and j denote the electrons, k and l denote the nuclei, m_e and m_k, respectively, are the mass of the electron and nucleus, e is the electronic charge, Z is an atomic number, and r_{ab} is the distance between the particles a and b. In a purely

qualitative inspection of this expression it can be seen that even for a simple diatomic system, the number of terms in the expansion becomes large. In fact, the largest system for which a closed form analytic solution has been obtained is for a hydrogenic atom. In this solution the Coulomb forces are used for the angular momentum coupling between the charges and the reduced mass is used as a mathematical tool to reduce the system to two, one-body problems. A full derivation of this can be found in Reference [49]. Consequently, in order to continue to use the Schrödinger equation an approximation must be made.

The Born-Oppenheimer approximation (BOA) in its essence is relatively simple, though its proof as a principle is subtle and deceptively difficult. In a most basic explanation in terms of its use, the BOA asserts that as the mass of an electron is small in comparison to the mass of a nucleus, electrons can be seen to instantaneously move with nuclear motion. As a consequence, the nuclei can be considered to be stationary in comparison to the motion of the electrons. In order to gain information such as the minimum energy conformation of a system, static electric potentials are then calculated at stationary geometric arrangements of the nuclei, and a potential energy surface (PES) is constructed.

Perhaps the simplest approach to defining the BOA in mathematical terms is to first consider the hydrogen molecule ion (H_2^+) in one dimension, here defined as the y-axis. If the Hamiltonian presented in Equation 2.19 is considered in the context of this system, it can easily be seen that the Laplacian will be reduced to a single partial derivative in y and, by reducing the three terms of the potential to one term, yields a Hamiltonian of the form

$$\hat{H} = -\frac{\hbar^2}{2m_e}\frac{\partial^2}{\partial y^2} - \sum_{k=1,2}\frac{\hbar^2}{2m_k}\frac{\partial^2}{\partial Y_k^2} + V(y, Y_1, Y_2) \tag{2.20}$$

where Y_k, with k=1,2, is the location of the nuclei, and y the location of the electron. The Schrödinger equation can, hence, be written

$$\hat{H}\Psi(y, Y_1, Y_2) = E\Psi(y, Y_1, Y_2) \tag{2.21}$$

In an attempt to gain a solution, the wavefunction can then be separated

$$\Psi(y, Y_1, Y_2) = \psi(y; Y_1, Y_2)\chi(Y_1, Y_2) \tag{2.22}$$

where χ is the nuclear wavefunction and ψ is the electronic wavefunction. The expression for the electronic wavefunction here is dependent on the position of the electron, which is itself parametrically dependent on the coordinates of the nuclei. If this trial form of the wavefunction is then substituted into Equation 2.21, the Schrödinger equation now has the form

$$\hat{H}\psi\chi = \chi T_e\psi + \psi T_N\chi + V\psi\chi + W = E\psi\chi \tag{2.23}$$

in which the parameter W is defined as

$$W = -\sum_{k=1,2} \frac{\hbar^2}{2m_k}\left(2\frac{\partial\psi}{\partial Y_k}\frac{\partial\chi}{\partial Y_k} + \frac{\partial^2\psi}{\partial Y_k^2}\chi\right) \tag{2.24}$$

This "W" accounts for non-adiabatic effects, discussed later in Chapter 2.5. Despite the fact that the elements of ψ are dependent on the nuclear coordinates, and hence the partial $\partial\psi/\partial Y_k$ is non-zero, the (relatively large) nuclear masses (m_k) in the denominator render the full expression to be almost zero so are neglected. Accordingly, Equation 2.23 can be rearranged and rewritten as

$$\psi T_N\chi + (T_e\psi + V\psi)\,\chi = E\psi\chi \tag{2.25}$$

In order to solve this equation the definition of this system must first be considered. If the electron is treated as existing in a potential dependent on the coordinates of the nuclei, a Schrödinger equation for the electron can be written

$$\begin{aligned} T_e\psi + V\psi &= E_e\left(Y_1, Y_2\right)\psi \\ &(= E_e\psi) \end{aligned} \tag{2.26}$$

In this expression for the electronic wavefunction, the eigenvalue E_e is dependent on the sum of the electronic contribution to the total energy and the internuclear repulsion potential at a given stationary coordinate for the nuclei. The implicit relation of E_e to the nuclear coordinate is what allows the potential energy surfaces to be constructed with a set of calculations for stationary nuclei.

Substituting Equation 2.26 into Equation 2.25 gives

$$\psi T_N\chi + E_e\psi\chi = E\psi\chi \tag{2.27}$$

As ψ is dependent on the electronic coordinate, and the system is normalised where

$$\int \psi^*\psi dy = 1 \tag{2.28}$$

ψ can be hence be integrated out (Scheme 2.29) and cancelled (Equation 2.30), giving the Schrödinger equation for the nuclear wavefunction χ

$$\left.\begin{aligned} \psi^*\psi T_N\chi + \psi^* E_e\psi\chi &= \psi^* E\psi\chi \\ \int \psi^*\psi T_N\chi dy + \int \psi^* E_e\psi\chi &= \int \psi^* E\psi\chi \\ T_N\chi\int \psi^*\psi dy + E_e\chi\int \psi^*\psi dy &= E\chi\int \psi^*\psi dy \end{aligned}\right\} \tag{2.29}$$

$$T_N\chi + E_e\chi = E\chi \tag{2.30}$$

when the nuclear potential has the form of a potential energy curve. The neglect of the coupling term, W, (Equation 2.24) and the definition of the electronic Schrödinger equation, $E_e\psi$, (Equation 2.26) is the BOA. It can hence be stated that the eigenvalue E in Equation 2.30 is the total energy of the system, within the Born-Oppenheimer approximation.

2.5 Beyond the Born-Oppenheimer Approximation

Although the Born-Oppenheimer approximation is a key approximation to be used when considering static molecules or, more specifically, systems in which the nuclear and electronic motion is not coupled, systems in which this coupled motion is significant presents a case where the Born-Oppenheimer approximation breaks down. Before the nuances of so-called beyond Born-Oppenheimer systems can be discussed it is helpful to first introduce some key concepts. The Born-Oppenheimer approximation as written in Equation 2.30 can also be termed as the time-independent adiabatic representation, the time-dependent form of which is

$$(T_N + V)\chi = i\hbar\frac{\partial\chi}{\partial t} \tag{2.31}$$

where T_N is the kinetic energy of the nuclei and V is the potential as a result of the electronic motion. The adiabatic representation uses the eigenfunctions of the electronic Schrödinger equation at a particular nuclear geometry. Consequently there is a set of surfaces, one for each eigenvalue, and the potential energy surfaces are ordered in energy. An example of adiabatic surfaces for the lowest two energies of butatriene cation along the C-C symmetric stretching and CH_2 twisting modes are given in Figure 2.1b).

An alternative representation of the energetic states of a system is the diabatic representation. The adiabatic electronic wavefunctions of a system can be treated by a unitary transformation at each point in space

$$\tilde{\boldsymbol{\psi}} = \boldsymbol{S}(\mathbf{r})\boldsymbol{\psi} \tag{2.32}$$

where $\tilde{\psi}$ is the electronic wavefunction in the diabatic representation, ψ is the electronic wavefunction in the adiabatic representation, and $\boldsymbol{S}(\mathbf{r})$ is the unitary

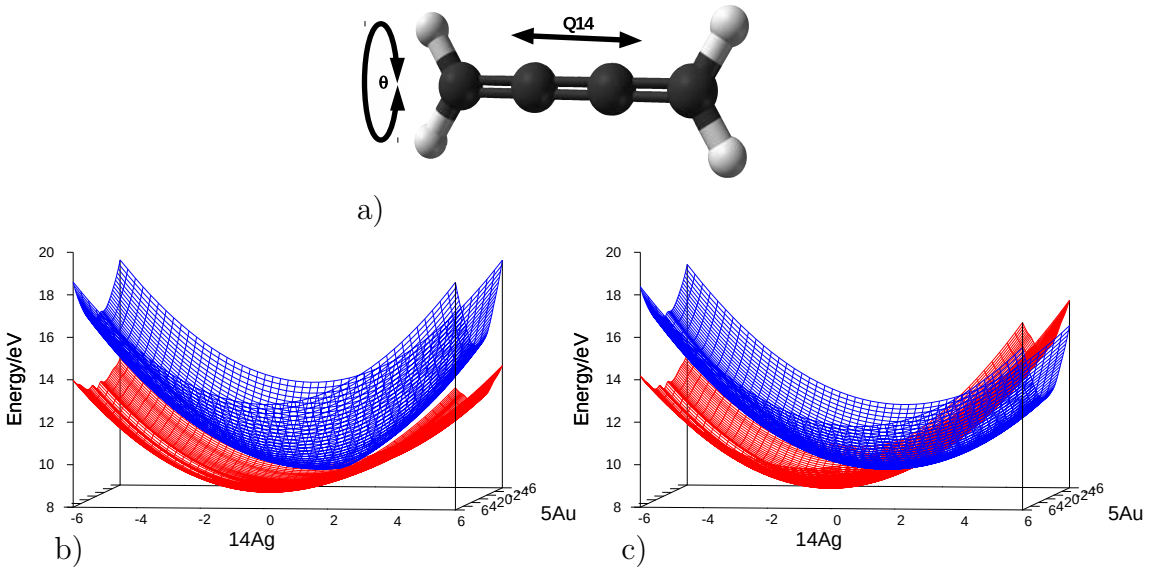


Fig. 2.1: The adiabatic (b) and diabatic (c) surfaces of the butatriene molecule (a) along the C-C symmetric stretching mode (14Ag) and the CH_2 twisting (5Au) modes.

transformation matrix which is a function of the nuclear coordinates $\mathbf{r}$. This results in the form for the Schrödinger equation in the diabatic representation as

$$(T_N\mathbf{1} + \boldsymbol{W})\boldsymbol{\chi} = i\hbar\frac{\partial\boldsymbol{\chi}}{\partial t} \tag{2.33}$$

It should be noted that the nuclear wavefunction χ is not the same as the wavefunction in the adiabatic representation (the difference defined in Equation 2.32). In this representation the energetic states are ordered by the character of the electronic state and consequently crossings between states can occur. It is this representation that is most intuitive to chemists. An example of the diabatic surfaces of butatriene along the C-C symmetric stretching and CH_2 twisting modes are given in Figure 2.1c). Here it can be seen that a degree of coupling between states is experienced by the systems as said electronic states may cross in energy. This is not "true" coupling between states as these crossings are essentially as a result of the matrix elements used in the unitary transformation.

However, although these two representations provide valuable information about the energetics and contributions of the electronic structure of the system, it is the

non-adiabatic coupling term (as in Equations 2.23 and 2.24) which contains the behaviours where the BOA breaks down. Using the full Born representation for the wavefunction

$$\Psi = \sum_i \chi_i \psi_i \tag{2.34}$$

which is a superposition of all of the eigenfunctions, ψ_i, of the electronic Hamiltonian operator, $\hat{H}_{el}$, Equation 2.33 can be recast

$$(\hat{T}_N \mathbf{1} + \boldsymbol{V} - \boldsymbol{\Lambda})\boldsymbol{\chi} = i\hbar \frac{\partial \boldsymbol{\chi}}{\partial t} \tag{2.35}$$

where $\boldsymbol{\Lambda}$ is the non-adiabatic coupling operator. If this equation is expanded to an infinite set, it includes all adiabatic electronic states and all coupling between these states. As coupling usually only occurs between neighbouring states in close proximity, these non-adiabatic coupling terms can be truncated, resulting in a finite set of decoupled (and coupled adiabatic) states, the term for which is the group Born-Oppenheimer approximation. This results in the matrix $\boldsymbol{\Lambda}$ having dimensions equal to the number of states in the sets. It can easily be seen that if this coupling term is ignored completely, the adiabatic representation is returned.

It is these non-adiabatic coupling terms that allow the coupling between the nuclear and electronic motion, commonly referred to as vibronic coupling. Following from Equation 2.24, the matrix elements of the non-adiabatic coupling operator, $\boldsymbol{\Lambda}$, are

$$\left.\begin{aligned} \hat{\Lambda}_{ij} &= \frac{\hbar^2}{2m}\left(2\left\langle \psi_i | \boldsymbol{\nabla} \psi_j \right\rangle \boldsymbol{\nabla} + \left\langle \psi_i | \boldsymbol{\nabla}^2 \psi_j \right\rangle\right) \\ &= \frac{\hbar^2}{2m}(2\boldsymbol{F}_{ij} \cdot \boldsymbol{\nabla} + \boldsymbol{G}_{ij}) \end{aligned}\right\} \tag{2.36}$$

where $\boldsymbol{F}_{ij}$ are the derivative couplings and $\boldsymbol{G}_{ij}$ are the scalar couplings [50, 51]. It can be seen that the diagonal terms of the derivative coupling matrix are

$$F_{ii} = \langle \psi_i | \boldsymbol{\nabla} \psi_i \rangle = 0 \tag{2.37}$$

while the off-diagonal terms can be written

$$F_{ij} = \langle \psi_i | \boldsymbol{\nabla} \psi_j \rangle = \frac{\langle \psi_i | \boldsymbol{\nabla} \hat{H} | \psi_j \rangle}{V_j - Vi} \tag{2.38}$$

where the adiabatic energies for states i and j, V_i and V_j, are at nuclear geometry $\mathbf{r}$. If the adiabatic states are degenerate, i.e. $V_j - V_i = 0$, then it can easily be seen that these off-diagonal terms in the derivative coupling matrix become infinite ($F_{ij} \to \infty$). It is at these points where conical intersections occur.

A conical intersection, also sometimes referred to as a photochemical funnel, molecular funnel or diabolic point, is a feature particularly common in the photochemistry of polyatomic molecules due to the high number of nuclear degrees of freedom as well as a manifold of close-lying electronic states. Conical intersections present a pathway for ultrafast, femtosecond timescale, interstate crossings allowing for radiationless decay from excited states to lower energy states or the ground state of a system. In the region of a conical intersection, the Born-Oppenheimer approximation breaks down as the energies of the states become singular and the off-diagonal coupling terms become infinite. A schematic of a conical intersection is shown in Figure 2.2a) and a conical intersection can be seen in the butatriene potential surfaces in Figure 2.1a).

As a conical intersection is a three-dimensional representation of two degrees of freedom (dimensions) of a system and the associated energy, conical intersections can only occur in systems where the number of degrees of freedom, N, is 2 or greater, i.e. $N \geqslant 2$. This provides an explanation as to why only avoided crossings can be present in a one-dimensional system, or in two dimensional (one degree of freedom, and energy) representation of a N-dimensional system. In Figure 2.2a) it can be seen that the conical intersection is point where the degeneracy of the

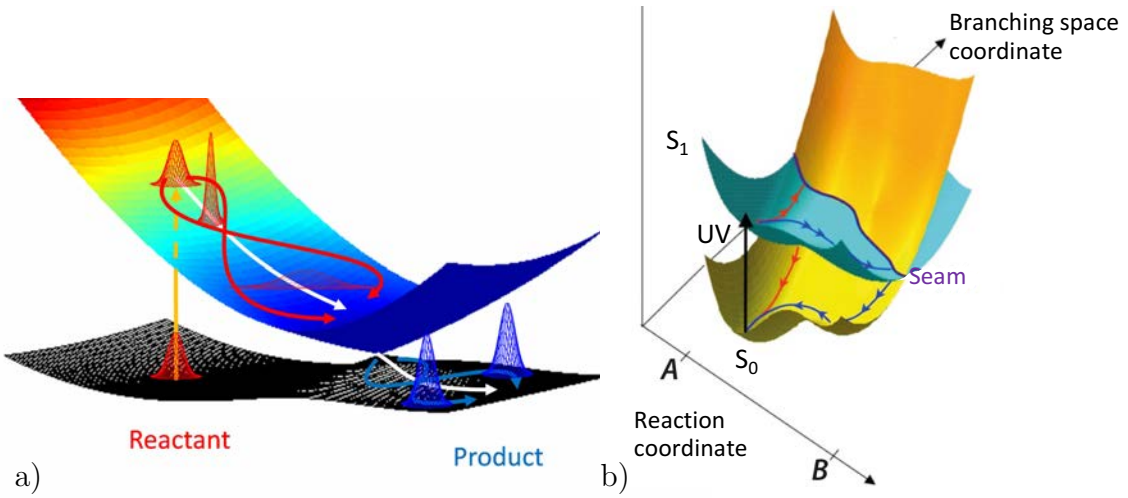


Fig. 2.2: An example of a conical intersection (a) and a seam of intersection (b). a) This conical intersection shows an example of where a wavepacket is excited at the Franck-Condon point to a higher lying state, the wavepacket moves towards the conical intersection (the red and white arrows), the wavepacket passes through the intersection and either continues (white arrow) or bifurcates (blue arrows). b) This seam of intersection shows two possible relaxation pathways (red and blue arrows) through the seam, after an UV excitation from the S_0 to the S_1 states. The branching space is labelled orthogonal to the seam of intersection.

states is lifted in two directions, or vectors, which is commonly referred to as the branching space. Figure 2.2b) is a representation of a seam of intersection, which occurs when the number of degrees of freedom in a system is greater than 2, i.e. $N > 2$. As a seam of intersection has an inherent greater dimensionality than a conical intersection, any point along the seam of an intersection between two states can be represented in the branching space as a conical intersection.

As stated previously, the mathematical description of points of intersection between states, also known as singularities, is inherently difficult and a representation is required with the aim to remove these singularities. In the diabatic representation it has been found [50, 52, 53] that if the relationship

$$\boldsymbol{\nabla S} = -\boldsymbol{FS} \tag{2.39}$$

is used, these singularities can be removed. In this case, however, the diabatic representation can only be described globally as all configurations are dependent

upon the neighbouring configuration. In addition, in order for the diabatic representation to be completely defined, it is necessary to specify $\boldsymbol{S}$ at a particular point where it can be chosen to make the diabatic and adiabatic representations equivalent. This is dependent on the problem at hand but is typically chosen to be the Franck-Condon point. It should also be noted that in order to avoid errors in the truncation of the matrix $\boldsymbol{F}$, Equation 2.39 requires a complete basis set, the feasibility of which is highly system dependent and will be discussed later in Chapter 3.

Chapter 3

Methodology

3.1 Introduction

In order to apply the theories outlined in Chapter 2 attention must first be turned towards the definition of the electronic and nuclear structure. The focus of Section 3.2 will be on the methods for the definition of electronic structure within the confines of the BOA, known commonly as *ab initio* methods, and consequently solving the TISE. In Section 3.3, methods for solving the TDSE, and hence situations in which the BOA breaks down, will briefly be presented and the methods used in this work will be presented in further detail.

3.2 Electronic Structure

Electronic structure, within the field of quantum chemistry, is a catch-all term for the methods by which the state of motion of electrons in a electrostatic field, created by the motionless nuclei, are described incorporating both the electronic wavefunction and associated energies. To reiterate, calculations of this type utilise the BOA and are hence used to solve the TISE. There are many methods for the description and hence calculation of electronic structure, most of which fall into three categories:

- Valence Bond Theory

- Molecular-Orbital Theory
- Density Functional Theory

Valence bond (VB) theory, for a chemist, is perhaps the most intuitive of these three categories as its foundation is directly analogous to the Lewis structure [54] representation of chemical bonds. Based on Heitler-London theory [55], which yielded the first quantum mechanical calculation of the bonding properties of the H_2 molecule, and further extended by Pauling to include the idea of Lewis pair bonding to develop definition of resonance and orbital hybridisation [56] [57], VB theory was the de rigueur method for the explanation and prediction of chemical bonding in the late 1920s to the mid 1950s [58].

Molecular-orbital (MO) theory, originally known as Hund-Mulliken theory [59], is the most familiar of the three categories due to its extensive development in computational chemistry and its applications as a illustrative tool in organic chemistry. Founded on Lennard-Jones' interpretation of the bonding in triplet oxygen [60], further extended by Hückel to include the ideas of the rotational hindrance and conjugation of π systems [61,62] and the description of unsaturated bonds [63,64], MO theory was refined in a theoretical setting by Robert Mulliken [65, 66] for which he gained the Nobel Prize in 1966 [67].

Electronic band structure is a concept most familiar to solid-state physicists, though its application as a model has been widely adopted within the field of computational materials chemistry in the form of density functional theory (DFT). DFT is a more generalised interpretation of physical systems essentially describing a mean free field of electrons acted upon by perturbations (for example nuclei) to the field. It was constructed by building upon the two Hohenberg-Kohn theorems [68] relating to the definition, within the confines of the Thomas-Fermi [69,70] model, of electronic motion influenced by an external potential. It was further developed by Kohn and Sham [71] to reduce the problem to a simplified model

describing non-stationary non-interacting electrons affected by an external potential.

Historically speaking, the acceptance of VB theory, 1928, only narrowly predates MO theory, 1931. The reasons for the eventual prevalence of MO theory, or to be more accurate the decline of VB theory, are nicely outlined in Chapter 1 of Shaik and Hiberty's "*A Chemist's Guide to Valence Bond Theory*" [72]. The conception of the Thomas-Fermi model, 1928, occurred in the same time period as the development of VB and MO theory, though the development of DFT in its own right occurred decades later, in the 1960s to 1970s.

Although interest in VB theory has had a resurgence in recent years [73, 74], MO theory's use in computational studies is more widely developed and electronic structure calculations in this work are carried out using this method.

3.2.1 The Hartree-Fock Self-Consistent Field Method

After the publication of Schrödinger's seminal paper [47], D. R. Hartree attempted to bring the ideas of the "old" quantum theory of Bohr [75] into the new, wave mechanical picture [76,77] based upon the central-field method. In the central-field method an electron is acted upon by an average charge distribution described as a spherically averaged sum of charge over the electrons which would then be seen as spherically symmetric around an atom centre. [78]

Hartree used the idea of a spherically averaged potential acting upon each electron, but instead trial electronic wavefunctions are selected, charge densities and potentials calculated, and the Schrödinger equation solved. This process is then repeated, using the result of the previous step as the trial wavefunction for the next. This iterative approach, carried out over a number of cycles, results in wavefunctions which are self consistent and hence can be described as converged. Hartree called this process the "*self-consistent field*" method [76].

Before going into details of this iterative method it is first necessary to introduce

some key concepts.

The Hartree Product and the Antisymmetry Problem

To apply this in the simplest terms, if the Hamiltonian is separable and considered as a sum over all of the one-electron terms it can be written as

$$\hat{H} = \sum_{i=1}^{N} \hat{h}_i \tag{3.1}$$

where N is the number of electrons and $\hat{h}_i$ is the operator describing the kinetic and potential energy of electron i. This expression is the full electronic Hamiltonian, if the electron-electron repulsion is ignored. The description here is analogous to, but not the same as, the expression in Equation 2.17.

By defining an orbital as a wavefunction for a single particle, an electron, each orbital can be expressed as either a spacial orbital, $\psi_i(\mathbf{r})$, or, for reasons which will become clear, as a spin orbital, $\chi_i(\mathbf{x})$. They are related by the definition

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ or \\ \psi(\mathbf{r})\beta(\omega) \end{cases} \tag{3.2}$$

where the spacial coordinates $\mathbf{r}$ are related to $\mathbf{x}$ by $\mathbf{x} = \{\mathbf{r}, \omega\}$, ω is the spin variable or coordinate, and $\alpha(\omega)$ and $\beta(\omega)$ are the spin functions related, respectively, to up and down spin. Consequently it can be seen that for every spacial orbital, two spin orbitals can be written. If the spacial molecular orbitals are assumed to form an orthonormal set (Equation 3.3) and the α and β spin functions are orthonormal, it follows that the spin orbitals are also orthonormal (Equation 3.4).

$$\int d\mathbf{r}\psi_i^*(\mathbf{r})\psi_j(\mathbf{r}) = \delta_{ij} \tag{3.3}$$

$$\int d\mathbf{x}\chi_i^*(\mathbf{x})\chi_j(\mathbf{x}) = \langle\chi_i|\chi_j\rangle = \delta_{ij} \tag{3.4}$$

With Equation 3.1 in hand, it can then be stated that the operator $\hat{h}$ has a set of eigenfunctions that can be taken to be a set of spin orbitals $\{\chi_n\}$

$$\hat{h}_i \chi_n(\mathbf{x}_i) = \epsilon_n \chi_n(\mathbf{x}_i) \tag{3.5}$$

As the Hamiltonian operator is separable, the many-electron eigenfunction can be constructed as a product of spin orbital wavefunctions for each electron. This can be written as

$$\Psi^{HP}(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) = \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2)\cdots\chi_n(\mathbf{x}_N) \tag{3.6}$$

where Ψ^{HP} is the eigenfunction of $\hat{H}$. If the eigenvalue E is also a sum of spin orbital energies of the spin orbitals of Ψ^{HP}

$$E = \epsilon_i + \epsilon_j + \cdots + \epsilon_n \tag{3.7}$$

the many-electron wavefunction may be written

$$\hat{H}\Psi^{HP} = E\Psi^{HP} \tag{3.8}$$

This form of a many-electron wavefunction is called a *Hartree Product* where electron one is described by spin orbital χ_i, electron two by χ_j, and so on. By application of the variational principle, this can be iteratively solved.

It should be noted that the Hartree Product is an independent-electron, or uncorrelated, wavefunction. This can be shown where if

$$|\Psi^{HP}(\mathbf{x}_1, \cdots, \mathbf{x}_N)|^2 d\mathbf{x}_1 \cdots d\mathbf{x}_N \tag{3.9}$$

defines the simultaneous probability of finding electron one in the volume element $d\mathbf{x}_1$ centred at $\mathbf{x}_1$, electron two in $d\mathbf{x}_2$ centred at $\mathbf{x}_2$, and so forth, given Equation 3.6 it can therefore be written

$$|\chi_i(\mathbf{x}_1)|^2 d\mathbf{x}_1 |\chi_j(\mathbf{x}_2)|^2 d\mathbf{x}_2 \cdots |\chi_n(\mathbf{x}_N)|^2 d\mathbf{x}_N \tag{3.10}$$

In this form it can be seen that the Hartree Product is a product of the probabilities that electron one is in $d\mathbf{x}_1$, multiplied by the probability that electron two is in $d\mathbf{x}_2$, etc.

However, in order for an exact wavefunction to be correct it not only must satisfy the Schrödinger equation, but must also satisfy the Pauli exclusion principle, also called the antisymmetry principle. In general terms the Pauli exclusion principle, initially applied as a new, two-valued quantum number [79] but later extended to any fermion [80], states that no two identical fermions may simultaneously occupy the same quantum state in a quantum system. In addition, on exchange of fermions the wavefunction must change sign. This can be expressed mathematically as

$$\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_j, \cdots, \mathbf{x}_N) = -\Phi(\mathbf{x}_1, \cdots, \mathbf{x}_j, \cdots, \mathbf{x}_i, \cdots, \mathbf{x}_N) \tag{3.11}$$

where the many-electron wavefunction, Φ, is antisymmetric with respect to the interchange of the spacial and spin coordinate, $\mathbf{x}$, of any two electrons. What may be immediately obvious here is that there is no explicit account of the indistinguishability of the electrons in the Hartree Product as written.

Slater Determinants

The Hartree product's failure to satisfy the antisymmetry principle was pointed out by J. C. Slater [81] and V. A. Fock [82] independently. Although Fock himself reformulated the Hartree Product in terms of group theory, a more widely accessible reformulation was published by Hartree [83] utilising Slater determinants.

If a two electron system in which the electrons occupy spin orbitals χ_i and χ_j is first considered, the Hartree product can be written in two ways. If electron one is put into χ_i and electron two into χ_j the Hartree product has the form

$$\Psi_{12}^{HP}(\mathbf{x}_1, \mathbf{x}_2) = \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) \tag{3.12}$$

If electron one is placed in χ_j and electron two in χ_i the Hartree product has the form

$$\Psi_{21}^{HP}(\mathbf{x}_1, \mathbf{x}_2) = \chi_i(\mathbf{x}_2)\chi_j(\mathbf{x}_1) \tag{3.13}$$

These Hartree products clearly distinguish between the two electrons. However, if the appropriate linear combinations of these two expressions are taken, it is possible to obtain a wavefunction which does obey the antisymmetry principle.

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = 2^{-1/2}(\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)) \tag{3.14}$$

Here a normalisation factor, $2^{-1/2}$, is used, and the minus sign is to ensure the wavefunction is antisymmetric with respect to the interchange of coordinates. Explicitly

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = -\Psi(\mathbf{x}_2, \mathbf{x}_1) \tag{3.15}$$

which is a reduced form of Equation 3.11. It can also be very simply proven that the normalisation factor is $2^{-1/2}$ by taking the integral of the normalised wavefunction, squared

$$\begin{aligned}
\langle\Psi|\Psi\rangle &= N^2\langle\chi_i\chi_j - \chi_j\chi_i|\chi_i\chi_j - \chi_j\chi_i\rangle \\
&= N^2\left(\langle\chi_i\chi_j|\chi_i\chi_j\rangle + \langle\chi_j\chi_i|\chi_j\chi_i\rangle - \langle\chi_i\chi_j|\chi_j\chi_i\rangle - \langle\chi_j\chi_i|\chi_i\chi_j\rangle\right) \\
&= N^2\left(1 + 1 + 0 + 0\right) \\
&= 2N^2 \\
\therefore N &= \frac{1}{\sqrt{2}}
\end{aligned} \tag{3.16}$$

In the form of a Slater determinant, the antisymmetric wavefunction of Equation 3.14 can be written

$$\Psi(\mathbf{x}_1, \mathbf{x}_2) = 2^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) \end{vmatrix} \tag{3.17}$$

It is useful to use this notation when increasing to an N electron system as

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \cdots & \chi_n(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \cdots & \chi_n(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \cdots & \chi_n(\mathbf{x}_N) \end{vmatrix} \tag{3.18}$$

Assuming the electron and spin orbital labels are in the same order, for convenience this is often abbreviated to include the normalisation constant and to show only the diagonal elements,

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N) = |\chi_i \chi_j \cdots \chi_n\rangle \tag{3.19}$$

Using this shorthand notation, Equation 3.15 should be written

$$|\cdots \chi_k \cdots \chi_m \cdots\rangle = -|\cdots \chi_m \cdots \chi_k \cdots\rangle \tag{3.20}$$

Constructing the Fock Operator

Having achieved a functional form for the wavefunction, the form of the Hamiltonian should next be considered. A one electron operator can be defined as

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_k \frac{Z_k}{r_{ik}} \tag{3.21}$$

where i denotes the electron and k the nuclei, r_{ik} the electron nuclear distance, and Z_k the nuclear mass. A two electron operator can be defined as

$$\hat{v}(i,j) = \frac{1}{r_{ij}} \tag{3.22}$$

where r_{ij} is the distance between electrons i and j. The electronic Hamiltonian can hence be written

$$\hat{H}_e = \sum_i \hat{h}(i) + \sum_{i<j} \hat{v}(i,j) + V_{kk} \tag{3.23}$$

where V_{kk} is considered as a constant (BOA) so in the context of electronic structure calculations is calculated as a classical term, added at the end of any calculations, and can be ignored at this point.

As a functional form of the Hamiltonian and wavefunction have been obtained, by assuming the Slater determinant notation in Equation 3.19, it can then be written

$$E_e = \langle \Psi | \hat{H}_e | \Psi \rangle \tag{3.24}$$

Applying the variational principle, which states that the energy will always be greater than or equal to the true energy, or

$$\frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geqslant E \tag{3.25}$$

it can be seen that by varying the parameters of each trial wavefunction so as to minimise the eigenvalue E the wavefunction will itself be a better approximation to the total wavefunction within a given functional space. Consequently the minimum energy configuration of orbitals is correct as it minimises the total electronic energy E_e. This is commonly referred to as the Rayleigh ratio.

Assuming, as stated previously, that the Hamiltonian is separable, the expression in Equation 3.24 for a two electron system, using Equations 3.21 and 3.22 can be written

$$\begin{aligned}\hat{H} &= \left(-\frac{1}{2}\nabla_1^2 - \sum_k \frac{Z_k}{r_{1k}}\right) + \left(-\frac{1}{2}\nabla_2^2 - \sum_k \frac{Z_k}{r_{2k}}\right) + \frac{1}{r_{12}} \\ &= \hat{h}(1) + \hat{h}(2) + \frac{1}{r_{12}} \\ &= \hat{O}_1 + \hat{O}_2\end{aligned} \tag{3.26}$$

where the Hamiltonian has been separated into the one and two electron parts as

$$\begin{aligned}\hat{O}_1 &= \hat{h}(1) + \hat{h}(2) \\ \hat{O}_2 &= r_{12}^{-1}\end{aligned} \tag{3.27}$$

In this expression, $\hat{h}(1)$ and $\hat{h}(2)$ are the *core Hamiltonians* for electron one and two respectively, describing both the kinetic and potential energy experienced by each (separately) in the field of the nuclei, or *core*.

If the one electron expression, $\hat{O}_1$, is evaluated in integral notation, using the expression for the wavefunction in Equation 3.14, and assuming the orthonormality of spin orbitals

$$\begin{aligned}\left\langle \Psi_0 \left| \hat{h}(1) \right| \Psi_0 \right\rangle &= \int d\mathbf{x}_1 d\mathbf{x}_2 \left[2^{-1/2}(\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2))\right]^* \\ &\quad \times \ \hat{h}(\mathbf{r}_1) \left[2^{-1/2}(\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) - \chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2))\right] \\ &= \frac{1}{2}\int d\mathbf{x}_1 d\mathbf{x}_2 \Big\{ \chi_1^*(\mathbf{x}_1)\chi_2^*(\mathbf{x}_2)\hat{h}(\mathbf{r}_1)\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) \\ &\qquad + \ \chi_2^*(\mathbf{x}_1)\chi_1^*(\mathbf{x}_2)\hat{h}(\mathbf{r}_1)\chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2) \\ &\qquad - \ \chi_1^*(\mathbf{x}_1)\chi_2^*(\mathbf{x}_2)\hat{h}(\mathbf{r}_1)\chi_2(\mathbf{x}_1)\chi_1(\mathbf{x}_2) \\ &\qquad - \ \chi_2^*(\mathbf{x}_1)\chi_1^*(\mathbf{x}_2)\hat{h}(\mathbf{r}_1)\chi_1(\mathbf{x}_1)\chi_2(\mathbf{x}_2) \Big\} \\ &= \frac{1}{2}\int d\mathbf{x}_1 \left\{ \chi_1^*(\mathbf{x}_1)\hat{h}(\mathbf{r}_1)\chi_1(\mathbf{x}_1) + \chi_2^*(\mathbf{x}_1)\hat{h}(\mathbf{r}_1)\chi_2(\mathbf{x}_1) \right\}\end{aligned} \tag{3.28}$$

By the same process it can be shown that

$$\left\langle \Psi_0 \left| \hat{h}(1) \right| \Psi_0 \right\rangle = \left\langle \Psi_0 \left| \hat{h}(2) \right| \Psi_0 \right\rangle \tag{3.29}$$

so it can hence be written

$$\left\langle \Psi_0 \left| \hat{O}_1 \right| \Psi_0 \right\rangle = \int d\mathbf{x}_1 \chi_1^*(\mathbf{x}_1) \hat{h}(\mathbf{r}_1) \chi_1(\mathbf{x}_1) + \int d\mathbf{x}_1 \chi_2^*(\mathbf{x}_1) \hat{h}(\mathbf{r}_1) \chi_2(\mathbf{x}_1) \tag{3.30}$$

Assuming the notation

$$\left\langle i \left| \hat{h} \right| i \right\rangle = \left\langle \chi_i \left| \hat{h} \right| \chi_j \right\rangle = \int d\mathbf{x}_1 \chi_i^*(\mathbf{x}_1) \hat{h}(\mathbf{r}_1) \chi_j(\mathbf{x}_1) \tag{3.31}$$

the expression for the one electron operator for a two electron system, $\hat{O}_1$, can be written as

$$\left\langle \Psi_0 \left| \hat{O}_1 \right| \Psi_0 \right\rangle = \left\langle 1 \left| \hat{h} \right| 1 \right\rangle + \left\langle 2 \left| \hat{h} \right| 2 \right\rangle \tag{3.32}$$

In general terms, the one electron operator for a many electron system can be written

$$\left\langle \Psi_0 \left| \hat{h}(i) \right| \Psi_0 \right\rangle = \sum_i \left\langle i \left| \hat{h} \right| i \right\rangle \tag{3.33}$$

If the two electron operator is then evaluated in the same manner as in Scheme 3.28, the terms for $\hat{h}(\mathbf{r}_1)$ exchanged for r_{12}^{-1}, and since $r_{12} = r_{21}$, the integral for the two electron operator can be written

$$\begin{aligned} \left\langle \Psi_0 \left| \hat{O}_2 \right| \Psi_0 \right\rangle = & \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_1^*(\mathbf{x}_1) \chi_2^*(\mathbf{x}_2) r_{12}^{-1} \chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) \\ & - \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_1(\mathbf{x}_1) \chi_2(\mathbf{x}_2) r_{12}^{-1} \chi_2(\mathbf{x}_1) \chi_1(\mathbf{x}_2) \end{aligned} \tag{3.34}$$

Assuming the notation

$$\langle ij \mid lm \rangle = \langle \chi_i \chi_j \mid \chi_l \chi_m \rangle = \int d\mathbf{x}_1 d\mathbf{x}_2 \chi_i^*(\mathbf{x}_1) \chi_j^*(\mathbf{x}_2) r_{12}^{-1} \chi_l(\mathbf{x}_1) \chi_m(\mathbf{x}_2) \tag{3.35}$$

the expression for the two electron operator for a two electron system, $\hat{O}_2$, can be written as

$$\left\langle \Psi_0 \left| \hat{O}_2 \right| \Psi_0 \right\rangle = \langle 12 \mid 12\rangle + \langle 12 \mid 21\rangle \tag{3.36}$$

In general terms, the two electron operator for a many electron system can be written

$$\left\langle \Psi_0 \left| r_{12}^{-1} \right| \Psi_0 \right\rangle = \sum_i \sum_{j>i} \langle ij \mid ij\rangle - \langle ij \mid ji\rangle \tag{3.37}$$

With Equations 3.33 and 3.37 an expression for the Slater determinant energy of a system can be given as

$$E_{SD} = \sum_i \langle i|\hat{h}|i\rangle + \sum_i \sum_{j>i} \langle ij \mid ij\rangle - \langle ij \mid ji\rangle \tag{3.38}$$

It is appropriate here, for completeness, to revisit the use of spin orbitals. If spacial orbitals are used a more intuitive representation of an atomic orbital can be presented. Revisiting the definition of the spin orbital, Equation 3.2, and revising

$$\begin{aligned} \chi_1(\mathbf{x}) = \psi_1(\mathbf{x}) &= \psi_1(\mathbf{r})\alpha(\omega) \\ \chi_2(\mathbf{x}) = \bar{\psi}_1(\mathbf{x}) &= \psi_1(\mathbf{r})\beta(\omega) \end{aligned} \tag{3.39}$$

In this notation it can be seen that every spacial orbital ψ_i is occupied by electrons of spin α and β, thus the Aufbau Principle is followed. Using this notation for an N electron, closed-shell system the Slater determinant form of the wavefunction can be written

$$\begin{aligned} |\Psi_0\rangle &= |\chi_1\chi_2\chi_3\chi_4\cdots\chi_{N-1}\chi_N\rangle \\ &= |\psi_1\bar{\psi}_1\psi_2\bar{\psi}_2\cdots\psi_{N/2}\bar{\psi}_{N/2}\rangle \end{aligned} \tag{3.40}$$

In order to recast Equation 3.38 in terms of spacial orbitals the definition of the one and two electron terms must be considered. By inspection of the new definition of the spacial orbital (Equation 3.39) it can easily be seen here

$$\sum_i^N \chi_i = \sum_i^{N/2} \psi_i + \sum_i^{N/2} \bar{\psi}_i \tag{3.41}$$

that the sum for the spin orbitals is over the number, N, of electrons. In the case of spacial orbitals, however, the terms accounting for the spin up (α) and down (β) are explicitly stated separately, and hence the number of electrons covered by each is $N/2$. As a consequence, the one electron operator becomes

$$\sum_{i}^{N}\langle i|\hat{h}|i\rangle = \sum_{i}^{N/2}\langle i|\hat{h}|i\rangle + \sum_{i}^{N/2}\langle \bar{i}|\hat{h}|\bar{i}\rangle = 2\sum_{i}^{N/2}\langle \psi_i|\hat{h}|\psi_i\rangle \tag{3.42}$$

and the two electron operator becomes

$$\begin{aligned}\sum_{i}^{N}\sum_{j>i}^{N}\langle ij|ij\rangle - \langle ij|ji\rangle = &\sum_{i}^{N/2}\sum_{j>i}^{N/2}(\langle ij|ij\rangle - \langle ij|ji\rangle)\\ &+\sum_{i}^{N/2}\sum_{j>i}^{N/2}(\langle i\bar{j}|i\bar{j}\rangle - \langle i\bar{j}|\bar{j}i\rangle)\\ &+\sum_{i}^{N/2}\sum_{j>i}^{N/2}(\langle \bar{i}j|\bar{i}j\rangle - \langle \bar{i}j|j\bar{i}\rangle)\\ &+\sum_{i}^{N/2}\sum_{j>i}^{N/2}(\langle \bar{i}\bar{j}|\bar{i}\bar{j}\rangle - \langle \bar{i}\bar{j}|\bar{j}\bar{i}\rangle)\\ =&\sum_{i}^{N/2}\sum_{j}^{N/2}2\langle \psi_i\psi_j|\psi_i\psi_j\rangle - \langle \psi_i\psi_j|\psi_j\psi_i\rangle\end{aligned} \tag{3.43}$$

Using the abbreviated notation as used in Equation 3.38 results in the energy of the closed shell ground state, using spacial orbitals, expressed as

$$E_0 = 2\sum_{i}^{N/2}\langle i|\hat{h}|i\rangle + \sum_{ij}^{N/2}2\langle ij|ij\rangle - \langle ij|ji\rangle \tag{3.44}$$

If the notation as in Equation 3.35 is used, the first term of the two electron operator can presented as

$$\begin{aligned}\langle ij|ij\rangle &= \int d\mathbf{r}_1 d\mathbf{r}_2 |\psi_i(\mathbf{r}_1)|^2\, r_{12}^{-1}\, |\psi_j(\mathbf{r}_2)|^2\\ &= J_{ij}\end{aligned} \tag{3.45}$$

This two electron integral is a classical coulombic repulsion between the charge clouds $|\psi_i(\mathbf{r}_1)|^2$ and $|\psi_j(\mathbf{r}_2|^2$. The expression J_{ij} is hence called a Coulomb operator. If the second term of the two electron operator is considered in the same way

$$\begin{aligned}\langle ij|ji\rangle &= \int d\mathbf{r}_1 d\mathbf{r}_2 \psi_i^*(\mathbf{r}_1)\psi_j(\mathbf{r}_1)\, r_{12}^{-1}\, \psi_j^*(\mathbf{r}_2)\psi_i(\mathbf{r}_2) \\ &= K_{ij}\end{aligned} \tag{3.46}$$

the resulting expression, K_{ij}, is known as the exchange operator. This exchange operator represents a modification of the Coulomb operator, taking into account the effects of spin correlation.

Having defined the expressions for energy of the system for both spin and spacial orbitals, the Fock operator can be written for closed shell spacial orbitals

$$\hat{f}(\mathbf{r}_1) = \hat{h}(\mathbf{r}_1) + \sum_{j}^{N/2} 2J_j - K_j \tag{3.47}$$

and for spin orbitals

$$\hat{f}(\mathbf{x}_1) = \hat{h}(\mathbf{x}_1) + \sum_{i}^{N} \mathscr{J}_i - \mathscr{K}_i \tag{3.48}$$

where $\mathscr{J}$ and $\mathscr{K}$ are spin analogues for J and K. The change in subscript between the spacial and spin expressions of the Fock operator are merely to emphasise that these expressions are only approximately, not exactly, equivalent. The Fock operator, in general terms, is an effective one electron operator which can be expressed as

$$\hat{f}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{k=1}^{M} \frac{Z_k}{r_{ik}} + \nu^{HF}(i) \tag{3.49}$$

The term $\nu^{HF}(i)$ is called the Hartree-Fock potential. As asserted previously, this potential describes the field of the other electrons acting upon the ith electron. As

this field is dependent on the orbitals of the other electrons, the Fock operator is dependent upon its eigenfunctions.

The Hartree Fock equation can similarly be defined for both spacial orbitals

$$\hat{f}(\mathbf{r}_1)\psi_j(\mathbf{r}_1) = \varepsilon_j\psi_j(\mathbf{r}_1) \tag{3.50}$$

and for spin orbitals

$$\hat{f}(\mathbf{x}_1)\chi_i(\mathbf{x}_1) = \varepsilon_i\chi_i(\mathbf{x}_1) \tag{3.51}$$

where ε_j and ε_i are expressed as the singular values of Equations 3.44 and 3.38 respectively, the j and i subscripts used for clarity. As the Fock operator is dependent on its eigenfunctions, the Hartree-Fock equation is hence non-linear and must be solved iteratively. This iterative method is the process introduced at the beginning of this Section (3.2.1), the self-consistent field (SCF) method.

3.2.2 Basis Sets

At this point, the Hartree-Fock equation can be solved numerically as it is relatively straightforward to describe atoms given their spherical symmetry. However, the application of the Hartree-Fock SCF method is highly numerically complex in the calculation of the molecular orbitals described by Mulliken and Hund. In 1951 C. C. J. Roothaan [84] and G. G. Hall [85] independently developed a procedure by which the molecular orbitals of a system are constructed as a linear combination of atomic orbitals, an approach now abbreviated to the LCAO.

Linear Combination of Atomic Orbitals

If the Hartree-Fock equation, in the terms of Equation 3.50, is considered, the spacial orbitals, ψ_i, can be thought of as molecular orbitals each of which can be expressed as a linear combination of N_o number of atomic orbitals, ϕ_o,

$$\psi_i = \sum_{o=1}^{N_o} c_{oi}\phi_o \qquad\qquad i = 1, 2, \cdots, N_o \tag{3.52}$$

where the coefficient, c_{oi} is, as yet, unknown. It can be stated that the atomic orbitals, ϕ_o, are basis functions, the sum over which, as described above, is a (basis) set for the description of the molecular orbital ψ_i. When the number of terms in the expansion $N_o \longrightarrow \infty$, the definition of the molecular orbital is exact. As this is not feasible, a set of the ϕ_o basis functions, must be chosen to give a reasonably accurate expansion for the exact description of the LCAO basis set, and hence the set of ψ_i molecular orbitals. The truncated expansion is exact only in the space spanned by the basis functions.

The concept of a function, described by a set of basis functions, which is itself described by a set of basis functions (and so on... until a final basis set of functions is reached) is an important one, and is reiterated in various theoretical methods, including those described later.

Using Equation 3.52, the ground state MO of the H_2 molecule can be represented as a linear combination of the individual hydrogenic 1s atomic orbitals, and hence from a set of N_o basis functions, N_o number of linearly dependent spacial wavefunctions are obtained. Consequently the difficulty in the calculation of the wavefunctions has been reduced to the calculation of the coefficients c_{oi}.

By substituting the linear expansion of Equation 3.52 into Equation 3.50, the Hartree-Fock equation becomes a integro-differential equation

$$\hat{f}(\mathbf{r}_1)\sum_{o=1}^{N_o} c_{oi}\phi_o(\mathbf{r}_1) = \varepsilon_i \sum_{o=1}^{N_o} c_{oi}\phi_o(\mathbf{r}_1) \tag{3.53}$$

Multiplying on the left by $\phi_{o'}(\mathbf{r}_1)$ and integrating over all space

$$\sum_{o=1}^{N_o} c_{oi} \int d\mathbf{r}_1 \phi^*_{o'}(\mathbf{r}_1)\hat{f}(\mathbf{r}_1)\phi_o(\mathbf{r}_1) = \varepsilon_i \sum_{o=1}^{N_o} c_{oi} \int d\mathbf{r}_1 \phi^*_{o'}(\mathbf{r}_1)\phi_o(\mathbf{r}_1) \tag{3.54}$$

turns the integro-differential equation into a matrix equation. By defining the overlap matrix, **S**, with the elements

$$S_{o'o} = \int d\mathbf{r}_1 \phi^*_{o'}(\mathbf{r}_1)\phi_o(\mathbf{r}_1) \tag{3.55}$$

and the Fock matrix, **F**, with the elements

$$F_{o'o} = \int d\mathbf{r}_1 \phi^*_{o'}(\mathbf{r}_1)\hat{f}(\mathbf{r}_1)\phi_o(\mathbf{r}_1) \tag{3.56}$$

Equation 3.54 becomes

$$\sum_{o=1}^{N_o} F_{o'o}c_{oi} = \varepsilon_i \sum_{o=1}^{N_o} S_{o'o}c_{oi} \tag{3.57}$$

which is one set of N_o simultaneous equations, a set for each value of o', known as the Roothaan, or Roothaan-Hall, equations. Although it is assumed that the basis functions, ϕ_o, are normalised and linearly independent, they are not necessarily orthogonal to each other. As a result, the overlap matrix **S** with values of the magnitude $0 \leqslant |S_{o'o}| \leqslant 1$ is unity in the diagonal elements and the off-diagonal elements are values less than 1. The Fock operator, $\hat{f}(\mathbf{r}_1)$, is a one electron operator, hence any set of one electron functions defines the operator in a matrix representation. Both the Fock matrix and the overlap matrix are $N_o \times N_o$ Hermitian, although often are real symmetric, matrices.

The full set of equations can, in compact notation, be written as a single matrix equation

$$\mathbf{Fc} = \boldsymbol{\varepsilon}\mathbf{Sc} \tag{3.58}$$

where **c** is a $N_o \times N_o$ square matrix of the expansion coefficients, c_{oi}, and $\boldsymbol{\varepsilon}$ is a diagonal matrix of the energies of the orbitals ε_i of the form

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_1 & & & \\ & \varepsilon_2 & & \mathbf{0} \\ \mathbf{0} & & \ddots & \\ & & & \varepsilon_{N_0} \end{pmatrix} \tag{3.59}$$

The Roothaan equations are a set of simultaneous equations for the coefficients which only have non-trivial solutions if the secular equation

$$|\mathbf{F} - \boldsymbol{\varepsilon}\mathbf{S}| = 0 \tag{3.60}$$

is satisfied. The solution to this equation cannot directly be found as the elements of the Fock matrix involve integrals over the Coulomb and exchange operators, which themselves are dependent upon the spacial wavefunction. If the elements of the Fock matrix are evaluated, returning to Equation 3.47, expanding the Coulomb and exchange terms using Equation 3.56, implementing the LCAO as Equation 3.52, and finally simplifying the notation as per Equation 3.35

$$\left.\begin{aligned} F_{o'o} = & \int d\mathbf{r}_1 \phi^*_{o'}(\mathbf{r}_1)\hat{h}(\mathbf{r}_1)\phi_o(\mathbf{r}_1) \\ & + 2\sum_j \int d\mathbf{r}_1 d\mathbf{r}_2 \phi^*_{o'}(\mathbf{r}_1)\psi^*_j(\mathbf{r}_2) r_{12}^{-1} \phi_o(\mathbf{r}_1)\psi_j(\mathbf{r}_2) \\ & - \sum_j \int d\mathbf{r}_1 d\mathbf{r}_2 \phi^*_{o'}(\mathbf{r}_1)\psi^*_j(\mathbf{r}_2) r_{12}^{-1} \psi_j(\mathbf{r}_1)\phi_o(\mathbf{r}_2) \\ = & \, h_{o'o} + 2\sum_{i,q,r} c^*_{qi}c_{ri} \int d\mathbf{r}_1 d\mathbf{r}_2 \phi^*_{o'}(\mathbf{r}_1)\phi^*_q(\mathbf{r}_2) r_{12}^{-1} \phi_o(\mathbf{r}_1)\phi_r(\mathbf{r}_2) \\ & - \sum_{i,q,r} c^*_{qi}c_{ri} \int d\mathbf{r}_1 d\mathbf{r}_2 \phi^*_{o'}(\mathbf{r}_1)\phi^*_q(\mathbf{r}_2) r_{12}^{-1} \phi_r(\mathbf{r}_2)\phi_o(\mathbf{r}_1) \\ = & \, h_{o'o} + \sum_{i,q,r} c^*_{qi}c_{ri} \left(2\langle o'q|or\rangle - \langle o'q|ro\,\rangle\right) \end{aligned}\right\} \tag{3.61}$$

yields a form for the Fock matrix elements no longer expressed in terms of a spacial wavefunction but upon the basis functions defined within confines of the LCAO.

It is useful at this point to introduce the matrix $\mathbf{P}$ to account for the expansion coefficients, the matrix elements of which are given by

$$P_{qr} = \sum_i c_{qi}^* c_{ri} \tag{3.62}$$

In the overlap region of ϕ_q and ϕ_r, as the matrix elements are a sum over all ψ_i occupied molecular orbitals, the elements P_{qr} are interpreted as the total electron density and are hence called the density matrix elements. In the case of $q \neq r$, the density matrix element P_{qr} is the bond order between q and r whereas in the case of $q = r$, the density matrix element P_{qq} is the electron density on the atom q.

Upon substitution of the expression for the density matrix elements into the final term of Scheme 3.61, the Fock matrix elements can be written as

$$F_{o'o} = h_{o'o} + \sum_{q,r} P_{qr} \left(2\langle o'q|or\rangle - \langle o'q|ro\rangle\right) \tag{3.63}$$

When this expression of the Fock matrix elements is used to carry out calculations it can be seen that the one electron matrix elements, $h_{o'o}$, are unchanged in each iteration of the calculation, meaning they only need to be calculated once. However, as the number of two electron integrals to be evaluated scales as N_o^4, and as the density matrix elements, dependent on the expansion coefficients, must be re-evaluated at each iteration of the calculation, it can easily be seen that there is a significant challenge in carrying out a Hartree-Fock SCF calculation. Fortunately, due to symmetry, some integrals may be identically zero, some non-zero integrals may be equal and, in the case of widely spaced nuclei, some integrals may be negligibly small, thus alleviating some of this challenge.

It should be noted that the Fock operator is limited in a chemical sense by the fact that it is ostensibly one electron in nature and aside from the exchange operator, which accounts for Fermi correlation or the correlation effect unique to electrons of the same spin, all other correlation effects are as yet unaccounted. A further discussion as to the nature of, and methods developed to account for, these other correlation effects will be discussed in the following section (Section 3.2.3).

Basis Sets

The development of the LCAO, and the Roothaan equations, enabled much larger systems to be considered in calculations. However, the bottleneck with respect to the size of the basis set, due to the quartic (N_o^4) scaling behaviour, presented a major issue that needed to be overcome. Consequently two schools of thinking developed in order to make further progress.

The '*semiempirical*' MO theories identify inherent chemical flaws in the Hartree-Fock equations whilst recognising the power of the method. The philosophy of these methods involve the introduction of approximations in order to simplify the solutions to the Hartree-Fock equations with the aim of improving accuracy. Typically this is carried out by the use of parameterizations, focusing on the reproduction of experimental properties, hence the name *semiempirical*. In the interest of brevity, the details of the development of 'semiempirical' methods will not be discussed further. More information on these methods can be found in standard references such as Refs. [86, 87].

The '*ab initio*' Hartree-Fock theory focuses on the fact the limitation of the accuracy of the Hartree-Fock method is in the number of terms used to form the basis set. As stated previously, the LCAO is correct, except for the correlation energy, if the number of terms in the expansion $N_o \longrightarrow \infty$, in other words, an infinite basis set. Consequently, anticipating the development of computational technology, no further approximations are made, and the size, definition and complexity of the basis set is increased, limited only by the computational capacity. The *ab initio* method also uses experimental data, but only as a guide to the selection of the computational model as opposed to direct parameterization. It is the focus of this philosophy that will be discussed.

In the context of the aforementioned N_o^4 basis function scaling issue, *ab initio* methods attempt to make a balance between minimising the number of basis

functions used, whilst choosing a functional form of the basis set such that the Hartree-Fock equations can be evaluated in a computationally efficient manner. The basis functions must also be chosen to a large amplitude in regions the electron probability density is large, and vice versa, in order to be useful in a chemical sense.

As the theory presented thus far utilises one electron approximations, no definitive analytical form can be given for atomic orbitals. Notwithstanding this primitive definition, approximate atomic orbitals can be constructed to model the actual wavefunction. One set of these orbitals, Slater type orbitals (STOs), are defined by constructing an orbital with quantum numbers n, l, and m_l, belonging to an atomic nucleus in atom centred polar coordinates as

$$\vartheta_{\zeta n l m_l}(r, \theta, \varphi) = N r^{n-1} e^{-\zeta r} Y_{l m_l}(\theta, \varphi) \tag{3.64}$$

where N is a normalisation constant, ζ is the orbital exponent the size of which is dependent on the atomic number and $Y_{l m_l}$ is a spherical harmonic which is related to the angular momentum. The conventional physical spherical coordinate notation is used here where r is the radial distance of the electron from the nucleus, θ is the polar angle or the angle from the z-axis, and φ is the azimuthal angle or the anticlockwise projected vector from the x-axis onto the xy plane.

STOs present a number of attractive features. In the STO representation the orbitals present an exponential decay with increasing radial distance, the angular component is hydrogenic, and the 1s-orbital has a cusp at the nucleus. From a practical perspective the most attractive feature of the STO representation is that overlap integrals between two STOs, as a function of interatomic distance, are easily calculated.

However, difficulties arise in the use of STOs as orbitals with differing principal quantum number values (n) but the same angular (l) and magnetic (m_l) quantum numbers are non-orthogonal. It should also be noted that $n > 1$ s-orbitals have

zero amplitude at the nucleus. The most significant limitation of STOs is that there is no known analytical solution to the general four electron index that appears in the Fock matrix elements (Scheme 3.61 as in Equation 3.35).

In 1950 S. Francis Boys [88] presented a new method for the representation of atomic orbitals, which proved to be a major step towards computationally feasible calculations. In the simplest terms, this newer type of orbitals, Gaussian type orbitals (GTOs), replaced the radial decay term, e^{-r}, of the STO representation with e^{-r^2} resulting in an analytical solution to the four index integral. In terms of the same coordinate system as presented in the functional form of the STOs, spherical polar coordinates, the GTOs have the form

$$\vartheta_{\zeta nlm_l}(r,\theta,\varphi) = Nr^{2n-2-l}e^{-\zeta r^2}Y_{lm}(\theta,\varphi) \tag{3.65}$$

More commonly, the functional form of GTOs are given in atom centred Cartesian coordinates as

$$\vartheta_{\zeta l_x l_y l_z}(\mathbf{r}) = Nx^{l_x}y^{l_y}z^{l_z}e^{-\zeta r^2} \tag{3.66}$$

where, as previously, N is a normalisation constant and the values of l determine the angular component, hence the sum of the components l_x, l_y and l_z determine the type of orbital being represented. $l_x + l_y + l_z$ is equal to 0 for an s-orbital, 1 for a p-orbital, and 2 for a d-orbital.

However, the transition between spherical polar coordinates to Cartesian coordinates is not as straightforward as it may appear. When a GTO defines a d-type orbital, in spherical coordinates there are five components (Y_{22}, Y_{21}, Y_{20}, Y_{2-1} and Y_{2-2}) while in Cartesian coordinates there appear to be six components (x^2, y^2, z^2, xy, xz and yz). Consequently the six Cartesian components are transformed into the spherical d-functions, with one additional s-function of the form $x^2+y^2+z^2$. In the same way, the ten Cartesian f-functions are transformed into the seven spher-

ical functions with one additional set of p-functions. From a practical perspective this transformation yields a saving in computational effort when many d-functions or functions with higher angular momentum are to be calculated. Additionally, the problem of linear dependence in larger basis sets is reduced when only the spherical components are used.

The major advantage of the use of GTOs lies in the mathematical behaviour of Gaussian functions. The product of two Gaussians, the centres of which lie a distance apart, is equivalent to a single Gaussian function centred at a point between the two, ergo the two electron integrals for three or four atomic centres can be reduced to integrals over two centres. This presents a significant reduction in computational effort.

Nevertheless, the use of GTOs has two unfortunate drawbacks. As stated previously, one of the advantages in the use of STOs is that the nuclear cusp is well defined. The gradient of the Gaussian function is zero at the nucleus, resulting in near nucleus behaviour being poorly defined. Additionally, the Gaussian function falls off rapidly at distances far from the nucleus. Consequently the computational advantage in the description of GTOs is almost lost as a much larger basis set of GTOs is required to obtain comparable accuracy to the use of STOs.

The denouement of this issue is achieved by the use of several GTOs grouped together to form what are known as contracted Gaussian functions given by

$$\phi_o = \sum_i d_{oi}\vartheta_i \tag{3.67}$$

In this form, each contracted Gaussian, ϕ_o, is a linear combination of contraction coefficients, d_{oi}, multiplied by so called primitive Gaussians, ϑ. The parameters of ϑ and d are constants during the calculation. If this is then fed into the LCAO equation (Equation 3.52) the molecular orbitals are now represented by a linear combination of contracted Gaussians. When contracted Gaussians are used in

place of primitive Gaussians the number of unknown expansion coefficients, c_{om}, to be calculated in a Hartree-Fock calculation is reduced. If the choice of contracted Gaussians is good, the decrease in the number of coefficients to be calculated is not to the detriment of the accuracy of the calculation.

The simplest form of basis set uses n number of primitive Gaussians to construct the contracted Gaussians, the exponents of which are fitted by a least squares method so as to form an approximation to a set of STOs optimised in an atomic SCF calculation. The basis sets constructed in this manner are hence known as the STO-nG basis sets. It has been found that for $n > 3$ no significant improvement is gained in the resulting wavefunction and energies, in that the result is far from the Hartree-Fock limit. Consequently the STO-3G basis set, and due to its primitive representation of orbitals, is commonly referred to as a minimal basis set.

In the early 1980's Pople and coworkers [89] published a series of papers presenting the construction of a more complex series of basis sets, in general terms given as a k-$nlmG$ basis set. These basis sets, commonly referred to as split-valence basis sets, are constructed by representing orbitals that are involved in bonding, or valence, differently to the way in which inner-shell or core orbitals are defined. The justification for this is simply that the inner-shell orbitals are not directly involved in the bonding in a molecule and consequently contribute only to the energetics of the system. The k hence refers to the number of primitive Gaussians that are used to form a contracted Gaussian representing the inner-shell orbitals. If the number m is omitted, the valence orbitals are split into two parts, the inner part of the valence orbitals given as a contraction of n number of primitive Gaussians and the outer valence orbitals given as a contraction of l number of primitive Gaussians. This is referred to as a (valence-)double-ζ basis set. If the number m is included, the valence orbitals are split into three parts, referred to as triple split-valence,

where the outermost valence orbitals are given as a contraction of m number of primitive Gaussians. These basis sets are referred to as a (valence-)triple-ζ.

In addition these split-valence basis sets can be further extended to include more diffuse functions (+ for heavy atoms, ++ for heavy atoms and hydrogen) and polarisation functions (* for heavy atoms, ** for heavy atoms and hydrogen), so the general form can be given as k-nlm++G**.

In the simplest terms, if the row or period of the atom in question is denoted as γ, every $\gamma - 1(> 0)$s/p/d/...-orbital is each described by one contracted Gaussian function comprising k primitives. All γ-orbitals are described by two, in the case of nl, or three, in the case of nlm, contracted Gaussians, the first comprising n primitive Gaussians, the second of l primitives, etc. The polarisation function, $*$, includes six d-type primitive Gaussians for heavy atoms, and for $**$, in addition to the six d-types, three p-type primitives for each hydrogen. The diffuse function, +, adds a set of s- and p-type primitives for heavy atoms, and for ++, in addition to the sp-primitives, a set of s-type primitives for each hydrogen.

Between 1965 and 1980 [89] a different contraction scheme for the construction of basis sets was developed, generally named the Dunning basis sets. The Dunning, correlation consistent basis sets, in general notation, are termed as cc-p$V$$N$Z. The letters 'cc-p' stand for 'correlation-consistent polarised', while the VNZ stands for 'valence N zeta' where N can be D, for double(-ζ), T, for triple(-ζ), etc. up to 6. As indicated by the 'p', this type of basis set inherently includes the polarization functions depicted by the '$**$' notation for Pople basis sets. The N, as with the Pople basis sets, indicates that there is a split valence contraction scheme. Unlike the Pople basis sets, the contraction schemes are fixed where double-ζ has 6,1,1,1 for all s-type orbitals and 4,1 for p-type orbitals, and triple-ζ has 6,2,1,1,1 for s-type orbitals and 4,1,1 for p-type orbitals. The prefix 'aug' can be added before the basis set to indicate augmented, or diffuse, functions have been added. Due to

the definition of the contraction scheme, Dunning basis sets have somewhat more flexibility in the description of electronic structure than Pople basis sets, though are understandably computationally more expensive.

There exists a variety of other types of basis sets, including Karlsruhe, plane-wave, real-space, and many more. As only the Pople and Dunning basis sets are used in this work, further discussion of these other basis set types can be found in standard references [86] [87].

3.2.3 Beyond Hartree-Fock

In the case of an infinite basis-set, the Hartree-Fock calculated, minimum energy of a system will always be a higher energy than the exact solution. The difference in energy between the Hartree-Fock minimum and the true minimum is the correlation energy. The definition of the correlation energy is perhaps most easily explained by looking at the construction of Hartree-Fock theory.

It is important to see that the Fock operator is not a single operator but is a construction of a set of interdependent one electron operators, used to build one electron molecular orbitals. These molecular orbitals are used to populate the terms in the Slater determinant, and the Hartree-Fock wavefunction is formed. Due to the application of the variation principle, the resultant Hartree-Fock energy is the upper bound to the exact energy of the system. As the Hamiltonian operator returns the electronic energy for a many electron system, it can be seen that the difference between this and the Hartree-Fock energy must be due to behaviours as yet unaccounted by the one electron approximations. It is the many electron behaviours that are the correlation energy, the sign of which is negative when used as a correction to the Hartree-Fock energy.

As Hartree-Fock is a reasonable starting point for the calculation of the exact energy of a system, and by definition is a single determinant method, it is therefore reasonable to assume that the total wavefunction of a system can be represented

by a linear combination of multiple determinants

$$\Psi = c_o \Psi_{HF} + \sum_{i=1} c_i \Psi_i \tag{3.68}$$

The coefficients, c_i, determine the weight of the determinants, Ψ_i, and ensure normalisation. It is the selection and manipulation of these determinants that allow the correlation energy to be included. Consequently, electronic structure methods including varying degrees of correlation energy, or numbers of determinants, are termed post- or beyond-Hartree-Fock.

Electron correlation is somewhat difficult to define in general terms as it is dependent on the frame of reference, or how the electronic structure is termed. It is hence useful to describe electron correlation in terms of fundamental spin states, and in terms of molecular orbitals to, in essence, paint a picture of the nature of the problem. Due to the Pauli exclusion principle, it is not possible to have two electrons of the same spin in the same orbital, and hence intraorbital correlation is not possible. However, two electrons of the same spin in different orbitals (interorbital) are correlated and this type of same spin correlation is termed Fermi correlation. When electrons of differing spins are considered, it can be seen that their behaviour is correlated in both the intra- and interorbital cases. This type of, opposite spin, correlation is termed as Coulomb correlation. If these types of correlation are considered in terms of electron density surrounding a particular electron, the reduced probability of finding electron density in the immediately surrounding region is termed the "Fermi hole" for same spin, and "Coulomb hole" for opposite spin cases. The Fermi and Coulomb hole as terminology and concepts are more commonly used in DFT methods.

As an alternative to this description, electron correlation can also be distinguished by consideration of the more chemically intuitive orbitals. If the H_2 molecule is considered, at its dissociation limit the bonding and anti-bonding or-

bitals become degenerate. In a Hartree-Fock calculation the Slater determinant can only account for an electron configuration populating one of the orbitals, and the calculation continues to an optimisation of occupied orbitals. However, it can easily be seen that if the other orbital is populated, due to the degeneracy, the calculation will achieve the same result. Consequently, if this is taken into account by two determinants of the form of Equation 3.68, these determinants will have roughly equal weight. It can also be seen that if two states are near degeneracy, the weight of the higher state in a two determinant system will be significant in comparison to that of the lower state. This type of correlation, where a single determinant method is insufficient to describe degenerate or near degenerate states, is commonly known as *static* or *non-dynamical* correlation.

The type of calculations considered in electronic structure methods are, by definition, solving the time independent Schrödinger equation, so it can readily be seen why the account of these degeneracies are termed *static* correlation. *Dynamic* correlation can hence be described as the instantaneously correlated motion of an electron with all other electrons. Consequently, when using an expression for the many electron wavefunction as in Equation 3.68, though the contributions of the correlation in each of the $i \geqslant 1$ determinants may be small in comparison to the weight of the Hartree-Fock determinant, the linear combination of them may be large and hence significant.

As with the treatment of electronic structure, and the construction of basis sets, there are many ways by which electron correlation can be treated.

Full Configuration Interaction

Perhaps the most useful method for taking into account correlation effects, in terms of the concepts presented, is the full configuration interaction, or full CI method. Expanding the first few terms of the linear combination of Equation 3.68, and recasting

$$|\Psi\rangle = c_0|\Psi_{HF}\rangle + \sum_{ar} c_a^r|\Psi_a^r\rangle + \sum_{\substack{a<b\\r<s}} c_{ab}^{rs}|\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c\\r<s<t}} c_{abc}^{rst}|\Psi_{abc}^{rst}\rangle + \sum_{\substack{a<b<c<d\\r<s<t<u}} c_{abcd}^{rstu}|\Psi_{abcd}^{rstu}\rangle + \cdots \tag{3.69}$$

where a, b, c, and d indicate occupied orbitals, r, s, t, and u indicate virtual orbitals, Ψ_a^r is the singly excited determinant, Ψ_{ab}^{rs} the doubly excited, etc. gives the form of the full CI wavefunction. This can be given in more simplified notation,

$$|\Psi\rangle = c_0|\Psi_0\rangle + c_S|S\rangle + c_D|D\rangle + c_T|T\rangle + c_Q|Q\rangle + \cdots \tag{3.70}$$

where $|S\rangle$ represents the single excitation terms, $|D\rangle$ the double excitation terms, etc. Presenting this in matrix form, called the full CI matrix, some interesting properties can be seen.

$$\begin{array}{c} \\ |\Psi_0\rangle \\ |S\rangle \\ |D\rangle \\ |T\rangle \\ |Q\rangle \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} |\Psi_0\rangle & |S\rangle & |D\rangle & |T\rangle & |Q\rangle & \cdots \end{array} \\ \begin{bmatrix} \langle\Psi_0|\hat{H}|\Psi_0\rangle & 0 & \langle\Psi_0|\hat{H}|D\rangle & 0 & 0 & \cdots \\ & \langle S|\hat{H}|S\rangle & \langle S|\hat{H}|D\rangle & \langle S|\hat{H}|T\rangle & 0 & \cdots \\ & & \langle D|\hat{H}|D\rangle & \langle D|\hat{H}|T\rangle & \langle D|\hat{H}|Q\rangle & \cdots \\ & & & \langle T|\hat{H}|T\rangle & \langle T|\hat{H}|Q\rangle & \cdots \\ & & & & \langle Q|\hat{H}|Q\rangle & \cdots \\ & & & & & \ddots \end{bmatrix} \end{array} \tag{3.71}$$

Here $\langle S|\hat{H}|T\rangle \leftrightarrow \langle\Psi_a^r|\hat{H}|\Psi_{cde}^{tuv}\rangle$, $\langle D|\hat{H}|D\rangle \leftrightarrow \langle\Psi_{ab}^{rs}|\hat{H}|\Psi_{de}^{uv}\rangle$ etc., and for simplicity only the upper triangle is included. When full CI is presented in this form, the interactions of all configurations of the ground to n-tuply excited state determinants can be seen hence the name (full) *configuration interaction*.

The first significant feature of the full CI matrix is that there is zero coupling between the Hartree-Fock ground state and the singly excited determinant. This is as a consequence of Brillouin's theorem which states that singly excited determinants will not directly interact with a reference Hartree-Fock determinant. This can be justified by a simple proof. If only the leading correction in Equation 3.69 is considered, a matrix eigenvalue problem can be presented as

$$\begin{pmatrix} \langle\Psi_0|\hat{H}|\Psi_0\rangle & \langle\Psi_0|\hat{H}|\Psi_a^r\rangle \\ \langle\Psi_a^r|\hat{H}|\Psi_0\rangle & \langle\Psi_a^r|\hat{H}|\Psi_a^r\rangle \end{pmatrix} \begin{pmatrix} c_0 \\ c_a^r \end{pmatrix} = E_0 \begin{pmatrix} c_0 \\ c_a^r \end{pmatrix} \tag{3.72}$$

It is clear, in this matrix, that the mixing of states occurs in the off-diagonal elements hence evaluation of these terms should be carried out. As the matrix elements of the Fock operator (Equation 3.63) can be written, with spin orbitals, as

$$\langle\chi_i|f|\chi_j\rangle = \langle i|h|j\rangle + \sum_b \langle ib|jb\rangle - \langle ib|bj\rangle \tag{3.73}$$

and the off-diagonal matrix elements of Equation 3.72 can also be written as

$$\langle\Psi_0|\hat{H}|\Psi_a^r\rangle = \langle a|h|r\rangle + \sum_b \langle ab|rb\rangle - \langle ab|br\rangle \tag{3.74}$$

it is therefore possible to write

$$\langle\Psi_0|\hat{H}|\Psi_a^r\rangle = \langle\chi_a|f|\chi_r\rangle \tag{3.75}$$

A requirement of the off-diagonal elements of the Fock matrix is that they should be zero except where the index values are equal, i.e. $\langle\chi_i|f|\chi_j\rangle = 0, (i \neq j)$. Consequently, in the case of the mixing of Hartree-Fock and singly excited determinants, as a is an occupied orbital and r is a virtual orbital, a can never equal r, i.e. $a \in \{occ\}$, $r \in \{virt\}$: $\{occ\} \cap \{virt\} = 0 \therefore a \neq r$, hence the mixing is always 0. Having said this, the singly excited determinants can, however, indirectly interact with the Hartree-Fock determinant through their mutual mixing with the doubly excited determinant.

As the matrix elements of the Hamiltonian between any Slater determinants that differ by more than 2 spin orbitals are zero, the Hartree-Fock determinant does not mix with triply and quadruply excited determinants. Singly excited determinants, similarly, do not mix with quadruply excited determinants. This can be shown by a similar method to the proof of Brillouin's theorem.

It should also be noted that although, for example, doubly and quadruply excited determinants mix, the number of terms with a non-zero value are relatively small. In other words, the block diagonal matrix representing this interaction is sparsely populated. This is due to the fact that in order for $\langle \Psi_{ab}^{rs} | \hat{H} | \Psi_{cdef}^{tuvw} \rangle$ to have a non-zero value, a and b must be included in the set of c, d, e and f, and at the same time r and s must also be included in the set of t, u, v and w, i.e. $\langle \Psi_{ab}^{rs} | \hat{H} | \Psi_{cdef}^{tuvw} \rangle \neq 0 \iff (a, b \in \{c, d, e, f\}) \wedge (r, s \in \{t, u, v, w\})$.

In order to be able to obtain a value for the correlation energy in full CI a technique called *intermediate normalisation* can be used. It is reasonable to assume that for a ground state calculation, if the Hartree-Fock determinant is considered an acceptable approximation to the exact wavefunction, the coefficient leading the Hartree-Fock determinant will be larger in comparison to the coefficients of the other determinants. Consequently, Equation 3.69 can be rewritten

$$|\Psi\rangle = |\Psi_{HF}\rangle + \sum_{ct} c_c^t |\Psi_c^t\rangle + \sum_{\substack{c<d \\ t<u}} c_{cd}^{tu} |\Psi_{cd}^{tu}\rangle + \sum_{\substack{c<d<e \\ t<u<v}} c_{cde}^{tuv} |\Psi_{cde}^{tuv}\rangle + \sum_{\substack{c<d<e<f \\ t<u<v<w}} c_{cdef}^{tuvw} |\Psi_{cdef}^{tuvw}\rangle + \cdots \tag{3.76}$$

The wavefunction here is not normalised. In this case a property exists where the overlap between the Hartree-Fock and exact wavefunction is equal to 1. This allows Equation 3.69 to be normalised if each term in the expansion is multiplied by a constant.

If the TISE is written in the Slater determinant form

$$\hat{H}|\Psi\rangle = E_0|\Psi\rangle \tag{3.77}$$

and $E_{HF}|\Psi\rangle$ is subtracted from both sides

$$(\hat{H} - E_{HF})|\Psi\rangle = (E_0 - E_{HF})|\Psi\rangle = E_{corr}|\Psi\rangle \tag{3.78}$$

where E_{corr} is the correlation energy, then multiplied by $\langle\Psi_{HF}|$

$$\langle\Psi_{HF}|\hat{H} - E_{HF}|\Psi\rangle = E_{corr}\langle\Psi_{HF}|\Psi\rangle = E_{corr} \tag{3.79}$$

the reason for the use of the intermediate normalisation becomes clear.

By combining Equation 3.76, the features as described in the full CI matrix, and Equation 3.79

$$\begin{aligned}\langle\Psi_{HF}|\hat{H} - E_{HF}|\Psi\rangle &= \langle\Psi_{HF}|\hat{H} - E_{HF}\left(|\Psi_{HF}\rangle + \sum_{ct} c_c^t|\Psi_c^t\rangle + \sum_{\substack{c<d\\t<u}} c_{cd}^{tu}|\Psi_{cd}^{tu}\rangle + \cdots\right)\\ &= \sum_{\substack{a<b\\r<s}} c_{ab}^{rs}\langle\Psi_{HF}|\hat{H}|\Psi_{ab}^{rs}\rangle = E_{corr}\end{aligned} \tag{3.80}$$

obtains a value for the correlation energy dependent only on the coefficients of the double excitations. If the same process (Equations 3.78 - 3.80) is then repeated, substituting the singly excited determinant in place of the Hartree-Fock determinant

$$\begin{aligned}\langle\Psi_{HF}|\hat{H} - E_{HF}|\Psi\rangle &= \sum_{ct} c_c^t\langle\Psi_a^r|\hat{H} - E_{HF}|\Psi_c^t + \sum_{\substack{c<d\\t<u}} c_{cd}^{tu}\langle\Psi_a^r|\Psi_{cd}^{tu}\rangle + \sum_{\substack{c<d\\t<u}} c_{acd}^{rtu}\langle\Psi_a^r|\Psi_{acd}^{rtu}\rangle\\ &= E_{corr}c_a^r\end{aligned} \tag{3.81}$$

there is now a coupling between the singly, doubly and triply excited state determinants. Repeating this process, substituting $\langle\Psi_{ab}^{rs}|$, $\langle\Psi_{abc}^{rst}|$, and so on, results in a set of equations to be solved simultaneously in order to obtain the correlation energy. In full CI, as all excited state determinants and mixing of determinants are included, all correlation is included.

What is immediately obvious from both the construction of the full CI matrix and the accompanying coupling equations is that the number of terms becomes exorbitantly large even when applied to small molecules. As a consequence, full CI

is highly impractical to apply directly. Despite this, full CI remains a pedagogically valuable method, hence various methods have been developed to truncate full CI expansion.

The simplest truncated full CI method involves ignoring the triply, and higher, excited determinants, resulting in a significant reduction in the coupling terms. This method is known as Configuration Interaction Singles and Doubles, or CISD. The CISD method has been further extended by reducing the truncation to include the triply excited determinants (CISDT) and quadruply excited determinants (CISDTQ), and to where the reference Hartree-Fock configuration is taken to be in a high-spin state with Spin-Flips allowed in the excited state determinants (SF-CISD). Further details of these methods can be found in standard references [86] [87].

Multiconfiguration Self-Consistent Field Theory

When full CI calculations are performed, as with the truncated methods briefly described, the Hartree-Fock determinant is used as a reference determinant about which the excited determinants are constructed. This means that the coefficients, c_{oi}, of the LCAO equation (Equation 3.52) after the Hartree-Fock SCF calculation has converged, remain fixed when the CI part of the calculation is carried out.

However, when the linear combination of Equation 3.68 is written in a more general form

$$\Psi = \sum_{J} c_J \Psi_J \tag{3.82}$$

the Hartree-Fock determinant is no longer the leading term. By releasing the constraint of using the Hartree-Fock determinant as a reference state, the requirement of fixing the c_{oi} coefficients is relaxed. The simultaneous optimisation of both the expansion coefficients of the determinants (variationally) and the expansion coefficients of the basis functions leads to multiconfiguration self-consistent field

theory (MCSCF). In this way MCSCF methods are seen to be a combination of CI and Hartree-Fock. MCSCF, as indicated by "SCF," is similar to Hartree-Fock in that optimisation is carried out iteratively, hence when there is only one multi configuration, the result is the same as Hartree-Fock.

Although Hartree-Fock SCF optimisations usually find the minimum energy of a system, the result of the calculation can be checked by calculating and diagonalising the matrix of the second derivatives of coefficients, c_{oi}. This diagonalisation of the second derivative matrix is, however, necessary in MCSCF calculations, as it is much less certain that the MCSCF optimised energy is a minimum i.e. it may be a transition state, or a minimum in some but not all coordinates. Consequently the second derivatives return a negative or positive value and the nature of the MCSCF calculated minimum may be ascertained, for example, only positive values are a minimum.

It should also be noted that despite a certain degree of static and dynamic correlation accounted for in MCSCF calculations, it is not possible to separate their contributions to an energy lowering in comparison to the result of a Hartree-Fock calculation. Accordingly, MCSCF methods are often used solely for the purpose of generating a qualitatively correct wavefunction and it is assumed that only static correlation is included.

A major difficulty in carrying out MCSCF calculations is deciding which configurations are necessary to be included for the investigation of a particular molecule or property, especially when excited states are to be included in the calculation. A common approach is the complete active space self-consistent field method (CASSCF or just CAS). In this approach the orbitals of the system of interest are divided into three groups, termed *inactive*, *active* and *virtual* orbitals. The inactive or core orbitals are necessarily doubly occupied in all included determinants, usually selected to be the lowest energy spacial orbitals. The virtual orbitals are

necessarily unoccupied in all determinants and are high energy spacial orbitals. The active orbitals, or CAS orbitals, are in an intermediate region between the inactive and virtual orbitals. It is over the active orbitals only that a "full" CI calculation is performed, allowing all possible excitations of the active orbitals, hence the occupation numbers of the orbitals are no longer restricted to 0 and 2.

When a CASSCF calculation is carried out, the number of electrons in the active space and the number of orbitals are specified by the user. The choice of orbitals to be included in the active space presents a major consideration to the successful calculation of the properties of a molecule. Consequently it is often useful to have considered the bonding and hence the expected orbitals in addition to the completion of some preliminary Hartree-Fock calculations to gain an idea of the energetic ordering and the computational description of the shape of the orbitals. Not only should the character of the active orbitals be considered (a poor choice may insufficiently describe particular bonding properties) but also the number of active orbitals (too large and the number of determinants in the expansion can become unfeasibly large). As a result, this selection of the size and composition of the active space requires experimentation often with the view of gaining experience and intuition as well as achieving the desired result.

In the case of excited state calculations, in addition to the numbers of electrons and orbitals, the number of excited states and their weights (coefficients) are specified by the user. Consequently, when carrying out excited state calculations, the character and number of active orbitals may need to be reassessed in comparison to an equivalent ground state calculation. Including too large a number of excited states may also render the calculation unfeasible. In spite of these drawbacks, CASSCF remains one of the most popular post Hartree-Fock methods.

A variation of CASSCF, called the restricted active space self-consistent field method (RASSCF or just RAS), exists whereby the active space is further subdi-

vided into three sections named RAS1, RAS2, and RAS3. As with the CASSCF method, the total number of electrons to be included in the active space must be specified, but in this case is a total number across the 3 RAS regions. The RAS1 region is populated by a specified minimum number of electrons, while the RAS3 is populated by a specified maximum number of electrons. The RAS2 region has no restrictions to the number of electrons. It is this RAS2 region that is most similar to the active space in CASSCF calculations. As a result of the specified minimum and maximum occupation numbers in the RAS1 and RAS3 regions, the number of excitations from RAS1 to RAS2, and from RAS2 to RAS3 is restricted.

Perturbation Theory

An alternative to including correlation effects in calculations, as opposed to treating the full CI expansion with some kind of approximation or truncation, is perturbation theory. The general idea of this class of methods is that when the Hamiltonian is too complex, more simple, known Hamiltonians can be used with the addition of extra terms, or additional simple Hamiltonians, which are seen as perturbations to the original one chosen. The category of time-independent, as opposed to timedependent, perturbation theory is commonly referred to as Rayleigh-Schrödinger (RS) perturbation theory as Schrödinger was the first to propose the method [90], who in turn credited Rayleigh's investigation into the effect of small inhomogeneities (perturbations) to a harmonically vibrating string. If a Hamiltonian operator, to which a weak physical disturbance (a small perturbation) is added, the total Hamiltonian can be expressed as

$$\hat{H} = \hat{H}_0 + \lambda \hat{V} \tag{3.83}$$

where $\hat{V}$ is the perturbation and λ is a parameter controlling the perturbation size. Consequently the wavefunction and, hence, energy of the system are now perturbed and can be expressed in terms of a power series, as

$$\Psi = \Psi^0 + \lambda\Psi^1 + \lambda^2\Psi^2 + \cdots \tag{3.84}$$

and

$$E = E^0 + \lambda E^1 + \lambda^2 E^2 + \cdots \tag{3.85}$$

From these two equations, a general form for the energy and the wavefunction can be written as

$$\begin{aligned} \Psi^k &= \frac{1}{k!}\frac{d^k\Psi}{d\lambda^k}\bigg|_{\lambda=0} \\ E^k &= \frac{1}{k!}\frac{d^kE}{d\lambda^k}\bigg|_{\lambda=0} \end{aligned} \tag{3.86}$$

where $k \to \infty$ and the specific values of k represent the kth order perturbation. In this expression it can be seen that when $\lambda = 0$ the first, unperturbed, term in the series remains. As stated previously, the perturbation is weak or small so, as the order increases, the terms become rapidly smaller.

Substituting Equations 3.83, 3.84 and 3.85 into the TISE yields the expression

$$(\hat{H}_0 + \lambda\hat{V})(\Psi^0 + \lambda\Psi^1 + \lambda^2\Psi^2\cdots) = (E^0 + \lambda E^1 + \lambda^2 E^2 + \cdots)(\Psi^0 + \lambda\Psi^1 + \lambda^2\Psi^2\cdots) \tag{3.87}$$

It can be seen that if this equation is expanded and the coefficients of each power of λ are compared, an infinite series of simultaneous equations would be obtained. It is also clear that the zero-order equation returns the unperturbed Schrödinger equation. A full explanation of the details of the consequences of normalisation etc. can be found in [86].

The most commonly used perturbation method is Møller-Plesset (MP) perturbation, referred to, in general terms, as MPn, where n defines the order of the perturbation to be included in the calculation. In this method the unperturbed Hamiltonian is defined as a shifted Fock operator of the form

$$\hat{H}_0 = \hat{f} + \langle \Psi_0 | \hat{H} - \hat{f} | \Psi_0 \rangle \tag{3.88}$$

The perturbation operator can then be defined as

$$\hat{V} = \hat{H} - \hat{H}_0 = \hat{H} - (\hat{f} + \langle \Psi_0 | \hat{H} - \hat{f} | \Psi_0 \rangle) \tag{3.89}$$

which is the termed the correlation potential. As seen previously, the normalised Slater determinant is the lowest eigenstate of the the Fock operator

$$\begin{aligned} \hat{f}\Psi_0 &= \sum_{i=1}^{N} \hat{f}(i)\Psi_0 \\ &= 2\sum_{j=1}^{N/2} \varepsilon_j \Psi_0 \end{aligned} \tag{3.90}$$

where N is the number of electrons. This equation is written both in terms of the one electron Fock operator, $\hat{f}(i)$, and in terms of the energy of the doubly occupied spacial orbital. It can be seen here that the Hamiltonian becomes a sum over N non-interacting, one electron Fock operators, or the sum of occupied orbital energies. As the Slater determinant is an eigenstate of the Fock operator, it implies that the zero order energy is the expectation value of the Hamiltonian with respect to the normalised Slater determinant, or, more simply, the Hartree-Fock energy.

$$\hat{f}\Psi_0 - \langle \Psi_0 | \hat{f} | \Psi_0 \rangle \Psi_0 = 0 \implies \hat{H}_0 \Psi_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle \Psi_0 \tag{3.91}$$

It is easily shown by Brillouin's theorem that as the singly excited determinant does not mix with the ground, or Hartree-Fock, determinant the MP1 energy is zero. When considering a second order perturbation in this method, MP2, as the singly excited determinant cancels, the doubly excited determinant must be defined, the energy of which can be given as

$$E_{MP2} = \sum_{\substack{a<b \\ r<s}} \frac{(\langle\psi_a\psi_b|\psi_r\psi_s\rangle - \langle\psi_a\psi_b|\psi_s\psi_r\rangle)^2}{\varepsilon_a + \varepsilon_b - \varepsilon_r - \varepsilon_s} \tag{3.92}$$

where, as previously, the subscripts a and b are over occupied orbitals, and the subscripts r and s are over occupied orbitals. Consequently, the energies at each level of MPn perturbation must be added to the Hartree-Fock energy as

$$E_{tot} = E_{HF} + E_{MP2} + E_{MP3} + \cdots + E_{MPn} \tag{3.93}$$

It should be noted that as this method is not variational, the addition of the MP3 correction is not necessarily an improvement, or lower in energy, that the MP2. Given rapidly increasing complexity of the calculation with increasing numbers of n in MPn calculations, MP6 calculations are only possible for small molecules, the usual limit being MP4. It should also be noted that standard MP calculations cannot be used for the calculation of excited states, though one method involving a combination of CASSCF and MP2 exists that is able to treat excited states, called CASMP2, or more commonly CASPT2 (CAS with second order perturbation theory) [91].

Coupled Cluster Theory

Another widely used post-Hartree-Fock method which, in a sense, uses both the ideas of perturbation theory and the (full) CI, is coupled cluster theory (CC). Developed as a perturbative variant of Sinanoğlu's Many Electron Theory [92], CC methods differ from usual perturbation methods in that rather than including all corrections to a given order, they include all corrections of a given type to infinite order [86]. In this method the full CI wavefunction can be defined as

$$\Psi = e^{\mathbf{T}}\Psi_{HF} \tag{3.94}$$

where $\mathbf{T}$ is the excitation operator, given as

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \cdots + \mathbf{T}_N \tag{3.95}$$

where N is the number of electrons. Consequently it is these $\mathbf{T}_i$ excitation operators, or cluster operators, that generate all of the i-tuply excited determinants from a reference. The exponential operator can hence be written in a series expansion as

$$e^{\mathbf{T}} = 1 + \mathbf{T} + \frac{1}{2!}\mathbf{T}^2 + \frac{1}{3!}\mathbf{T}^3 + \cdots \tag{3.96}$$

and, for example, the $\mathbf{T}_2$ operator acting upon the Hartree-Fock wavefunction would be given as

$$\mathbf{T}_2\Psi_{HF} = \sum_{\substack{a<b \\ r<s}} t_{ab}^{rs}\Psi_{ab}^{rs} \tag{3.97}$$

Here, as previously, the a and b refer to occupied orbitals, r and s refer to virtual orbitals, and the expansion coefficients t are commonly referred to as the amplitudes. These amplitudes are equivalent to the c_i coefficients in Equation 3.68.

What may be immediately obvious is that there appears to be little advantage to the use of this exponential function as opposed to the full CI equations. The principal difference between the two methods is that the excitation operator is truncated in a different manner to the truncated CI methods. If only the double excitation operator is considered, then $\mathbf{T} = \mathbf{T}_2$, and the expansion of the exponential function is given as

$$\Psi_{CCD} = \left(1 + \mathbf{T}_2 + \frac{\mathbf{T}_2^2}{2!} + \frac{\mathbf{T}_2^3}{3!} + \cdots\right)\Psi_{HF} \tag{3.98}$$

If only the double excitations were included in a CI calculation (CID) it would be expressed as the first two terms in the parenthesis of this equation. As this CCD

expansion includes the products of 2, 3, 4, etc. $\mathbf{T}_2$ terms, the quadruple, hextuple, octuple, etc. excitations are also included.

Consequently, the major concern of the CC methods is the determination of the amplitudes, t. This is carried out in a similar manner to the Hartree-Fock method, where trial wavefunctions, expressed as Slater determinants, are calculated which produces a set of non-linear, coupled equations in the amplitudes. These are then usually solved by an iterative method. The CC energy can then be calculated as

$$\langle \Psi_{HF} | \hat{H} | e^{\mathbf{T}} \Psi_{HF} \rangle = E_{CC} \tag{3.99}$$

The most commonly used CC method is CCSD, where the single and double excitations are included. It should be noted that the inclusion of the single excitations in this method does not present a major expense in calculation time or memory, by comparison to CCD, as the increased accuracy is desirable.

As with all previously mentioned methods, there are various other CC methods, the most common in usage are CCSD(T), where triple excitations are included perturbatively and equations-of-motion CCSD (EOM-CCSD), whereby the treatment and study of electronically excited states can be included.

Consequently, it can be seen that there is a wide variety of electronic structure methods, each holding differing starting principles and presenting different layers of approximations. It is instructive to look into the details of a variety of electronic structure methods not only with an aim toward carrying out calculations, but also as a study of the mathematical techniques and principles used.

3.3 Nuclear Dynamics

The main focus of electronic structure methods is in solving the TISE, generally within the bounds of the Born-Oppenheimer approximation. However, as informative as single point calculations are, molecules are very rarely completely

stationary. Consequently, a variety of methods have been developed with a view of solving the TDSE. These methods are with a focus on moving nuclei, hence the term dynamics. Generally dynamics methods fall into three categories:

- Classical
- Semi-classical
- Quantum

Classical dynamics, usually referred to as molecular mechanics, methods are methods by which the atoms, or in some cases groups of atoms, of a system are treated as single classical particles defined with a radius, polarizability and net charge. The bonding between these particles is described not by overlap of orbitals but by a precalculated or experimentally defined bond distance between two particles with the bond itself being represented as a spring. The term *force field* in this context refers to the functional form and set of parameters used to calculate the potential energy, and hence potential energy surfaces of the system. As the particles in the system are parametrically defined and the equations of motion are based in classical dynamics, the system size is limited only by computational capacity available. There are many software packages offering molecular mechanics simulations, some examples of commonly used softwares being CHARMM, AMBER, GROMACS.

Semi-classical methods approach the simulation of molecular dynamics by incorporating quantum mechanical behaviours into classical dynamics simulations. These methods are generally used to include the presence and effects of excited states, as opposed to classical dynamics methods which are generally restricted to the ground state of a system. There are many different methods and approaches to the inclusion of quantum mechanical behaviours resulting in perhaps the most varied of these three categories. The Ehrenfest, or mean-field, method takes a

linear combination of the adiabatic states of a system and the molecular dynamics runs on an average potential energy surface. While in the BOA electronic motion is explicitly governed by the nuclear motion and is not time dependent, in the Ehrenfest representation the electronic motion is implicitly related to the nuclear motion and is explicitly time dependent. More details of this method will be discussed in Chapter 6.

Surface hopping methods are methods by which an ensemble, or swarm, of trajectories are propagated on adiabatic surfaces of a system, supplied by an external electronic structure method or precomputed surfaces, and in regions where the quantum amplitude is high, such as regions dominated by non-adiabatic effects, hopping can occur between surfaces. Consequently a variety of methods have been developed with differing methods by which the trajectories are described and propagated, the manner by which the hopping occurs, and the effect a successful hop has on the ensemble. Examples of surface hopping codes currently in use are AIMS, Newton-X and SHARC.

Another semi-classical treatment of molecular systems, based on Feynman path integrals [93], is path-integral molecular dynamics (PIMD). In this approach the BOA is used to separate the nuclear and electronic parts of the wavefunction. The effective Hamiltonian, derived from Feynman's path integral, is used to govern a classical representation of a number of artificially constructed particles. Sets of these particles, connected by harmonic potential springs, are used to represent each of the nuclei resulting in classical, albeit complex, system. There are a number of the path integral based methods including RPMD, CMD and FK-QCW.

A third method currently in development for the semiclassical treatment of molecular systems is exact factorization. As in the Ehrenfest method, the exact factorization method has time dependent electronic motion. In addition, this electronic motion is explicitly coupled to, but not governed by, the nuclear coordinate.

This method, however, has only been applied to one-dimensional systems but could prove to be a promising alternative semi-classical method. Recent developments of this method can be found in References [94–97]

Quantum molecular dynamics methods attempt to describe the time dependent evolution of a chemical system at the atomic level including all quantum effects. This is achieved by attempting to directly solve the TDSE. In this category of methods, instead of describing the nuclei as particles, they are described by a wavefunction. At different points in time the energy of the system can be found, which can be generalised for all molecular configurations by a potential energy hypersurface. As the nuclei are described as a nonstationary superposition of eigenstates, or wavepackets, they are hence able to interact and are subject to interference effects. Regions on the surfaces where high density is found indicates points where the molecule is likely to be.

3.3.1 Multi-Configuration Time-Dependent Hartree Method

Wavepacket dynamics requires powerful numerical methods to describe the time-evolution of the system. While the pictorial representation of the evolving wavepacket can provide some interesting chemical insight, valuable information can be extracted from said wavepacket over the course of a calculation. However, these exact methods suffer from exponential scaling with an increase in number of degrees of freedom, thus limiting said methods to no more than 2 atoms in 3 dimensions, or 6 atoms in one dimension. Consequently it has been necessary to develop more approximate methods.

Although time-dependent Hartree or self-consistent field methods have existed as a concept since at least the 1960s [98], and even as early as the 1930s [99], the multi-configuration time-dependent Hartree method, or MCTDH, achieved its first published results in 1990. [100] By expanding the wavefunction in terms of a set of time-dependent basis functions, and using variational equations of motion, the

basis functions, known as single-particle functions (SPFs), thus follow the evolving wavepacket. A desirable feature of the use of SPFs is that the result converges towards the exact result with increasing basis size.

To simplify notation, if the TDSE is recast, using atomic units where $\hbar = 1$, it has the form

$$i\dot{\Psi} = \hat{H}\Psi \tag{3.100}$$

The aim of the MCTDH method is to represent the wavefunction and Hamiltonian with a set of basis functions in order to be able to solve the TDSE. In the standard method, representing each of a system's degrees of freedom (DOFs), κ, by a one-dimensional time-independent basis set and where the wavefunction is expanded in a direct, Hartree product scheme, yields

$$\Psi(Q_1, \cdots, Q_f, t) = \sum_{j_1=1}^{N_1} \cdots \sum_{j_f=1}^{N_f} C_{j_1 \cdots j_f}(t) \prod_{\kappa=1}^{f} \chi_{j_\kappa}^{(\kappa)}(Q_\kappa) \tag{3.101}$$

where the number of DOFs is specified by f, $Q_1, \cdots, Q_f$ define the nuclear coordinates, N_κ denote the number of basis functions used to represent the κth DOF, and $C_{j_1 \cdots j_\kappa}$ are the timedependent expansion coefficients. This is the same as the full CI representation. Using the Dirac-Frenkel variational principle [99, 101] the equations of motion for the expansion coefficients can be derived as

$$\langle \delta\Psi | \hat{H} - i\frac{\partial}{\partial t} | \Psi \rangle = 0 \tag{3.102}$$

which leads to

$$i\dot{C}_J = \sum_L \hat{H}_{JL} C_L \tag{3.103}$$

Here, the multiindex $J = j_1 \cdots j_f$ has been established, as it is for L, and the Hamiltonian is given in the matrix representation of the product basis set as

$\hat{H}_{JL} = \langle \chi_{j_f}^{(1)} \cdots \chi_{j_f}^{(f)} | \hat{H} | \chi_{j_f}^{(1)} \cdots \chi_{j_f}^{(f)} \rangle$. The basis functions are usually replaced with a discrete variable representation (DVR) [102–104] so that the action of the Hamiltonian on the wavefunction can be evaluated both efficiently and accurately. As the kinetic energy part of the Hamiltonian can be written in tensor form, using fact that the potential energy on a DVR grid is diagonal, and assuming, for simplicity, that the same number of grid points is needed for each DOF ($N = N_1 = \cdots = N_f$), the computational effort required to carry out an evaluation of the R.H.S of Equation 3.103 is proportional to fN^{f+1}.

However, as a propagation algorithm also requires at least three wavefunctions in memory, and complex words require 16bytes of memory, the scaling of the memory requirement is $3 \times N^f \times 16\text{B}$. Consequently the standard method presented here would have a memory requirement of 48GB in the case of $N = 32$ and $f = 6$, and is therefore restricted to small system size.

The MCTDH method employs a time dependent, smaller basis of SPFs as an intermediate, the anstaz for the MCTDH wavefunction [100, 102, 105–108] is given as

$$\begin{aligned} \Psi(Q_1, \cdots, Q_f, t) &= \sum_{j_1=1}^{n_1} \cdots \sum_{j_f=1}^{n_f} A_{j_1 \cdots j_f}(t) \prod_{\kappa=1}^{f} \varphi_{j_\kappa}^{(\kappa)}(Q_\kappa) \\ &= \sum_{J} A_J \Phi_J \end{aligned} \tag{3.104}$$

where $n_\kappa < N_\kappa$. Implicitly defined above, the f-dimensional Hartree product of SPFs is given by Φ_J, and $A_J = A_{j_1 \cdots j_\kappa}$ are the MCTDH expansion coefficients. The SPFs, $\varphi_{j_\kappa}^{(\kappa)}$, are themselves represented by linear combinations of a primitive basis

$$\varphi_{j_\kappa}^{(\kappa)}(Q_\kappa, t) = \sum_{i_\kappa=1}^{N_\kappa} c_{i_\kappa\, j_\kappa}^{(\kappa)}(t) \chi_{i_\kappa}^{(\kappa)}(Q_\kappa) \tag{3.105}$$

Although this representation of time dependent SPFs and coefficients is not

unique, additional constraints placed on the SPFs allowing them to remain orthonormal throughout the propagation yielded unique equations of motion. This is achieved by the constraint

$$\langle \varphi_j^{(\kappa)}(0) | \varphi_l^{(\kappa)}(0) \rangle = \delta_{jl} \tag{3.106}$$

By introducing the single-hole functions

$$\Psi_l^{(\kappa)} = \langle \varphi_l^{(\kappa)} | \Psi \rangle = \sum_J^{\kappa} A_{J_l^\kappa} \prod_{\kappa' \neq \kappa} \varphi_{j_{\kappa'}}^{(\kappa')} \tag{3.107}$$

the equations of motion for both the coefficients and the SPFs can be written more concisely. Here J_l^κ denotes that the κth entry of the composite index of J is set to l, and the summation is over all DOFs excluding the κth. As a result of the single hole functions, the total wavefunction for any DOF, κ, can be written as

$$\Psi = \sum_l \varphi_l^{(\kappa)} \Psi_l^{(\kappa)} \tag{3.108}$$

It is this expansion that is used in the derivation of the equations of motion for the SPFs.

Defining the mean field as

$$\langle H \rangle_{jl}^{(\kappa)} = \langle \Psi_j^{(\kappa)} | \hat{H} | \Psi_l^{(\kappa)} \rangle \tag{3.109}$$

where the matrix elements are operators on the κth DOF, and the density matrices as

$$\rho_{jl}^{(\kappa)} = \langle \Psi_j^{(\kappa)} | \Psi_l^{(\kappa)} \rangle = \sum_J^{\kappa} A_{J_j^\kappa}^* A_{J_l^\kappa} \tag{3.110}$$

the MCTDH projector can, finally, be defined as

$$P^{(\kappa)} = \sum_{j=1}^{n_\kappa} |\varphi_j^{(\kappa)}\rangle \langle \varphi_j^{(\kappa)}| \tag{3.111}$$

Applying the Dirac-Frenkel variation principle, Equation 3.102, to the MCTDH ansatz of Equation 3.104, and applying some algebra, [102, 105–107] the equations of motion can compactly be written as

$$i\dot{A}_J = \sum_L \langle \Phi_J | \hat{H} | \Phi_L \rangle A_L \tag{3.112}$$

$$i\dot{\varphi}^{(\kappa)} = \left(1 - P^{(\kappa)}\right) \left(\rho^{(\kappa)}\right)^{-1} \langle \mathbf{H} \rangle^{(\kappa)} \varphi^{(\kappa)} \tag{3.113}$$

where the vector notation for the SPFs

$$\varphi^{(\kappa)} = \left(\varphi_1^{(\kappa)}, \cdots, \varphi_{n_\kappa}^{(\kappa)}\right)^T \tag{3.114}$$

has been adopted.

As a result, following directly from the variational principle [102, 105], the normalisation of the MCTDH equations and, in the case of time-independent Hamiltonians, the total energy is conserved. Following from this derivation it can be seen that when $n_\kappa = 1$, the MCTDH equations simplify to the time-dependent Hartree method. By increasing the value of n_κ, an increasing amount of correlation is included until when $n_\kappa = N_\kappa$, and the MCTDH method is equivalent to the standard method.

The MCTDH equations provide a useful starting point for further approximations. One such approximation is to replace some of the MCTDH standard flexible SPFs with Gaussian functions [43, 109] with a view of providing a more efficient treatment of potentially larger molecular systems. The MCTDH ansatz using Gaussian functions can hence be recast as

$$\Psi(Q_1, \cdots, Q_f, t) = \sum_{j_1=1} \cdots \sum_{j_f=1} A_{j_1 \cdots j_f}(t) \prod_{\kappa=1}^{d} \varphi_{j_\kappa}^{(\kappa)}(Q_\kappa, t) \prod_{\kappa=(d+1)}^{f} g_{j_\kappa}^{(\kappa)}(Q_\kappa, t) \tag{3.115}$$

Using a combination of SPFs and Gaussian parameters has been found to be suitable for the treatment of system-bath problems [110] as the system requires a highly quantum mechanical description while the bath is a set of oscillators well described by these Gaussian functions. The presence of the Gaussian functions in this way leads to the term Gaussian wavepackets, or GWPs, and this version of MCTDH is commonly referred to as G-MCTDH.

The Variational Multi-Configurational Gaussian Method

The natural progression from the G-MCTDH method is to replace all of the grid-based SPFs with Gaussian functions, yielding a purely GWP method. As the fully variational development of the MCTDH equations is used, the GWPs in what is known as the variational multi-configurational Gaussian (vMCG) [43, 110–112] method, do not follow classical trajectories, though the GWPs can be reduced to follow classical trajectories.

The vMCG ansatz has the form

$$\Psi(\mathbf{x}, t) = \sum_{j=1}^{n} A_j(t) g_j(\mathbf{x}, t) \tag{3.116}$$

where the basis functions and coefficients of the system are both time-dependent, and the coordinate has been simplified to the vector $\boldsymbol{x}$. These multidimensional Gaussian functions, with all DOFs combined, have the following form, in matrix notation

$$g_j(\boldsymbol{x}, t) = \exp(\boldsymbol{x}^T \cdot \boldsymbol{\varsigma}_j \cdot \boldsymbol{x} + \boldsymbol{\xi}_j \cdot \boldsymbol{x} + \eta_j) \tag{3.117}$$

the generally time-dependent, complex parameters of which are represented by a square matrix, $\boldsymbol{\varsigma}$ and a vector, $\boldsymbol{\xi}$, the dimensions of which are given by the number of DOFs, and a number, η. These numbers can all be collected together in a vector

$$\mathbf{\Lambda}_j = \left\{ \boldsymbol{\varsigma}_j, \boldsymbol{\xi}_j, \eta_j \right\} \tag{3.118}$$

The use of time-dependent GWPs was implemented by Heller and co-workers [113,114] in semi-classical dynamics calculations, the terminology of which will be followed here. The choice of the $\boldsymbol{\varsigma}_j$ matrix define the type of Gaussian function used: thawed, separable, or frozen. When a *thawed* Gaussian is used, the matrix $\boldsymbol{\varsigma}_j$ contains both diagonal and off-diagonal elements, meaning coupling between DOFs is allowed. A *separable* Gaussian is where the matrix $\boldsymbol{\varsigma}_j$ is diagonal. A *frozen* Gaussian refers to where the diagonal elements of the matrix $\boldsymbol{\varsigma}_j$ remain at fixed values during the wavepacket propagation. Generally, in vMCG calculations, frozen Gaussians are used, though the calculations appear to be sensitive to the choice of the width of the Gaussian.

The form of the Gaussians can more easily be seen by rewriting Equation 3.117 in linear notation, for the separable case

$$g_j(\mathbf{x}, t) = \exp\left(\sum_{\kappa} \varsigma_{j\kappa} x_{\kappa}^2 + \xi_{j\kappa} x_{\kappa} + \eta_j \right) \tag{3.119}$$

where, as previously, κ runs over the DOFs of the system. The use of a single function in this notation, as opposed to writing in terms of a product of one dimensional Gaussian functions, emphasises the coupling of the parameters in the dynamics calculations. It then follows that by defining the relationships

$$\left.\begin{aligned} \varsigma_{j\kappa} &= -a_{j\kappa} \\ \xi_{j\kappa} &= 2a_{j\kappa} q_{j\kappa} + i p_{j\kappa} \\ \eta_j &= \sum_{\kappa} (-a_{j\kappa} q_{j\kappa}^2 - i p_{j\kappa} q_{j\kappa}) + i\gamma_j \end{aligned}\right\} \tag{3.120}$$

Equation 3.119 can then be written in the more intuitive Heller form

$$g_j(\mathbf{x}, t) = \exp\left(\sum_{\kappa} -a_{j\kappa}(x_{\kappa} - q_{j\kappa})^2 + i p_{j\kappa}(x_{\kappa} - q_{j\kappa}) + i\gamma_j \right) \tag{3.121}$$

In this form it can more easily be seen that the ς parameters represent the widths of the Gaussian functions, ξ represents the momentum and coordinate of the centre of the functions, and η represents the remaining parameters including the phase, γ, of the functions. In simpler terms, the ς are the quadratic parameters, the ξ are the linear parameters and the η are the scalar parameters.

As with the MCTDH method, if the Dirac-Frenkel variational principle is applied to the wavefunction where variations to one of the time-dependent coefficients

$$\delta\Psi = \delta A_j g_j \tag{3.122}$$

and to one of the Gaussian parameters, $\mathbf{\Lambda}_j$,

$$\delta\Psi = \delta\lambda_{j\alpha} A_j \frac{\partial g_j}{\partial \lambda_{j\alpha}} \tag{3.123}$$

where $\lambda_{j\alpha}$ is a parameter of the jth Gaussian, g_j, result in the equations of motion. The equations of motion for the time-dependent coefficients has the form

$$i\dot{A}_j = \sum_{lm} [S_{jl}]^{-1} (H_{lm} - i\tau_{lm}) A_m \tag{3.124}$$

where the overlap matrix, $\boldsymbol{S}$, has the elements

$$S_{jl} = \langle g_j | g_l \rangle \tag{3.125}$$

and the Hamiltonian matrix, $\boldsymbol{H}$, has the elements

$$H_{jl} = \langle g_j | \hat{H} | g_l \rangle \tag{3.126}$$

In order to maintain the normalisation of the Gaussians during the propagation, the diagonal of the overlap time-derivative matrix, $\boldsymbol{\tau}$, must be zero. By considering the Gaussian parameters to be part of a single vector, $\mathbf{\Lambda}_j$, using the chain rule,

$\boldsymbol{\tau}$ can explicitly be written as a function of the time derivative of the Gaussian parameters

$$\tau_{jl} = \langle g_j | \dot{g}_l \rangle = \sum_{\alpha} \left\langle g_j \left| \frac{\partial g_l}{\partial \lambda_{l\alpha}} \dot{\lambda}_{l\alpha} \right. \right\rangle \tag{3.127}$$

If N_κ is the number of DOFs, the maximum values of the indices α define whether the width of the Gaussian is thawed ($\max(\alpha) = N_\kappa(N_\kappa+1)$), separable ($\max(\alpha) = 2N_\kappa + 1$), or frozen ($\max(\alpha) = N_\kappa + 1$).

The equations of motion for the Gaussian parameters can be written in a compact matrix form as

$$i\dot{\boldsymbol{\Lambda}} = [\boldsymbol{C}]^{-1}\boldsymbol{Y} \tag{3.128}$$

where this apparently simple notation belies the underlying complexity. By expanding the matrix $\boldsymbol{C}$ and the vector $\boldsymbol{Y}$

$$C_{j\alpha,l\beta} = \rho_{jl}\left(S_{jl}^{(\alpha\beta)} - [\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{S}^{(0\beta)}]_{jl}\right) \tag{3.129}$$

$$Y_{j\alpha} = \sum_{l} \rho_{jl}\left(H_{jl}^{(\alpha 0)} - [\boldsymbol{S}^{(\alpha 0)}\boldsymbol{S}^{-1}\boldsymbol{H}]_{jl}\right) \tag{3.130}$$

where ρ_{jl} is an element of the density matrix

$$\rho_{jl} = A_j^* A_l \tag{3.131}$$

the additional definitions of the elements of the $\boldsymbol{S}$ matrices and the Hamiltonian matrix $\boldsymbol{H}$

$$\left.\begin{aligned} S_{jl}^{(\alpha\beta)} &= \left\langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} \middle| \frac{\partial g_l}{\partial \lambda_{l\beta}} \right\rangle \\ S_{jl}^{(\alpha 0)} &= \left\langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} \middle| g_l \right\rangle \\ H_{jl}^{(\alpha 0)} &= \left\langle \frac{\partial g_j}{\partial \lambda_{j\alpha}} \middle| H \middle| g_l \right\rangle \end{aligned}\right\} \tag{3.132}$$

reveal some of the underlying complexity. Dependent on the α and β parameters, the S matrices correspond to different orders of Gaussian moments e.g. $M_{jl}^{\kappa} = \langle g_j|x_\kappa|g_l\rangle$, $M_{jl}^{\kappa\mu} = \langle g_j|x_\kappa x_\mu|g_l\rangle$. Of particular interest is the order up to which the matrix $\boldsymbol{S}^{(0\alpha)}$ can represent the Gaussian moments. In the case of thawed Gaussians, all moments can be represented up to second-order while for separable Gaussians all zero-order, linear and second-order moments can be represented. In the case of the use of frozen Gaussians, only the zero-order and linear moments can be represented. It is the representation of the Gaussian moments by the $\boldsymbol{S}^{(0\alpha)}$ matrix that proves to be important.

Progressing from the ideas of Chapter 2.3 it can be seen that the zeroth-order, or separable, part of the Hamiltonian can be expanded as a power series in terms of the Gaussian moments as

$$H_{jl} = M_{jl}^{(0)} X_l^{(0)} + \sum_{\kappa} M_{jl}^{(\kappa)} X_l^{(\kappa)} + \sum_{\kappa\mu} M_{jl}^{(\kappa\mu)} X_l^{(\kappa\mu)} + \cdots \tag{3.133}$$

If the local harmonic approximation (LHA) is used, the potential energy of a given multi-dimensional Gaussian can be expanded, up to second order, with respect to the time dependent centre coordinate, $\boldsymbol{q}_j(t)$, as

$$V_j(\boldsymbol{r}) = V_{j0} + \sum_{\kappa} V'_{j,\kappa}(x_\kappa - q_{j\kappa}) + \frac{1}{2}\sum_{\kappa\mu} V''_{j,\kappa\mu}(x_\kappa - q_{j\kappa})(x_\mu - q_{j\mu}) \tag{3.134}$$

where the centre of $g_j(\mathbf{r}, t)$ has been used for $\boldsymbol{q}_j(t)$, and the derivatives of the potential taken from this point. As the kinetic energy operator can be taken to have a separable form, i.e. $\hat{T} = \sum_j \frac{1}{2m_j}\frac{\partial^2}{\partial x_j^2}$ the X coefficients of Equation 3.133 can be written as

$$X_j^{(0)} = \sum_{\kappa}\left(\frac{\varsigma_{j\kappa}}{m_\kappa} + \frac{\xi_{j\kappa}^2}{2m_\kappa}\right) + V_{j0} - \sum_{\kappa} V'_{j,\kappa} q_{j\kappa} + \frac{1}{2}\sum_{\kappa\mu} V''_{j,\kappa\mu} q_{j\kappa} q_{j\mu} \tag{3.135}$$

$$X_j^{(\kappa)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_\mu V''_{j,\kappa\mu} q_{j\mu} \tag{3.136}$$

$$X_j^{(\kappa\mu)} = -\frac{2}{m_\kappa}\varsigma_{j\kappa}\varsigma_{j\mu} + \frac{1}{2}V''_{j,\kappa\mu} \tag{3.137}$$

The $\boldsymbol{Y}$-vector can also be separated, into two parts

$$\boldsymbol{Y} = \boldsymbol{Y}_0 + \boldsymbol{Y}_R \tag{3.138}$$

The part of the Hamiltonian that can be expressed in terms of $\boldsymbol{S}^{(0\alpha)}$ is represented by $\boldsymbol{Y}_0$, with the remaining part of the Hamiltonian expressed by $\boldsymbol{Y}_R$, known as the residual term. The residual term includes both the correlation terms, as well as the higher order terms of the separable part of the Hamiltonian. Using the relationship between the overlap matrix elements $\boldsymbol{S}^{(\alpha\beta)}$ and the Gaussian moments, M, and comparing Equations 3.129 and 3.130, Equation 3.138 can be rewritten as

$$Y_{j\alpha} = \sum_{l\beta} C_{j\alpha,l\beta} X_l^{(\beta)} + Y_{R,j\alpha} \tag{3.139}$$

Consequently, the equation of motion for the Gaussian parameters can be further simplified to

$$i\dot{\boldsymbol{\Lambda}} = \boldsymbol{X} + \boldsymbol{C}^{-1}\boldsymbol{Y}_R \tag{3.140}$$

The significance of this, so-called "CX formalism" is realised when the practical implications are considered. As the $\boldsymbol{C}$-matrix is inverted during propagation, the removal of part of the Hamiltonian from the $\boldsymbol{C}^{-1}\boldsymbol{Y}$-term results in a reduction of possible numerical errors and hence increases the stability of the propagation. Another desirable feature is that the parameter equations of motion can be divided into classical and non-classical parts.

The separation of classical and non-classical parts is achieved in the vMCG method by putting only the classical Gaussian propagation terms into $\boldsymbol{X}$ with the remaining (quantum) contributions kept in the $\boldsymbol{Y}_R$ term. However the method for

the separation of these terms is dependent on whether thawed, separable or frozen width Gaussian are being treated. In the case of thawed width Gaussians it is known that these Gaussians move classically in a harmonic potential and hence, in the above notation, a complete solution is achieved using $i\dot{\mathbf{\Lambda}} = \boldsymbol{X}$. [109,112,115]. In the case of frozen width Gaussians, the separation is not as straightforward.

When frozen width Gaussian basis functions are used, only the linear parameters of κ are varied with time. Consequently, the zeroth-order Hamiltonian can be written in terms of Equations 3.135 and 3.136 as

$$H_{jl} = S_{jl}^{(0)} X_l^{(0)} + \sum_{\kappa} S_{jl}^{(0\kappa)} X_l^{(\kappa)} + \sum_{\kappa\mu} M_{jl}^{(\kappa\mu)} X_l^{(\kappa\mu)} + ... \tag{3.141}$$

In this form, the Hamiltonian is easily separable into the $\boldsymbol{Y}_0$ and $\boldsymbol{Y}_R$ terms.

This is not the case for the equations of motion as the linear parameter of the zeroth-order part of the equations of motion have the form

$$\begin{aligned} i\dot{\xi}_{j\kappa}^{(0)} &= -i2\varsigma_{j\kappa}\dot{q}_{j\kappa} - \dot{p}_{j\kappa} = X_j^{(\kappa)} \\ &= -\frac{2}{m_\kappa}\varsigma_{j\kappa}\xi_{j\kappa} + V'_{j,\kappa} - \sum_{\mu} V''_{j,\kappa\mu} q_{j\mu} \end{aligned} \tag{3.142}$$

whereas the classical equations of motion for the position and momentum derivatives have the form

$$\left.\begin{aligned} \dot{q}_{j\kappa} &= \frac{p_{j\kappa}}{m_\kappa} \\ \dot{p}_{j\kappa} &= -V'_{j,\kappa} \end{aligned}\right\} \tag{3.143}$$

As a consequence, in order to obtain the zeroth-order linear parameter in terms of the classical equations of motion ($i\dot{\xi}_{j\kappa}^{(0)} = -i2\varsigma_{j\kappa}\frac{p_{j\kappa}}{m_\kappa} + V'_{j,\kappa}$), the terms $-\frac{2}{m_\kappa}\varsigma_{j\kappa}(\xi_{j\kappa} - ip_{j\kappa})$ and $-\sum_\mu V''_{j,\kappa\mu} q_{j\mu}$ must go into $\boldsymbol{Y}_R$. In the vMCG method, $\bar{\xi}_{j\kappa}^{(0)} = q_{j\kappa} + ip_{j\kappa}$ is propagated (as opposed to $\xi_{j\kappa}^{(0)} = -2\varsigma_{j\kappa}q_{j\kappa} + ip_{j\kappa}$) and when necessary, for example when τ_{jl} is calculated, a constant width parameter, $\varsigma_{j\kappa}$, is used to reconstruct $\xi_{j\kappa}^{(0)}$. All that remains is to obtain the remaining part of $\dot{\xi}_{j\kappa}$ from $\boldsymbol{Y}_R$.

$$i\dot{\xi}_{R,j\kappa} = \sum_{l\mu} C^{-1}_{j\kappa l\mu} Y_{R,l\mu} \tag{3.144}$$

This results in an overall equation of the form

$$i\dot{\bar{\xi}}_{j\kappa} = V'_{j,\kappa} - i\frac{p_{j\kappa}}{m_\kappa} + \Re(i\dot{\xi}_{R,j\kappa}) - \frac{1}{2\varsigma_{j\kappa}} i\Im(i\dot{\xi}_{R,j\kappa}) \tag{3.145}$$

If the last two terms are ignored, all GWPs will follow classical trajectories. It should also be stated that for coherent states in an appropriate width harmonic well, the last two terms of this expression cancel. Details on the performance and application of this method can be found in reference [43] and references therein.

The Direct-Dynamics Variational Multi-Configurational Gaussian Method

One of the major drawbacks to grid-based methods is the requirement that the potential energy surfaces must be computed or fitted before any calculations are performed. This presents a major bottleneck to the completion of quantum-chemistry calculations. A relatively recent development with a view of combating this bottleneck is the direct-dynamics variational multi-configurational Gaussian method (DD-vMCG) [43,116–118]. In this method the matrix elements of the Hamiltonian must be evaluated in order to integrate the equations of motion for the time dependent coefficients (Equation 3.124) and the Gaussian parameters (Equation 3.128). If the dynamics is run using Cartesian coordinates or normal modes, the kinetic energy operator has a simple analytical form, and by adopting the LHA (Equation 3.134) only the energies, gradients and Hessians at the centres of the Gaussians are required at each time step. As a result these values are easily evaluated on-the-fly via an interface with an external quantum chemistry software package.

Another advantage of the DD-vMCG method is in the form of the choice of coordinates. In the usual grid-based methods a careful choice of the use of coordinates is required when fitting the potential energy surfaces so as to avoid spurious

correlated motions. This can often lead to complex expressions for the kinetic energy operator. As the DD-vMCG method circumvents the requirement for simplified precomputed surfaces, a coordinate system can be chosen such that the kinetic energy operator is easily evaluated. Consequently studies have been performed using the DD-vMCG method in atomic Cartesian coordinates, [117, 119] Jacobi coordinates, [119] and in normal modes. [116, 120–124] The uses of these different coordinate systems are not without their drawbacks. When Cartesian coordinates are used, the translations and rotations of the total system are not easily separated from the vibrational motions. When using Jacobi coordinates, although the translations of the system can be separated, the initial representation of the wavepacket as a superposition of GWPs is not easily defined. The use of normal modes goes further by approximately removing rotational motion of the system, though the selection of specific nuclear motion can be difficult to resolve. This difficulty in the use of normal modes is, however, a relatively trivial issue.

As stated previously, the DD-vMCG method utilises external quantum chemistry software packages to calculate the potential energy surfaces. These electronic structure methods are used to calculate the energy, gradient and Hessian, so even relatively simple systems with a moderate basis-set and level of theory will take a significant amount of time. Consequently it is undesirable to carry out these time-consuming electronic structure calculations for every point reached by the GWPs. The DD-vMCG method bypasses this issue by constructing PESs using the LHA and creates a database of electronic energies and other calculated information. [119] The idea of using a database in this way has been used before in other classical trajectory [125, 126] and quantum trajectory [127] methods.

As a DD-vMCG calculation is carried out, a linear combination of multidimensional GWPs, each centred at a particular point in configuration space, is used to construct the nuclear wavepacket. Instead of calling the external elec-

tronic structure program at each time step, the database is then consulted. The Euclidean norm of the difference vector between all points of the new molecular geometry and all existing geometries in the database is then calculated as a method by which the distance between two structures can be measured. The parameter *dbmin* (database minimum), defined at the start of the calculation, is then compared to the lowest value of the calculated norm.

If this lowest value is higher than the predefined value of *dbmin* then an electronic structure program is called in order to calculate the energy, gradient and Hessian of the PES, and the information is stored in the relevant database. Additional information such as the dipole moments, derivative couplings and MO coefficients may also be stored. The PES is then expanded to the second order of a Taylor series using the stored data as a reference

$$\boldsymbol{V}(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x}_0) + \boldsymbol{g}(\boldsymbol{x}_0).(\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0).\boldsymbol{H}(\boldsymbol{x}_0).(\boldsymbol{x} - \boldsymbol{x}_0) \tag{3.146}$$

where the gradient, $\boldsymbol{g}(\boldsymbol{x}_0)$, and the Hessian, $\boldsymbol{H}(\boldsymbol{x}_0)$, of the adiabatic PES are evaluated at x_0 and are relative to a change in nuclear geometry. This expansion is then used as a basis for the LHA, which in turn is used to calculate the required matrix elements in the equations of motion. The expansion up to and including the second derivatives, however, is not only used in the construction of the PESs.

If the lowest value of the norm is lower than the predefined value of *dbmin* then a modified Shepard interpolation [127] is carried out in order to obtain the energy, gradient and Hessian for the LHA. This Shepard weighted interpolation has the form

$$V(\boldsymbol{q}) = \sum_i \omega_i(\boldsymbol{q}) T_i(\boldsymbol{q}) \tag{3.147}$$

where T_i is the ith database entry centred Taylor series expansion and ω_i is defined as

$$\omega_i(\boldsymbol{q}) = \frac{\nu_i(\boldsymbol{q})}{\sum_j \nu_j(\boldsymbol{q})} \tag{3.148}$$

where

$$\nu_i(\boldsymbol{q}) = \frac{1}{|\boldsymbol{q} - \boldsymbol{q}_i|^{2p}} \tag{3.149}$$

It has been found that setting $p = 2$ yields the best results. [43] This set of equations (Equation 3.147-3.149) is used to gain the interpolated energy, gradient and Hessian. However, the Taylor series as used in Equation 3.147 must be truncated to first order for the gradients and zeroth order for the Hessians. Using Equation 3.146 the PES is then formed using the second-order Taylor series for each database point, $\boldsymbol{x}_i$, and the energy for the new geometry is evaluated.

The choice of the value of *dbmin* appears to be system dependent [43], though running an exploratory set of calculations varying this value will yield informative results. It should also be noted, from a practical perspective, that if an electronic structure calculation fails to complete, this interpolation is also used.

Although the use of the database in this way can significantly reduce the amount of real time required for a full propagation, there remains the issue of the evaluation of the Hessian matrix, which is a computationally intensive process even for Hartree-Fock methods. As the energy, gradient vector and Hessian matrix are essential to the use of the LHA in the calculation of extrapolated points, as well as for the calculation of matrix elements, a method was required by which the explicit calculation of the Hessian at every database point could be avoided. As this is a common issue, many methods for the approximation of the Hessian exist requiring only a reference Hessian, and the current gradient information, known as Hessian update algorithms. Consequently, the Powell update algorithm [128] is used, where the updated Hessian is calculated using the following equation

$$\mathbf{H_{New}} = \mathbf{H_{Old}} + \frac{1}{\boldsymbol{\delta}.\boldsymbol{\delta}}(\boldsymbol{\epsilon} \otimes \boldsymbol{\delta} + \boldsymbol{\delta} \otimes \boldsymbol{\epsilon}) - \frac{\boldsymbol{\epsilon}.\boldsymbol{\delta}}{(\boldsymbol{\delta}.\boldsymbol{\delta})^2}\mathbf{H_{Old}}.\boldsymbol{\delta} \otimes \boldsymbol{\delta}.\mathbf{H_{Old}} \tag{3.150}$$

where $\mathbf{H}$, with the appropriate label, is the Hessian, $\boldsymbol{\delta}$ is the vector of the position difference, and $\boldsymbol{\epsilon}$ is the vector of the gradient difference between the "old" and "new" geometries.

At the first point in a DD-vMCG calculation, when it is performed using an empty database, an electronic structure calculation is carried out in order to obtain the energy, gradient and full Hessian, and these data are used as the reference. In the case where a populated database is used, the first entry in the database is used as the reference. The propagation continues and when the calculation reaches a point where a new electronic structure calculation is required, as opposed to an extrapolated point, only the energy and gradient at this new point are calculated. The DD-vMCG program then calculates the Euclidean norm distance between the reference and the new point, as well as calculating the distances between the reference and all other points in the database. The points in the database are then divided into two subsets. The first, "internal" subset of points comprises points which are closer to the reference than the new point. The second, "external" subset of points comprises the points which are further away from the reference than the new point. The gradient at the new point and the gradient and Hessian of each of the points in the internal subset in turn is used to obtain a set of Powell updated Hessians for the new point, $\mathbf{H}_i$. As the distances between the new point and of the internal points have been calculated, d_i, the Hessian at the new point is given by the weighted sum

$$\mathbf{H_{New}} = \frac{\sum_{i \in \mathbf{Internal}} d_i^{-4} \mathbf{H}_{\mathbf{Old}}^{(i)}}{\sum_{i \in \mathbf{Internal}} d_i^{-4}} \tag{3.151}$$

The new Hessian is then added to the database and each of the Hessians in the external subset are updated in a similar manner, including the new Hessian as

part of the internal subset. The performance of this Hessian update procedure has been tested, details of which can be found in reference [43]. It was found that although the use of the Hessian update procedure was dependent on the system being studied, it may provide a crucial time saving if a high level electronic structure method is being used to calculate new points.

A final important feature to note in the DD-vMCG method is the form of potential energy surface, specifically the representation of the states in the calculation. Although the external electronic structure programs calculate points along adiabatic surfaces, the wavepacket is propagated along the diabatic surfaces. Given the prerequisites for the DD-vMCG method, a method by which the diabatic surfaces can be calculated is required which can even-handedly account for on-the-fly calculated surfaces with multiple states with an unknown number of state crossings. The method by which this is carried out is known as *propagation diabatisation* and is achieved by the following scheme. It should be noted that for ease of reading some of the equations presented in Chapter 2.5 will be reiterated here.

As adiabatic states form an orthonormal set, the elements of the transformation matrix, $\boldsymbol{S}$, can be defined as

$$S_{ji} = \langle \psi_j | \varphi_i \rangle \tag{3.152}$$

which is simply the overlap between the adiabatic, ψ, and diabatic, φ, states, which is a scalar function of the position. Both the adiabatic and diabatic states form a complete set, so a gradient of these matrix elements can be taken

$$\nabla S_{ji} = \sum_{k=1}^{N} \langle \nabla \psi_j | \psi_k \rangle \langle \psi_k | \varphi_i \rangle + \sum_{k=1}^{N} \langle \psi_j | \varphi_k \rangle \langle \varphi_k | \nabla \varphi_i \rangle \tag{3.153}$$

the final term of which, in a strictly diabatic representation, is

$$\langle \varphi_k | \nabla \varphi_i \rangle = 0, \forall\, i, k \tag{3.154}$$

If it is assumed that the couplings between states in a diabatic representation are sufficiently small, it can hence be written

$$\nabla S_{ji} = \sum_{k=1}^{N} \langle \nabla \psi_j | \psi_k \rangle S_{ki} \tag{3.155}$$

If the definition of the nonadiabatic coupling vector is recalled, the elements of which are defined as

$$F_{ij} = \langle \psi_i | \nabla \varphi_j \rangle = -\langle \nabla \psi_i | \psi_j \rangle = \frac{\langle \psi_i | \boldsymbol{\nabla} \hat{H} | \psi_j \rangle}{V_j - V_i} \tag{3.156}$$

where V_i and V_j are the nonadiabatic energies of states i and j, the adiabatic states are orthonormal and it can hence be written

$$\nabla S_{ij} = -\sum_{k=1}^{N} \boldsymbol{F}_{ik} S_{kj} \tag{3.157}$$

assuming that a complete basis set is used. The key equation to this propagation diabatisation is the matrix form of this equation

$$\boldsymbol{\nabla S} = -\boldsymbol{FS} \tag{3.158}$$

where this $\boldsymbol{S}$ matrix is taken to be at a point where the adiabatic and diabatic surfaces are the same, typically the Frank-Condon point. It is the solutions to this differential equation that define the scheme of the diabatisation. Propagating over a short time step, the formal solution to this equation is

$$\boldsymbol{S}(\boldsymbol{q} - \Delta \boldsymbol{q}) = \exp\left(-\int_{\boldsymbol{q}}^{\boldsymbol{q}-\Delta \boldsymbol{q}} \boldsymbol{F} \cdot d\boldsymbol{q} \right) \boldsymbol{S}(\boldsymbol{q}) \tag{3.159}$$

However, in this form it can be seen that in order to solve this equation an exponential of a matrix expression must be taken. It can also be seen that the integral of a function with an unknown analytic form must be taken. The second of these issues can be overcome using a simple numerical integration along the straight line

between $\boldsymbol{q}$ and $\boldsymbol{q} + \Delta\boldsymbol{q}$ using the trapezium rule, resulting in a matrix of scalars parameters. The first issue is not as straightforward to solve.

The equation for the propagation, Equation 3.159, does not guarantee that a unitary matrix $\boldsymbol{S}(\boldsymbol{q})$ will return a unitary matrix for $\boldsymbol{S}(\boldsymbol{q} - \Delta\boldsymbol{q})$. Following the Esry and Sadeghpour [129] method, which uses the Cayley-Hamilton form of the propagator, the unitarity of the transformation matrix can be maintained, resulting in

$$\exp\left(\frac{1}{2}\int \boldsymbol{F}\cdot d\boldsymbol{q}\right)\boldsymbol{S}(\boldsymbol{q}-\Delta\boldsymbol{q}) = \exp\left(-\frac{1}{2}\int \boldsymbol{F}\cdot d\boldsymbol{q}\right)\boldsymbol{S}(\boldsymbol{q}) \tag{3.160}$$

which is a rearranged form of Equation 3.159. By utilising a Taylor series expansion of the exponentials, and the resultant matrix on the left-hand side is inverted, the final transformation matrix in terms of $\boldsymbol{S}(\boldsymbol{q} - \Delta\boldsymbol{q})$ is obtained. Full details of the method by which this implemented in the DD-vMCG software can be found in Reference [130].

As mentioned previously, this method is correct only for a complete basis set of states. However, tests to date indicate that the performance of this propagation diabatisation scheme is correct as smooth, diabatic potential energy surfaces are calculated. [130]

Chapter 4

Formimidic Acid Proton Transfer

4.1 Introduction

Formamide (FAM), also known as methanamide, over the years since its first published synthesis in 1863 [131] has been an important but, until relatively recently, often disregarded compound. Though its initial synthesis was by a reaction of ethyl formate with ammonia, advancements in technology enabled it to be formed in direct synthesis from the carbonylation of ammonia with an alcohol-dissolved alkoxide catalyst [132, 133]

$$\mathrm{CO} + \mathrm{NH_3} \xrightarrow{NaOCH_3} \mathrm{HCONH_2} \tag{4.1}$$

which requires a pressure of 0.8-1.7 MPa at 75-80 ^{o}C. Alternatively FAM can be formed via a two-step indirect synthesis involving the treatment of liquid methylformate with gaseous ammonia [134, 135]

$$\begin{gathered} \mathrm{CO} + \mathrm{CH_3OH} \xrightarrow{NaOCH_3} \mathrm{HCOOCH_3} \\ \mathrm{HCOOCH_3} + \mathrm{NH_3} \longrightarrow \mathrm{HCONH_2} + \mathrm{CH_3OH} \end{gathered} \tag{4.2}$$

an industrial example of which being the BASF Process [135, 136], with a requisite temperature 40-100 ^{o}C at 0.1-0.3 MPa.

A cursory investigation into the early history of research directly focused on the properties and uses of FAM yields relatively little information within the field of chemistry, the reasons for which are not immediately obvious. A more com-

prehensive exploration, however, produces a wealth of information the detailed description of which would be overly didactic within the context of this study. Consequently, a few specific areas of interdisciplinary interest will be discussed along with a short summary of current uses.

In the early attempts at the systematic classification of bacteria it was recognised that the identification of bacterial polysaccharides was essential [137] (later discovered to be specifically relevant to encapsulated bacteria such as Streptococcus pneumoniae). Consequently an efficient method for the solvation of bacteria without destroying the polysaccharide structures was required. Given the experimental techniques available at the time, it was proposed that the use of neat FAM at 150 ^{o}C dissolved the bacteria, leaving the polysaccharide macromolecules affected neither by dissolution nor by cross-reactions. [138] After this initial series of studies between 1923-1938 [139] relatively little work was carried out within the field of bacterial taxonomy until the 1960s [140]. This 20 year gap in research is noted in a now key review from 1980 [141] on the characteristics of Streptococcus mutans.

The break in research involving FAM during this time occurs in a number of fields and can superficially be explained by the outbreak of the World War II. Consequently scientific research during this time was focused on contributing to the war effort and in some cases can be difficult to trace, particularly if the research was for nefarious purposes. One of the more familiar chemical weapons used during World War II is the hydrogen cyanide (HCN) based pesticide Zyklon B [142]. Although the patent for Zyklon B was accepted in 1926 [143], there was a requirement for an efficient method by which HCN could be produced. Industrially, the main process for the production of HCN is the Andrussow process [144, 145], originally developed as a method of ammonia oxidation by Leonid Andrussow for the company I.G. Farben [146,147]. In the paper "The Manufacture of Hydrocyanic

Acid by the Andrussow Process" by J. M. Pirie [148] it is stated

> *A large plant for the production of hydrocyanic acid... was built in Germany during the war, but was never put into production.*

The chapter on formamides in Ullmann's Encyclopedia of Industrial Chemistry mentions a formamide vacuum synthesis of HCN [135], the reference for which is given as "I. G. Farben, FR 906 114, 1944" [149]. Pulling these references together reveals an interesting piece of history. During August and September 1944 the Eighth Air Force of the USAAF carried out 8 bombing runs, committing a total of 1083 B-17s and 41 B-24s, destroying plants referred to as "chemical works" and "oil refineries" in Oppau and Ludwigshafen [150, 151]. It is reported in the "Classified List of OTS Printed Reports" (1946) as listed under the Technical Reports Section of the Science and Technology Division of the Library of Congress, that FAM was being produced at a factory in Oppau [152], the details of which are unavailable. In the same report it lists an I. G. Farben plant in Frankfurt under the category of "Chemical Warfare". It is a reasonable conclusion to draw that a method of HCN production using FAM was never implemented as the Oppau facility providing the FAM based chemical feedstock for the plant in Frankfurt (around 80 km away) had been destroyed.

Due to its high degree of association (FAM is more highly hydrogen bonded than water at temperatures below 80 ^{o}C) its unusually high specific heat [153], its high dielectric constant and its bifunctionality, FAM has attracted commercial interest for an eclectic range of applications [135, 154]. Its use as a formylating agent [155], as a starting point for transamidation to form substituted formamides (*N,N*-dimethylformamides) [156] and in the preparation of nitrogen-containing heterocycles and formate esters [157] make it an ideal solvent or intermediate for many synthetic processes including the production of polymers [158–160], pharmaceuticals [161–163] and crop protecting agents [154]. FAM is the smallest,

most stable compound containing the four most abundant elements in the universe. FAM itself has been generally detected in the interstellar medium [164], in comets [165, 166], including Hale-Bopp, on icy grains around the protostellar object W33A [167], and from sources SgrA and SgrB2 [168, 169]. In recent years, support has been growing in the apperception of FAM as a key abiotic precursor to the synthesis of pyrimidines and nucleobases in prebiotic Earth, supporting the RNA World Hypothesis. A comprehensive review, "Formamide and the origin of life" by Saladino et al. [170], details the diverse conditions under which FAM has been used in the synthesis of nucleic bases and nucleosides, the phosphorylation of nucleosides, as a catalyst for oligomer- and polymerization of nucleotides, and the synthesis of pre-metabolic components. As Hadean Earth was not shielded from solar radiation it is reasonable to consider not only the effects of temperature, catalysis and concentration but the photoactivity of FAM in its viability as a precursor to prebiotic life. It is also, therefore, pertinent to consider how formamide may be formed in the interstellar medium. The paper "Mechanistical studies on the production of formamide (H_2NCHO) within interstellar ice analogs" of B. M. Jones et al. [171] provides a comprehensive review of studies to date pertaining to the behaviour of FAM in the celestial environment. Also covered are the possible pathways to formation, IR spectral identification within ice matrices, and analysis by mass spectrometry.

Despite this plethora of information regarding the behaviour, uses and function of FAM, to date there have been only a limited number of computational studies. [172–176] As yet there are no published full quantum dynamics studies on the behaviour of FAM.

4.2 Electronic Structure Calculations

As stated in Chapter 3.3, external quantum chemistry programs are needed in order to calculate the energy, gradient and Hessian required for the propagation in a DD-vMCG calculation. The use of a large basis set and a high level of theory would enable higher accuracy in the results; however, these computationally intensive calculations would present a severe bottleneck to the propagations over even relatively short timescales in terms of the memory and both the computational and real time for taken for the calculation. An important consideration, particularly relevant in quantum chemistry calculations involving a large number of excited states, is the stability of the quantum chemistry calculations to the movement of atoms away from equilibrium geometries. A final consideration is that the DD-vMCG method is relatively new and computational studies have, so far, been carried out testing the ability of the code to handle high numbers of degrees of freedom, but thus far no studies have been carried out examining the ability of the code to handle large numbers of excited states. Consequently a balance between accuracy, stability, computational expense, and time expense in the calculations carried out by an external quantum chemistry software must be considered from the outset of an investigation.

Due to the dearth of computational investigations into the excited states of FAM and FIM it was unknown as to how many excited states would be needed in the quantum chemistry calculations. Consequently an overall emphasis on efficiency in the quantum chemistry calculations was made, whilst maintaining the intention of gaining as accurate a description of the excited states of the molecules as possible within a reasonable time constraint.

4.2.1 Ground State Calculations

Basis Set Investigation

An initial set of calculations were carried out performing a ground state optimisation of formamide using increasing basis set size, and increasing level of theory. For consistency, the same starting geometry was used as the starting point for each of these calculations, as opposed to the usual method by which a previous lower level (basis set or theory level) calculation is used as starting point for higher level calculations. The energetic results of this initial set of calculations can be seen in Table 4.1, while the time taken for these calculations can be seen in Table 4.2. It should be noted that while the fourth-order perturbation theory method, MP4, is listed as being a lower theory level than CCSD, MP4 has been shown, using a cc-pVDZ basis set, to capture approximately 99% of electron correlation, whereas CCSD captures 98%. They are listed in this way as MP4 optimisations are not possible using the standard quantum chemistry software packages, due to the extensive coding required for these to be implemented. Consequently the MP4 energies and times are the calculated correction to the MP2 optimised geometry and in addition to the time taken for the MP2 calculations, respectively, and are listed this way as they are not optimisations.

In Table 4.1 it can be seen that there is a general trend that with increasing basis set size, across the Pople series and Dunning series, the energy at the optimised geometry is lowered. The decrease in energy across the Pople series can be seen to be stepwise, where the increase from 3- to 6-primitive Gaussians representing the inner-shell orbitals presents a first significant improvement, and the increase from double-ζ to triple-ζ presents a second significant improvement. Across the Dunning series, where only an increase from double- to triple- to quadruple-ζ is present, the same trend in decreasing energy is seen.

As expected, in Table 4.2 it can be seen that with increasing basis set size, the

	HF	MP2	MP4	CCSD
3-21G	-167.9849003	-168.3041414	-168.3291477	-168.3155118
3-21G∗	-167.9849003	-168.3041414	-168.3291477	-168.3155118
6-31G∗	-168.9307027	-169.3944734	-169.4280191	-169.4108379
6-31+G∗	-168.9385910	-169.4104353	-169.4446477	-169.4256122
6-311+G∗	-168.9800204	-169.4807448	-169.5170643	-169.4939978
cc-pVDZ	-168.9491075	-169.4311602	-169.4676206	-169.4497254
cc-pVTZ	-169.0032040	-169.6054492	-169.6467061	-169.6158302
cc-pVQZ	-169.0167414	-169.6626997	-169.7013452	x

Table 4.1: A comparison of the energies (in Hartrees) of the optimised structures of formamide showing the level of theory versus the basis set.

time taken for the optimisation also increased. The increase in calculation time across the Dunning series is, however, far more pronounced than the increase in calculation time across the Pople series.

Upon closer inspection it can also be seen that for the, almost, equivalent double-ζ (6-31+G∗ and cc-pVDZ) and triple-ζ (6-311+G∗ and cc-pVTZ) basis sets, the Dunning calculated energies are slightly lower. However, upon consideration of the times taken for these calculations it can be seen that while the calculation times for the double-ζ basis sets are comparable with increasing theory level, for the triple-ζ basis sets, the calculation times of the Dunning basis sets are significantly longer than that of the Pople basis sets. It should be noted that while the construction of these basis sets differs, as described in Chapter 3.2.2, the comparative time taken for the optimisation with increasing theory level, given the relatively small improvement in the calculated energy, presents a good initial indication that while the cc-pVDZ basis set may be suitable for the later Direct Dynamics calculations, the cc-pVTZ will not be suitable.

In Table 4.1 it can also be seen that at the lowest basis set level, 3-21G and 3-21G∗, the inclusion of the polarisation function on atoms other than the hydrogen atoms afforded energetically equivalent results. However, upon inspection of the

	HF	MP2	MP4	CCSD
3-21G	12.1	22.9	25.1	65.6
3-21G∗	12.6	23.3	25.3	65.8
6-31G∗	26.9	44.8	51.2	287.1
6-31+G∗	31.0	60.0	71.2	575.8
6-311+G∗	43.4	73.7	97.0	966.8
cc-pVDZ	31.9	65.3	73.8	418.1
cc-pVTZ	330.8	1483.2	1697.6	18771.3
cc-pVQZ	8439.0	22206.9	25639.3	x

Table 4.2: A comparison of the CPU times taken (in seconds) for each of the calculations showing theory level versus basis set. Gaussian 09 [46] was used on a Xeon ES-2640 8-core workstation.

time taken for these calculations it can be seen that the inclusion of the polarisation functions resulted in an increase in calculation time, the effect of which reduced with increasing theory level. Although these calculations were carried out only as for comparison to the higher basis sets, this was a somewhat unexpected result.

An inspection of the calculated energies with increasing level of theory, Table 4.1, shows that there is a significant drop in energy from the Hartree-Fock to the second-order perturbation correction method MP2. A smaller drop is energy is seen when the perturbation level is increased from second-order to fourth-order, MP4. Across all basis sets it can be seen that the energies calculated using the coupled-cluster method, CCSD, are between the MP2 and MP4 energies, and are in fact closer to the MP2 than the MP4. This is to be expected as the MP4 method accounts for more electron correlation than the CCSD calculation. However, the increase in calculation time from MP2 (and MP4) to CCSD is significant.

As a result of this exploration it can be seen that while the Dunning basis sets offer a higher degree of accuracy, the calculation times for higher than a double-ζ basis set would be undesirable. It can also be seen that the difference in calculation time for the Pople double-ζ to triple-ζ basis set is small at lower levels of theory, but at higher levels the difference becomes significant. Consequently it is clear

that the choice of basis set from a timing perspective will be between these Pople type basis sets, while the specific choice is dependent on the effect that the use of the theory level has on the calculation time as a balance with the increase in accuracy.

Planar vs Non-Planar Geometry

In order to carry out the Direct Dynamics calculations it is often more helpful to separate the molecular motions, or characterise the normal modes, when the molecule has symmetry. As can be seen in Table 4.3, the 6-311+G∗ CCSD optimisations (column "CCSD opt") of FIM and the transition state between FIM and FAM result in near-planar geometries, while the optimised geometry for FAM is more significantly non-planar. By changing the dihedral angles of FAM, the transition state and FIM to force the molecule to be planar, the molecules were then re-optimised using the same basis set and level of theory, the geometries for which are the columns titled "CCSD pl".

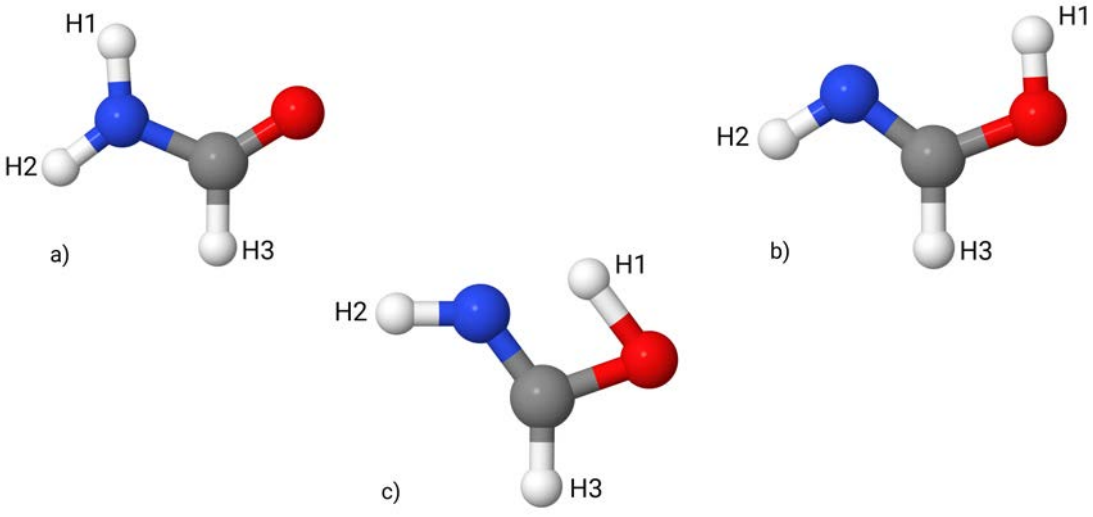


Fig. 4.1: The 6-311+G∗ CCSD planar optimised geometries of a) formamide, b) formimidic acid and c) the transition state. The labels H1, H2, H3 are for clarity in the interpretation of the data in Table 4.3.

It can readily be seen that the bond lengths and angles of the re-optimised, pla-

				FAM		TS		FIM	
				CCSD opt	CCSD pl	CCSD	CCSD pl	CCSD opt	CCSD pl
At1	At2	At3	At4	6-311+G*	6-311+G*	6-311+G*	6-311+G*	6-311+G*	6-311+G*
N	C			1.36851	1.36235	1.30276	1.30261	1.2687	1.26870
N	H2			1.00656	1.00468	1.01168	1.01167	1.01660	1.01660
C	H3			1.10608	1.10604	1.09050	1.09052	1.09270	1.09270
C	O			1.21059	1.21179	1.28020	1.28022	1.34524	1.34524
N	H1			1.00879	1.0072	1.35325	1.35301	2.31293	2.31292
O	H1			2.53242	2.55256	1.34061	1.34084	0.96663	0.96663
H2	N	C		119.210	121.555	125.421	125.409	110.673	110.673
H3	C	N		112.894	112.857	128.519	128.536	127.323	127.323
O	C	N		124.821	124.855	108.751	108.757	122.004	122.004
H1	N	C		117.235	119.435	73.8450	73.8530	54.1830	54.1830
H1	O	C		52.5940	52.5130	74.9940	74.9810	107.4690	107.469
H2	N	C	O	-164.338	180.000	179.717	180.000	179.968	180.000
H1	N	C	O	-13.061	0.000	0.050	0.000	-0.016	0.000
H1	O	C	N	5.783	0.000	-0.050	0.000	0.033	0.000

Table 4.3: The optimised geometries of formamide (FAM), formimidic acid (FIM), and the transition state (TS) between the two isomers. The "CCSD opt" columns are the results of the optimisations as seen in Table 4.1, while the "CCSD pl" columns represent the optimisations with a planar starting geometry. Here, H1 is the proton that transfers from the nitrogen in FAM to the oxygen in FIM, H2 is the proton bonded to the nitrogen in both FAM and FIM, and H3 is the proton bonded to the carbon. These labels can be seen in Figure 4.1.

	FAM[a]		TS[b]		FIM[c]	
	CCSD[1] 6-311+G*	CCSD pl[2] 6-311+G*	CCSD[1] 6-311+G*	CCSD pl[2] 6-311+G*	CCSD[1] 6-311+G*	CCSD pl[2] 6-311+G*
Energy	-169.493998	-169.493712	-169.404268	-169.404268	-169.471304	-169.471304
ΔE(1) to (2)		0.007780		0.000005		0.000000
ΔE(a) to (b)/(c)			2.44161	2.43382	0.618509	0.609729

Table 4.4: A comparison of the energies for the formamide and formimidic acid geometries calculated in Table 4.3. The energies of the optimised geometries are in Hartrees. The ΔE values are in eV and are the difference in energy between a) the relevant optimised, non-planar geometry and the planar geometry of formamide, b) the transition state, and c) the acid form.

nar structure of FIM are the same as those of the optimised, non-planar molecule. The bond lengths and angles of the re-optimised transition state, in comparison to the non-planar structure show small adjustments of the order of 10^{-4}Å and 10^{-2} degrees, respectively. As changing the dihedral angles of the FAM optimised structure represented a more significant alteration of the system, understandably the re-optimised, planar structure exhibited a more significant alteration across the molecular framework. The most significant changes in the FAM structure were seen in the positioning of the protons H1 and H2 (the protons bonded to the nitrogen) as these were the protons that were out of plane.

Table 4.4 shows the difference in the calculated optimisation energy between the "CCSD opt" to the "CCSD pl" structures. In this table it can be seen that the enforced planarization of FIM resulted in a calculated energy indistinguishable from that of the non-planar optimisation method. The change in energy from the non-planar to the planar transition state geometry was also negligible. Despite the more significant change in geometry of FAM, the resultant energy difference, of the order of 10^{-2} eV, is also small enough to be considered negligible. It can also be seen in Table 4.4 that the effect of the change from non-planar, to planar geometries resulted in a reduction of the barrier height to transition, and the energy difference between FAM and FIM was reduced, both of the order of 0.008 eV.

Consequently, as the effect of the use of planar geometries resulted in only small alterations in the energetics of the system, calculations from this point progressed using planar starting geometries.

4.2.2 Excited State Calculations

State-Averaged CAS Investigation

In order to be able to successfully carry out a Direct Dynamics calculation involving multiple states it is necessary for the non-adiabatic couplings between all of the

states to be calculated. This is particularly important when the system of interest has a manifold of excited states in close proximity between which multiple crossings and intersections may occur. Consequently the logical choice when considering the excited state dynamics of a system from a chemical perspective is the use of the multiconfigurational method of CASSCF, described in Chapter 3.2.3. All electronic structure calculations in this chapter used the Molpro 2015 program [177] which has a very efficient CASSCF procedure [178] and is able to provide all the derivative couplings between states as required for the Direct Dynamics.

The use of the CASSCF is, however, by no means a "black-box" method. It is known that a poor selection of active space size, occupancy and orbitals can lead to poorly converged calculations and instabilities. This problem is often exacerbated with the inclusion of excited states leading to even further instabilities in the calculation. It is these issues that present a major difficulty in the use of the DD-vMCG method. In addition, the inclusion of a nitrogen atom in the system is known to increase the difficulty in the accurate selection of the orbitals within the active space. As a result, as opposed to attempting to enforce a particular selection of orbitals on a specified number of excited states, a new analytical procedure for the systematic identification of the orbitals required alongside the number of excited states was developed in order to minimise the difficulties in this CASSCF selection process.

Before any active space orbitals and number of excited states can be chosen the chemical features of the system in question must be considered. It was decided that the initial selection of orbitals should have the capacity to describe a potential proton transfer between the two isomers. It is known that π-π^* transitions are optically bright so in addition to the description of a proton transfer, a sufficient description of the π-bonding network would be required. In addition, due to the system size, it was also decided that, in the interest of efficiency, a relatively

large active space could be chosen so as to maximise the number of configuration interactions to be calculated with the caveat that a smaller basis set would be required in order to minimise the time-scaling issue.

Consequently the orbital selection for FAM should include the description of the N-H(1) σ-bond and σ^*-bond, the lone-pair acceptor on the oxygen. and a description of the π bonding network containing at least one π-bonding molecular orbital and one π^*-bonding molecular orbital. The selection of orbitals to be included in the active apace for FIM should hence contain the complementary O-H(1) σ- and σ^* molecular orbitals, as well as the lone-pair acceptor on the nitrogen and the relevant π- and π^*-bonding molecular orbitals. While the choice of the 6-311+G$*$ basis set would yield more accurate results, it is known that with increasing basis-set size, the clear identification of the orbitals of a system is severely hampered, in addition to the increase in calculation time. In the interest of efficiency it was hence decided that given the choice of a large active space, the as yet unknown number of excited states required for the calculation and the limitations of the computer cluster on which the calculations would be performed, a 6-31G$*$ would be used as a compromise. A final decision made in this initial process was that the active space for FAM and FIM should be the same size.

After an initial inspection of the molecular orbitals of both FAM and FIM calculated in the initial Hartree-Fock level calculations it was decided that the active space size required would be either a CAS(10,9) (10 electrons, 9 orbitals), and CAS(10,8) (10 electrons, 8 orbitals) or a CAS(8,7) (8 electrons, 7 orbitals). Following the decision to sacrifice the size of the basis set in order to include a larger active space, it was decided that a study would be carried out using CAS(10,9) and CAS(10,8) active spaces, for comparison. Given the size of these active spaces and the limitations of the electronic structure calculation methods available, it was decided that between 2 and 10 excited states would be studied, with equal

weight on each state, at each of these levels of active space. The results of these two studies on FAM can be seen in Figures 4.2 and 4.3, which are the 6-31G* state averaged calculations using a CAS(10,9) and CAS(10,8) active space, respectively. As the method by which this study is presented is novel, a detailed discussion of how these data are interpreted will be described.

The column titles at the top of the page, SA3, SA4, SA5, and so-on refer to the total number of states included in the calculation where, for example, "SA3" means that this column represents a State Averaged calculation over a total of 3 states, including the ground state. These column titles refer to both the top and bottom sections of the figure.

The top section of the figure contains the information pertaining to the characterisation of the orbitals included in the active space and is hence called the *active space characterisation section* or the ASC. In this section the text above the dotted line lists a simplified characterisation of the occupied orbitals, whilst the text below the dotted line refers to the virtual orbitals. This simplification has been made so as to decrease the complexity of the data table.

The bottom section of this figure is called the *excited state characterisation section*, or XSC, which summarises four key pieces of information. The first line of the XSC is the calculated energy of the S_0 state in the relevant state averaged calculations. The numbers in the rows below are the excitation energies, in eV, from the relevant S_0 state to the excited state, hence the first row of numbers below the S_0 energy is the S_1, the second row is the S_2, and so on.

The coloured boxes below the excitation energies are the characterisations of the principal transition(s) contributing to the excitation, where the box on the left characterises the orbital from which the excitation occurs, and the box on the right characterises the orbital into which the electron is excited. In some instances, more than one type of excitation significantly contributes to the overall character

of the state. In these cases the numerical values were considered and the two or three most significant, listed in order, are included. It was considered that the emphasis in this type of analysis should be in the characterisation so while the numerical values are available, they have not been included. The explanation of the colours of the transitions are listed in a key to the right of the ASC.

The final piece of information included in the XSC are where the numbers of the excitation energies are written in red. These correspond to the states with an oscillator strength greater than 0.01. As with the transitions, the numerical values of these oscillator strengths are available though have not been included for simplicity.

It should also be noted that "OLP" and "NLP" are an abbreviation for "oxygen Lone Pair" and "nitrogen Lone Pair" respectively, and that in Figures 4.2 to 4.4 the numbering of the protons, for example NH(1) and NH(2), refer to the numbering used in the previous section as seen in Figure 4.1.

As an example, in Figure 4.2 the CAS(10,9) SA10 calculated transition from the S_0 to the S_2 has an excitation energy of 7.978 eV which comprises a combination of excitations from the OLP (yellow) to the NH_2 σ^* (green) and from the $\pi 2$ (red) to the π^* (blue). It can be seen that the oscillator strength of this transition is significant, likely due to the significant π-π^* character.

In Figure 4.2 the ASC for the SA7 calculation shows that the NH(1) σ- and σ^*-bonds, the oxygen lone pair, two π- and one π^*-molecular orbitals are represented, along with an additional molecular orbitals representing both the CN and NH_2 bonding as well as the CH, CN and O antibonding character. It can readily be seen in the ASC that while the active spaces for all of the state averaged calculations contain the prerequisite orbitals representing the π and OLP character, the SA7 calculation represents the only one of the CAS(10,9) calculations which fulfills the active space character criteria chosen.

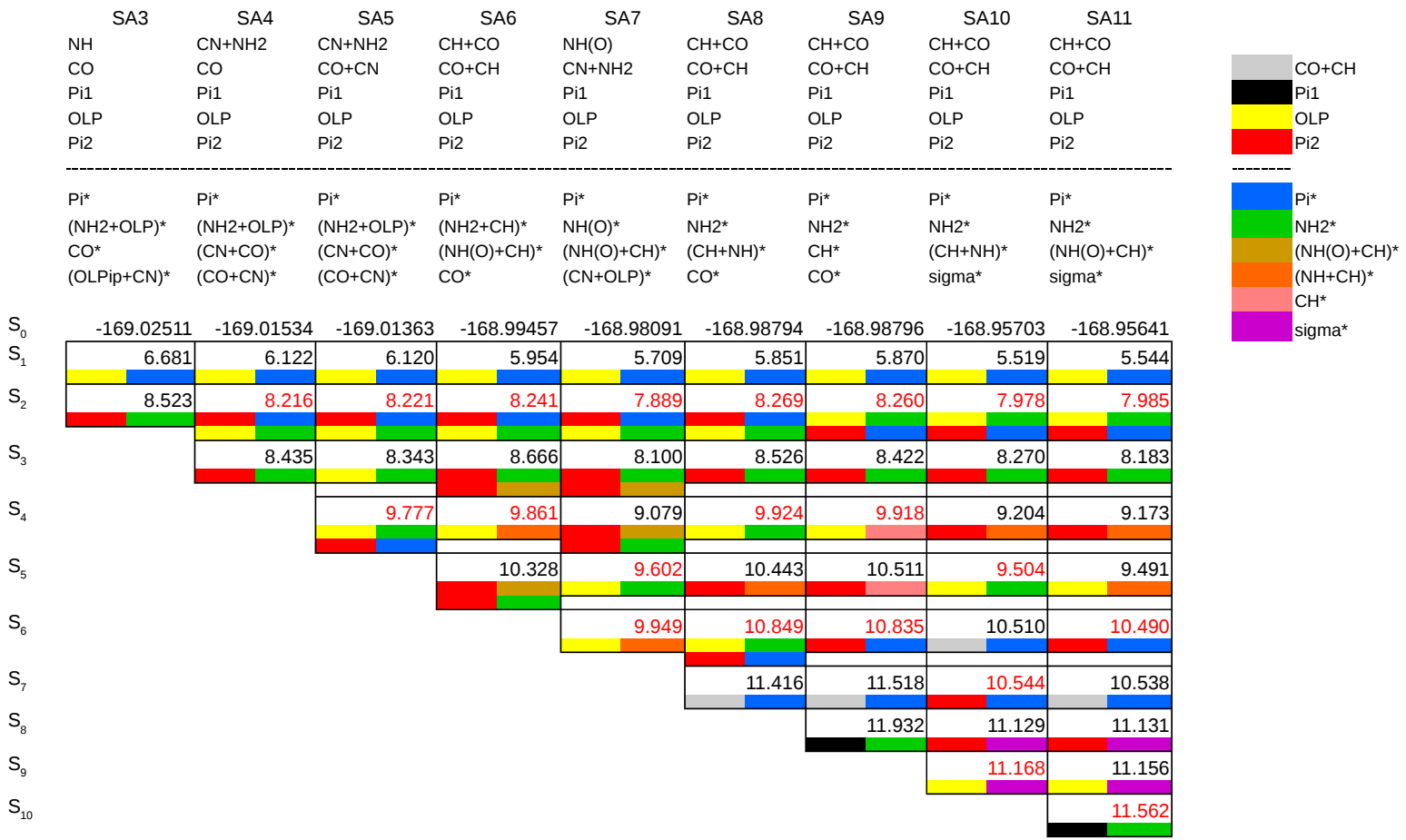


Fig. 4.2: Character of MO for formamide using a CAS(10,9) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01).

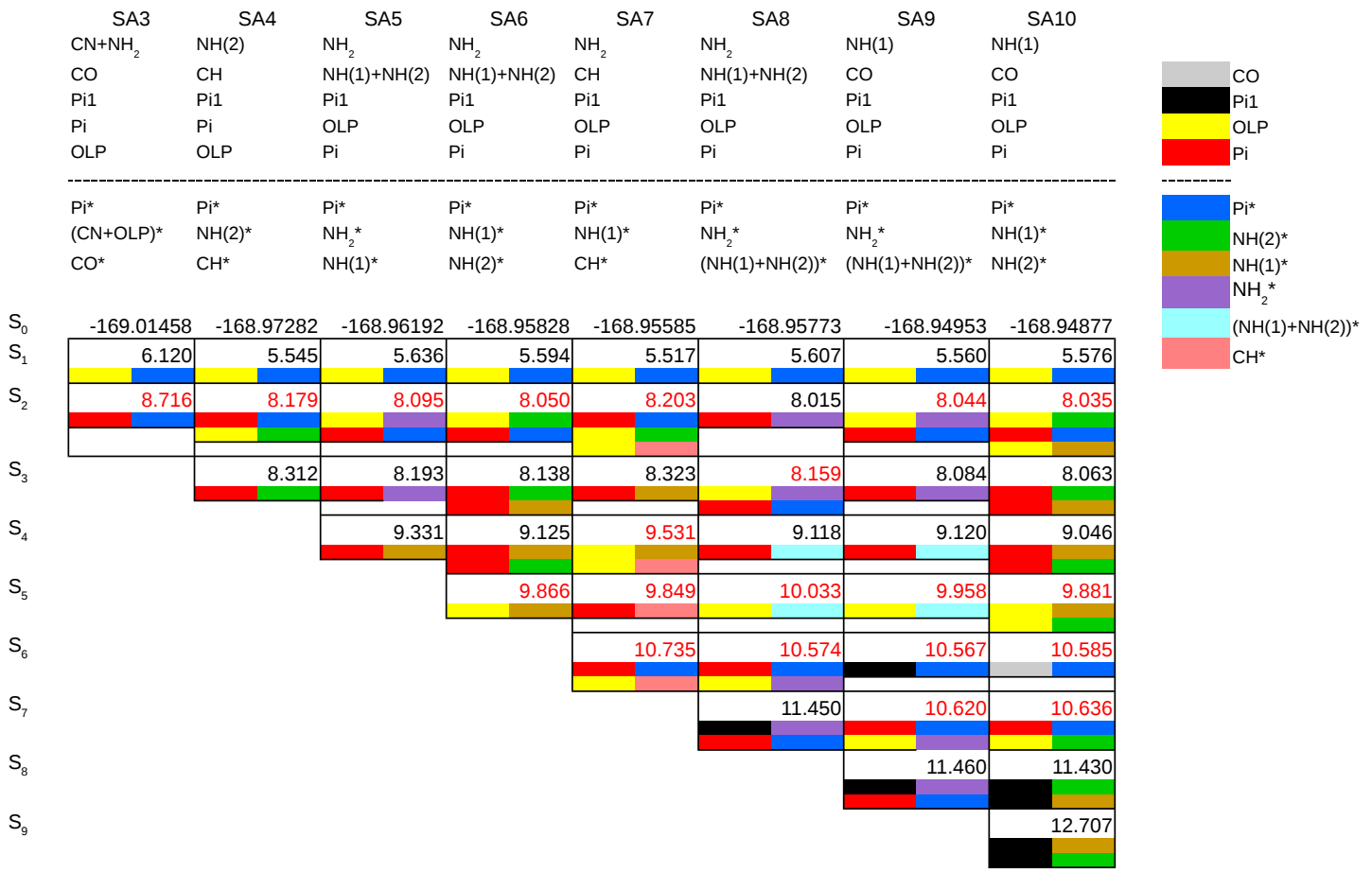


Fig. 4.3: Character of MO for formamide using a CAS(10,8) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01).

In the XSC of Figure 4.2 it can readily be seen that all of the state averaged calculations define the character of the S_1 state as being an OLP to π^* transition. Although the SA3 calculation defines the S_2 as being a π to NH_2^* transition, the rest of the state average calculations define this transition as being a combination of π to π^* and OLP to NH_2^* transitions. Upon analysis of the S_3 states it can be seen that this π to NH_2^* transition in the SA4-11 calculations more closely matches the energy of the S_2 in the SA3 calculation and hence is likely the same state.

Between the SA8-11 calculations it can be seen that the S_6 and S_7 states swap in character between a π to π^* transition and a CO+CH to π^* transition. However, although the energy difference between the S_6 and S_7 in the SA8 and 9 calculations is approximately 0.6 eV, this energy difference in the SA10 and 11 calculations reduces to approximately 0.035 eV and the character of the states swap in SA10.

Upon analysis of the XSC for the SA7 calculation it can be seen that although the character of the S_1, S_2 and S_3 states are the same as for the SA3-11, SA3-8 and SA6-11, respectively, the calculated energies of these states are significantly lower than in the neighbouring state averaged calculations. It can also be seen that the SA7 calculation is the only one that results in the character of the S_3 and S_4 states, π to NH_2^* and π to $(NH1+CH)^*$, being almost the same. In addition, the calculated energies of the higher lying excited states for the SA7 are significantly lower in energy than in any of the excitations in the neighbouring state averaged calculations. As a result of the ASC analysis the logical choice of the number of excited states for a CAS(10,9) calculation would be the SA7 calculation, though the analysis of the XSC would imply that this number of states may not be stable and hence unsuitable for further study.

For the CAS(10,8) study of FAM state averaged calculations were carried out using 2 to 10 excited states, the results of which can be seen in Figure 4.3. In the ASC it can be seen that states contain definition of the OLP and the π-bonding

molecular orbitals, while the SA4 is the only calculation which does not contain any degree of definition of the NH(1)σ-bond. Although all of the state averaged active spaces contain definition of a π^* molecular orbital, the SA3 and SA4 calculations do not contain definition of the NH(1)σ^* molecular orbital.

In the XSC of Figure 4.3 it can readily be seen that as with the CAS(10,9) calculations, all of the state averaged calculations define the character of the S_1 state as being an OLP to π^* transition. The SA5, SA8 and SA9 active spaces do not contain an orbital explicitly for the NH(2)* but by observation of the trends of the characterisation across the S_2 and S_3 states it can be seen that this NH(2)* character has been replaced by the NH_2^* character. Consequently it is reasonable to assume that the (NH(1)+NH(2))* character seen in the S_4 and S_5 states of the SA8 and SA9 calculations have been interchanged with the NH(1)* character.

Although, by comparison of the S_2 states it can be seen that the SA4-7 and SA9-11 calculations place the character of this transition as a combination of a π to π^* transition and an OLP to NH(2)*(NH_2^*) transition, the SA8 defines this state as a π to NH_2^* transition. However, if the S_3 states are also taken into account it can be seen that the S_2 and S_3 states are close in energy and it appears in the characterisation of SA8 that the S_2 and S_3 have swapped in comparison to the other state averaged calculations. Upon inspection of the SA7 calculations it can be seen that the presence of the CH-CH*-bonding definition results in the discontinuity of the π to NH(1)*((NH(1)+NH(2))*) transitional trend across the S_4 and the OLP to NH(1)*((NH(1)+NH(2))*) trend across the S_5 state.

If the region of the XSC covering the S_6 to S_8 is examined it is clear that the energy and characterisation of the SA8 S_7 and SA9 S_8 are concurrent. This leads to the observation that an additional lower order state may be present in the SA9 calculation. By inspection it can be seen that the character of the SA8 S_6 state and the SA9 and SA10 S_7 states are similar in energy and the same in character,

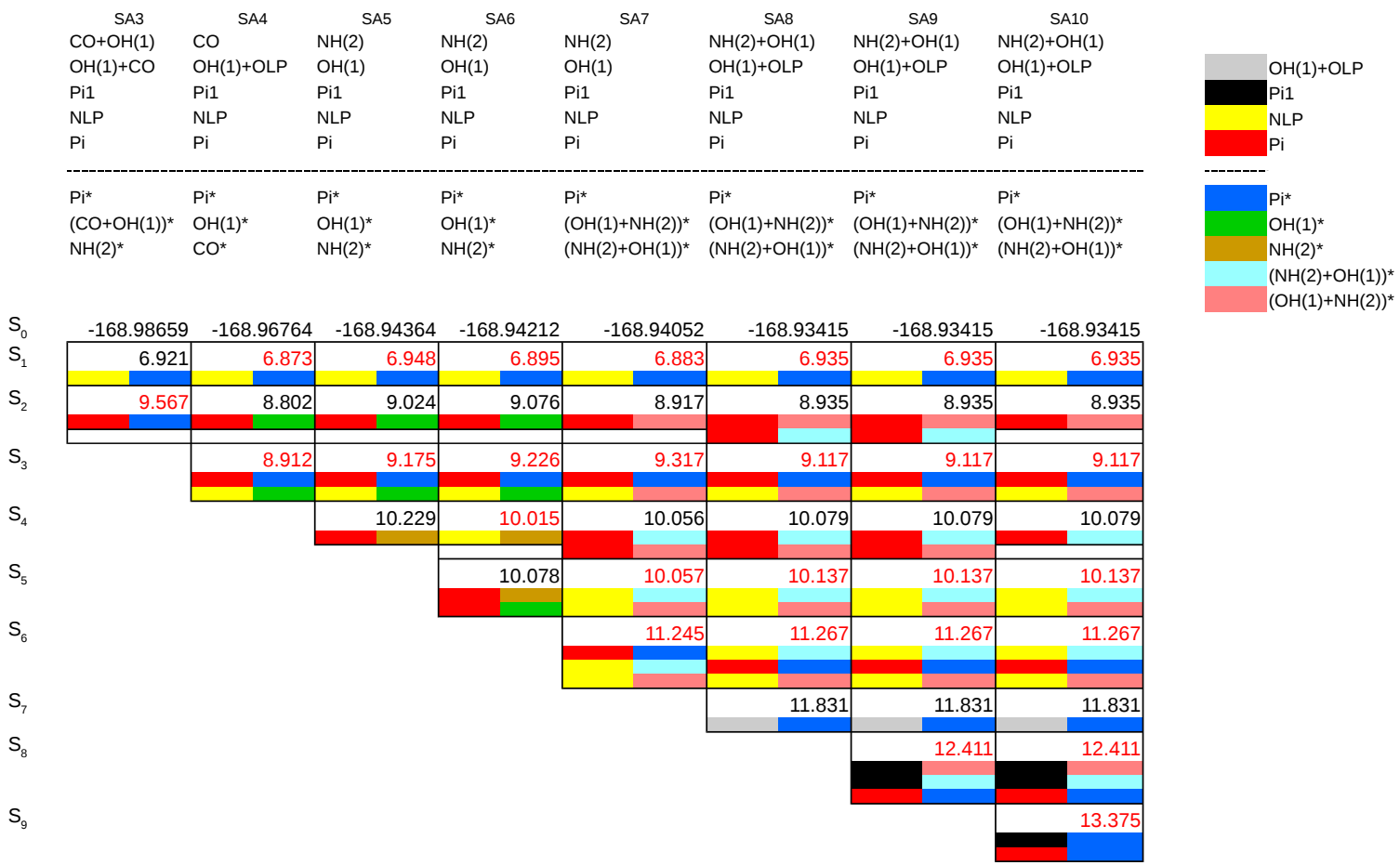


Fig. 4.4: Character of MO for formimidic acid using a CAS(10,8) and different state-averaging. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. Key to colours on the right hand side. The list of orbitals above the cells list those in the CAS space. Numbers in red are for states with significant oscillator strength (> 0.01).

which resolves the lack of a clear trend across the S_6 state. The presence of this additional state in the SA9, but particularly in the SA10 calculation, is likely due to the definition of the CO-bond in the active space.

As a result of the ASC and XSC analysis of the formamide CAS(10,8) state averaged calculations it can be stated that the most suitable states for further study are the SA5, SA6 and SA8 states. Due to the wider availability of states, and the apparent increased stability of these states, it was decided that the logical next step would be to carry out an investigation on the FIM using a CAS(10,8) active space, before making a final decision as to the number of states to be included in the Direct Dynamics study.

The results of the CAS(10,8) study of FIM can be seen in Figure 4.4 where state averaged calculations have been performed including between 3 and 10 states in total. In the ASC it can be seen that all of the active spaces contain the prerequisite OH-OH* molecular orbitals, the nitrogen lone pair, two π molecular orbitals and the single π^* molecular orbital. In order to directly map onto the FAM calculations it is required that there be additional significant character defining the NH(2) bonding. Consequently the SA3 and SA4 calculations can be immediately discounted as these states do not contain the additional bonding character defining the NH-bond.

In the XSC it can be seen across all states that the S_1 is characterised by an NLP to π^* transition. Upon inspection of the general trends across the characterisation it can be seen that the states with transitions to the OH(1)* molecular orbital become states with transitions to the (OH(1)+NH(2))*, while the states with transitions to the NH(2)* become the states with transitions to the (NH(2)+OH(1))* molecular orbital. Consequently it can be seen that aside from the SA3 calculated S_3 state and the SA6 calculated S_4 and S_5 states, the excited state characterisations are consistent with increasing numbers of excited states. This is particularly

	Energy (eV)	Oscillator Strength (au)	Main configurations
S_1	5.607	0.0008	0.69(OLP – Pi*)
S_2	8.015	0.0004	0.66(Pi – NH+NH)
S_3	8.159	0.0225	0.54(OLP – NH+NH) + 0.34(Pi – Pi*)
S_4	9.118	0.0000	0.66(Pi – NH+NH)
S_5	10.033	0.0710	0.63(OLP – NH+NH)
S_6	10.574	0.7258	0.44(Pi – Pi*) + 0.37 (OLP – NH+NH)
S_7	11.450	0.0013	0.55(Pi1 – NH+NH) + 0.36(Pi – Pi*)

Table 4.5: Formamide energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA8-CAS(10,8)/6-31G* calculation

	Energy (eV)	Oscillator Strength (au)	Main configurations
S_1	6.935	0.0105	0.687(NLP – Pi*)
S_2	8.935	0.0009	0.56(Pi – (OH+NH)*) + 0.35(Pi – (NH+OH)*)
S_3	9.117	0.1325	0.48(Pi – Pi*) + 0.37(NLP – (OH+NH)*)
S_4	10.079	0.0045	0.56(Pi – (NH+OH)*) + 0.36(Pi – (OH+NH)*)
S_5	10.137	0.1723	0.48(NLP – (NH+OH)*) + 0.47(NLP – (OH+NH)*)
S_6	11.267	0.4634	0.41(NLP – (NH+OH)*) + 0.39(Pi – Pi*) + 0.30(NLP – (OH+NH)*)
S_7	11.831	0.0000	0.676(OH+OLP – Pi*)

Table 4.6: Formimidic acid energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA8-CAS(10,8)/6-31G* calculation

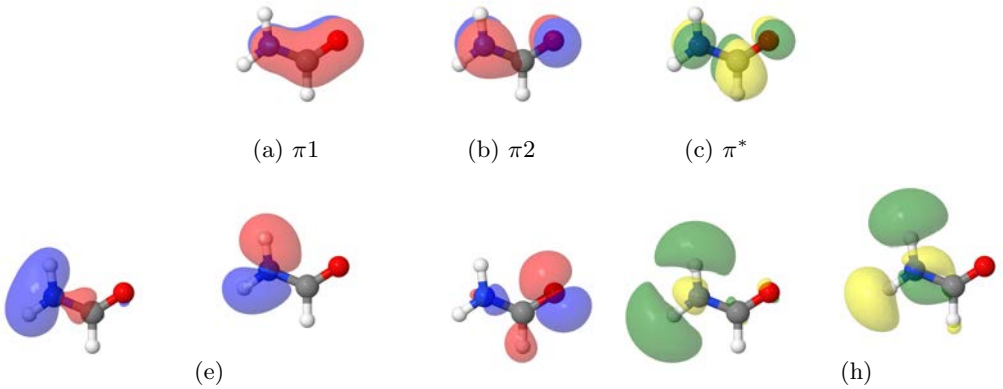
apparent in the SA8 to SA10 calculations where the calculated energies of the ground state and the excitation energies for all transitions from the ground state are concordant, indicating a high degree of stability in the calculations.

As a result of the ASC and XSC analysis of the formimidic acid CAS(10,8) state averaged calculations it can be stated that the SA5 and SA7 to SA10 are the most suitable states for further study.

If a comparison between the XSC of the FAM and FIM is carried out it can be seen that the characterisation of the S_1 and S_2 are equivalent, though the excitation energies differ by approximately 1.3 eV and 1.1 eV respectively.

From this analysis, in order to follow the prerequisite inclusion of the largest number of excited states it follows that the Direct Dynamics calculations should be

performed on both the FAM and FIM using a CAS(10,8) active space with state averaging over 8 states in total. The final results of the excitation energies, oscillator strengths and characterisations from the SA8-CAS(10,8)/6-31G* calculations of FAM and FIM are summarised in Tables 4.5 and 4.6 respectively. In Table 4.5 it can be seen that the S_6 state, characterised as a π-π^* transition, is the bright state, while the S_3 and S_5 have small but significant oscillator strengths. In Table 4.6 it can be seen that the S_6 is the bright state, and although this transition has significant π-π^* transition character, the principal component of this transition is characterised as NLP-(NH(2)+OH(1)*). It can also be seen that the oscillator strengths of the S_3 and S_5 are greatly increased in comparison to the same states in the FAM results. This comparative increase in the oscillator strength in the S_3 state is likely due to the dominant π-π^* transition in the FIM, which is only a secondary contribution to the equivalent state in FAM.



(a) π1 (b) π2 (c) π^*

(d) NH_2 (e) NH(1)+NH(2) (f) OLP (g) NH_2^* (h) (NH(1)+NH(2))*

Fig. 4.5: CAS molecular orbitals of formamide as a result of the CAS(10,8) SA8 calculation.

As a result of these studies it can be seen that not only is the character of the active space important to state averaged calculations, but also the number of states included in the calculation. This representation of the results of the state averaged CAS calculations serves both as a tool for the selection of the active space size and the number of excited states, and as a demonstration that the analysis of

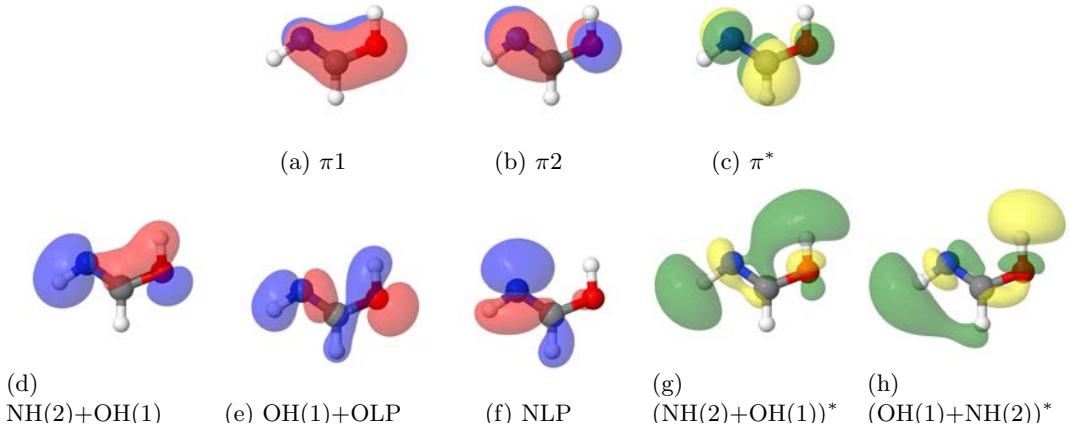


Fig. 4.6: CAS molecular orbitals of formimidic acid from the CAS(10,8) SA8 calculation.

the numerical and pictographic results of CAS calculations are intrinsically linked.

In Figures 4.5 and 4.6 the pictures of the CAS(10,8) SA8 active space orbitals of FAM and FIM can be seen, with the accompanying labels as seen in the state averaged CAS characterisations of Figures 4.3 and 4.4, respectively.

Vibrational Frequencies

As stated in Chapter 3.3.1, calculations using the DD-vMCG method require either the Cartesian coordinates or the normal modes of a system to be explicitly defined. A harmonic approximation is then taken using these coordinates or modes in order to provide the ground state potential energy surface. Additionally, they are also used to define the width of the total wavepacket, and by extension the width of the fully populated Gaussian basis function when the calculation is initiated.

It was decided that the normal modes would be used in the study of FAM and FIM, and hence a characterisation of the normal modes is required.

In Table 4.7 the normal modes of formamide, calculated at the 6-311+G$*$ CCSD, the 6-31G$*$ MP2 and the 6-31G$*$ CAS(10,8) SA8, are characterised accompanied by the relevant frequencies. As the Direct Dynamics calculations are to be carried out using CAS(10,8) SA8 externally calculated points, Table 4.9 contains a pictographic representation of each of these modes.

Label	Frequency (cm^{-1})			Character
	MP2 6-31G*	CCSD 6-311+G*	CAS(10,8) 6-31G* SA8	
$\nu_1(A'')$	137.48i	350.35i	320.66i	H1 H2 oop in-phase wagging
$\nu_2(A')$	566.17	576.24	598.06	NH_2(H1 H2) ip rocking
$\nu_3(A'')$	652.79	619.77	601.74	H1 H2 oop rocking
$\nu_4(A'')$	1055.40	1045.95	1014.98	H3 oop wagging
$\nu_5(A')$	1072.42	1079.99	1108.68	H2 ip rocking
$\nu_6(A')$	1309.13	1293.11	1356.26	H3 H1 ip out-of-phase rocking
$\nu_7(A')$	1455.15	1443.86	1397.43	H3 ip wagging
$\nu_8(A')$	1652.90	1680.88	1803.58	H1 H2 ip out-of-phase wagging
$\nu_9(A')$	1837.98	1827.88	2031.85	C H1 ip in-phase wagging
$\nu_{10}(A')$	3048.50	3008.20	2640.00	H3 ip stretching
$\nu_{11}(A')$	3635.63	3636.12	3855.27	H1 H2 ip in-phase stretching
$\nu_{12}(A')$	3777.92	3772.01	4029.16	H1 H2 ip out-of-phase stretching

Table 4.7: The normal mode frequencies of formamide calculated at different levels of theory at the optimised C_s structure where the numbering of the hydrogen atoms is as in Figure 4.1a), "ip" means in-plane and "oop" means out-of-plane.

In the first row of Table 4.7 it can be seen that the normal mode, ν_1, is an imaginary frequency. This is likely due to the fact that the starting structure, as discussed in Ground State Calculations section (Chapter 4.2.1) is not the optimised ground state minimum structure of FAM but the optimised ground state planar minimum structure. Comparing the MP2 and CCSD frequencies it can be seen that the CCSD imaginary frequency is significantly higher than the frequency calculated at the MP2 level. Further to this it can be seen that the frequencies calculated at these levels of theory are similar with the exceptions of ν_3, ν_8 and ν_{11} where, respectively, the MP2 is higher by 33 cm^{-1}, the CCSD is higher by approximately 28 cm^{-1} and the MP2 is higher by approximately 40 cm^{-1}.

Comparing the CAS(10,8) frequencies to the MP2 and CCSD calculated frequencies shows that the only trend is that the CAS(10,8) frequencies are more extreme than the other two sets of frequencies.

As the structure has C_s symmetry it can be seen that there exist two categories of normal modes representing in-plane and out-of-plane motion where only three

Label	Frequency (cm^{-1}) MP2 6-31G*	CCSD 6-311+G*	CAS(10,8) 6-31G* SA8	Character
$\nu_1(A'')$	590.72	596.58	630.43	H1 oop wagging
$\nu_2(A')$	650.36	611.55	635.19	NH(2) OH(1) ip in-phase wagging
$\nu_3(A'')$	850.84	846.36	911.91	H2 H3 oop in-phase wagging
$\nu_4(A'')$	1060.57	1066.07	1062.53	H3 H2 oop out-of-phase wagging
$\nu_5(A')$	1096.17	1097.61	1150.16	H2 ip wagging
$\nu_6(A')$	1217.85	1232.28	1284.95	H1 H3 in-phase, H2 out-of-phase, ip wagging
$\nu_7(A')$	1406.83	1404.31	1447.58	H1 H2 in-phase ip wagging
$\nu_8(A')$	1435.46	1437.46	1515.33	H3 ip wagging
$\nu_9(A')$	1750.28	1750.26	1854.47	CH(3) ip rocking
$\nu_{10}(A')$	3201.31	3155.30	3126.07	H2 stretching
$\nu_{11}(A')$	3551.66	3554.58	3379.82	H1 (H3) stretching
$\nu_{12}(A')$	3682.92	3776.70	3449.48	H3 (H1) stretching

Table 4.8: The normal mode frequencies of formimidic acid calculated at different levels of theory at the optimised C_s structure where the numbering of the hydrogen atoms is as in Figure 4.1b), "ip" means in-plane and "oop" means out-of-plane.

modes, ν_1, ν_3 and ν_4, represent this out-of-plane motion. It should be noted that despite the fact that motion of H(2) and H(3) appear to have their "own" vibrational modes, the motion of H(1) is always coupled to the motion of other atoms.

The pictographic representation of the CAS(10,8) calculations can be seen in Table 4.9. It is useful to include this representation of the normal modes in order to aid in the interpretation of the notation used in Table 4.7. It is also useful to include this representation of the normal modes in order for the later Direct Dynamics results to be more easily interpreted.

In Table 4.8 the normal modes of formimidic acid, calculated at the 6-311+G∗ CCSD, the 6-31G∗ MP2 and the 6-31G∗ CAS(10,8) SA8, are characterised accompanied by the relevant frequencies. In this table it can be seen that there are no imaginary modes present as the optimised ground state and optimised planar ground state of FIM are the same energy.

Table 4.9: Normal modes of formamide calculated at the CAS(10,8)/6-31G* SA8 level of theory at the optimised C_s structure. It should be noted that the arrow representing the CH stretching mode (ν_{10}) has been reduced in scale by 40%.

Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})
$\nu_1(A'')$		320.66i	$\nu_5(A')$		1108.68	$\nu_9(A')$		2031.85
$\nu_2(A')$		598.06	$\nu_6(A')$		1356.26	$\nu_{10}(A')$		2640.00
$\nu_3(A'')$		601.74	$\nu_7(A')$		1397.43	$\nu_{11}(A')$		3855.27
$\nu_4(A'')$		1014.98	$\nu_8(A')$		1803.58	$\nu_{12}(A')$		4029.16

Table 4.10: Normal modes of formimidic acid calculated at the CAS(10,8)/6-31G* SA8 level of theory at the optimised C_s structure. It should be noted that the arrow representing the OH stretching mode has been reduced in scale by 40%.

Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})
$\nu_1(A'')$		630.43	$\nu_5(A')$		1150.16	$\nu_9(A')$		1854.47
$\nu_2(A')$		635.19	$\nu_6(A')$		1284.95	$\nu_{10}(A')$		3126.07
$\nu_3(A'')$		911.91	$\nu_7(A')$		1447.58	$\nu_{11}(A')$		3379.82
$\nu_4(A'')$		1062.53	$\nu_8(A')$		1515.33	$\nu_{12}(A')$		3449.48

With the exceptions of ν_2, ν_{10} and ν_{12}, the vibrational frequencies calculated at the MP2 and CCSD levels of theory are much more similar than is seen in the same comparison of FAM. The three large frequency differences in FIM are comparatively much larger than those of FAM ν_2, ν_{10} and ν_{12}, where in ν_2 the MP2 is higher by approximately 39 cm^{-1}, in ν_{10} the MP2 is higher by approximately 46 cm^{-1} and in ν_{12} the CCSD is higher by approximately 94 cm^{-1}.

Comparing the CAS(10,8) frequencies to the MP2 and CCSD calculated frequencies shows that while the frequencies of ν_2 and ν_4 lie between the frequencies of the MP2 and CCSD calculations, modes ν_9 and ν_{13} are over 100 cm^{-1} higher than the MP2 and CCSD calculated frequencies.

As with FAM, it can be seen that there exist two categories of normal modes representing in-plane and out-of-plane motion where only three modes, ν_1, ν_3 and ν_4, represent this out-of-plane motion.

The pictographic representation of the CAS(10,8) calculations can be seen in Table 4.10. It is useful to include this representation of the normal modes in order for the later Direct Dynamics results to be more easily interpreted.

4.3 Direct-Dynamics

As the relevant analysis of the preliminary electronic structure calculations of formamide and formimidic acid have been completed, it is now necessary to describe the protocol by which the Direct Dynamics results were obtained.

4.3.1 The Direct-Dynamics Protocol

As a result of the exploration of the electronic structure of formamide and formimidic acid the Direct Dynamics calculations were carried out with a CAS(10,8) active space using equal weight on 8 states in total, which is the S_0 to S_7 states. A preliminary set of Direct Dynamics calculations were performed, initiated on the

S_1, S_2, S_3 and S_7 states. These calculations were propagated using the ground state normal modes of the systems (Tables 4.7 and 4.9 for FAM, 4.8 and 4.10 for FIM) from which the frequencies were used to define the widths of the Gaussian wavepackets. The initial distribution of the Gaussians comprises a single fully populated Gaussian at the Frank-Condon point around which the rest of the, zero populated, Gaussians are distributed, with an overlap of 0.5 (exactly 50% overlap). In order to ensure the stability of the initial electronic structure calculation the Gaussians were distributed in momentum space, as opposed to configuration space. The use of the single fully populated Gaussian at this point ensures that the initial wavepacket is the ground state vibrational wavefunction. As there were 8 states included in the preliminary electronic structure calculation, the initial calculation included 8 Gaussian basis functions, with a predefined propagation time of 150 fs.

An initial analysis of these results allowed the selection and placement of complex absorbing potentials (CAP(s)). A CAP, as defined and implemented in the MCTDH software package, is a negative, imaginary potential which is used to absorb a wavepacket. In grid based calculations, they are placed at the ends of grids to ensure that the wavepacket is not reflected off the end of the grid, resulting in the decoherence. In a Direct Dynamics calculation, if a rapid dissociation occurs, the dissociating atom or fragment gains momentum as it gets further away from the main molecule. This has a direct impact on the integrator in that rapidly decreasing time step sizes must be taken in order to gain a valid description of the system as a whole. Additional issues arise in that these rapidly changing geometries result in a larger number of points requiring electronic structure calculations. At these widely spaced geometries, the electronic structure calculations will take longer to run, and may fail. Consequently, the use of CAPs in the DD-vMCG method essentially provides a cut-off point to excessive motion along normal mode

coordinates.

The CAPs are defined as $-iW$, where

$$W = \eta\,\Theta(k(x - x_0))^n \tag{4.3}$$

Here, Θ is a Heaviside step function while the values of η, k, x_0 and n are defined by parameters in the input of the calculation. These parameters, ordered $[x_0, \eta, n, k]$, define the order, n, and strength, η, of the CAP positioned in this case at x_0 along a normal mode where $k = \pm 1$ indicates whether it is in the positive or negative direction. When the Gaussian reaches the CAP the motion continues classically until the population of that Gaussian is zero, after which it stops.

As mentioned in Chapter 3.3.1 the DD-vMCG method builds a database of geometries, energies and other information such as molecular orbitals along the course of a propagation. Multiple runs of a calculation, using the database constructed from the previous calculation(s), will result in fewer electronic structure calculations required, as well as the opportunity for further regions of configuration space to be explored. Consequently, the databases from this first set of calculations were used for subsequent calculations.

The related databases of the previous calculations, and with the relevantly defined CAPs, were then used in a second set of calculations propagated on the S_1, S_2, S_3 and S_7 states. Using the same CAPs and databases, in addition to the re-running of the 8 Gaussian propagations, sets of 24 and 48 Gaussian calculations were performed, each with a propagation time of 150 fs.

The databases constructed from the second set of propagations over the 4 states were then merged into one database which was then used for each of the final propagations on the four states, again using 8, 24 and 48 Gaussian wavepackets. As a sense of scale, each of the second set of calculations, at the end of the propagation, had databases comprising approximately 1500 points. The single, merged

database from all of the second set of calculations comprised approximately 6000 database points. The databases at the end of the final set of calculations contained approximately 6500 points.

It is the results of this third set of calculations that are used for the following analysis.

4.3.2 State Population Analysis

In Figures 4.7 and 4.9 the state populations results of formamide and formimidic acid, respectively, for each of the 8, 24 and 48 GWP propagations on the S_1, S_2, S_3 and S_7 states are shown. In Figures 4.8 and 4.10 the details of the results of the first 50 fs and 30 fs, respectively, of the 48 GWP propagations are shown, with a comparison to the total density. The state populations give information as to how the total wavepacket has been distributed, along the period of propagation, into the various states included in the calculation, as well as an idea of the performance of the calculation.

In Figure 4.7 it can be seen, most significantly in the S_1 plots, that the increase from 8 to 24 GWPs resulted in a more smooth representation of the decay and redistributions of the density into the different states, while the improvement from the 24 to 48 GWP is less significantly pronounced. This implies that the calculations are reaching convergence. It can also be seen that in the S_1 and S_7 48 GWP propagations the calculations have not yet reached a full propagation time. Although the increase in the number of GWPs presents a greater challenge to the Direct Dynamics code, it also represents a larger amount of external quantum chemistry calculations required, so it is not, as yet, possible to confirm with complete certainty as to whether the Direct Dynamics code is failing or taking a long time to read the database, or if the quantum chemistry calculations are failing, or taking a long time. It can also be seen in Figure 4.7 that the timescale for the decay from the S_1 state is significantly longer, approximately 20% of the

population is still in the S_1 state after 60 fs, than the timescale for the decay from the other states, approximately 80% decay has occurred within the first 10 fs. Upon inspection of the decay of the S_7 state it can be seen that the increase to 24 and 48 GWPs resulted in a slowing of the decay after 10 fs, a feature not observed in the 8 GWP calculation. What is clear from this set of results is that in all states at least a small amount of population is transferred to the other states of the system, even when the calculation is propagated on the S_1 state.

It should be noted, however, that the amount of the population absorbed by the CAPs cannot directly be separated from the population transfer in the analysis of this set of graphs. However, by also plotting the total density with the state populations the features of population transfer can more easily be resolved from the state population that is being absorbed by the CAP.

In Figure 4.8 the first 50 fs of the 48 GWP propagations are shown, with the change in the total density with time, the thick black line, also plotted. It can be seen that in the S_1 and S_7 states, density begins to flow into the CAPs after around 20 fs, while in the S_2 this begins to occur after 10 fs and in the S_3 this occurs after around 8 fs, indicating that the initial drop in the population at the beginning of the propagation is due to population transfer.

In the S_1 state, as mentioned previously, the rate of population transfer is significantly slower than in the other excited states. It can also be seen that the population transfer is relatively evenly distributed across all of the excited states in the calculation.

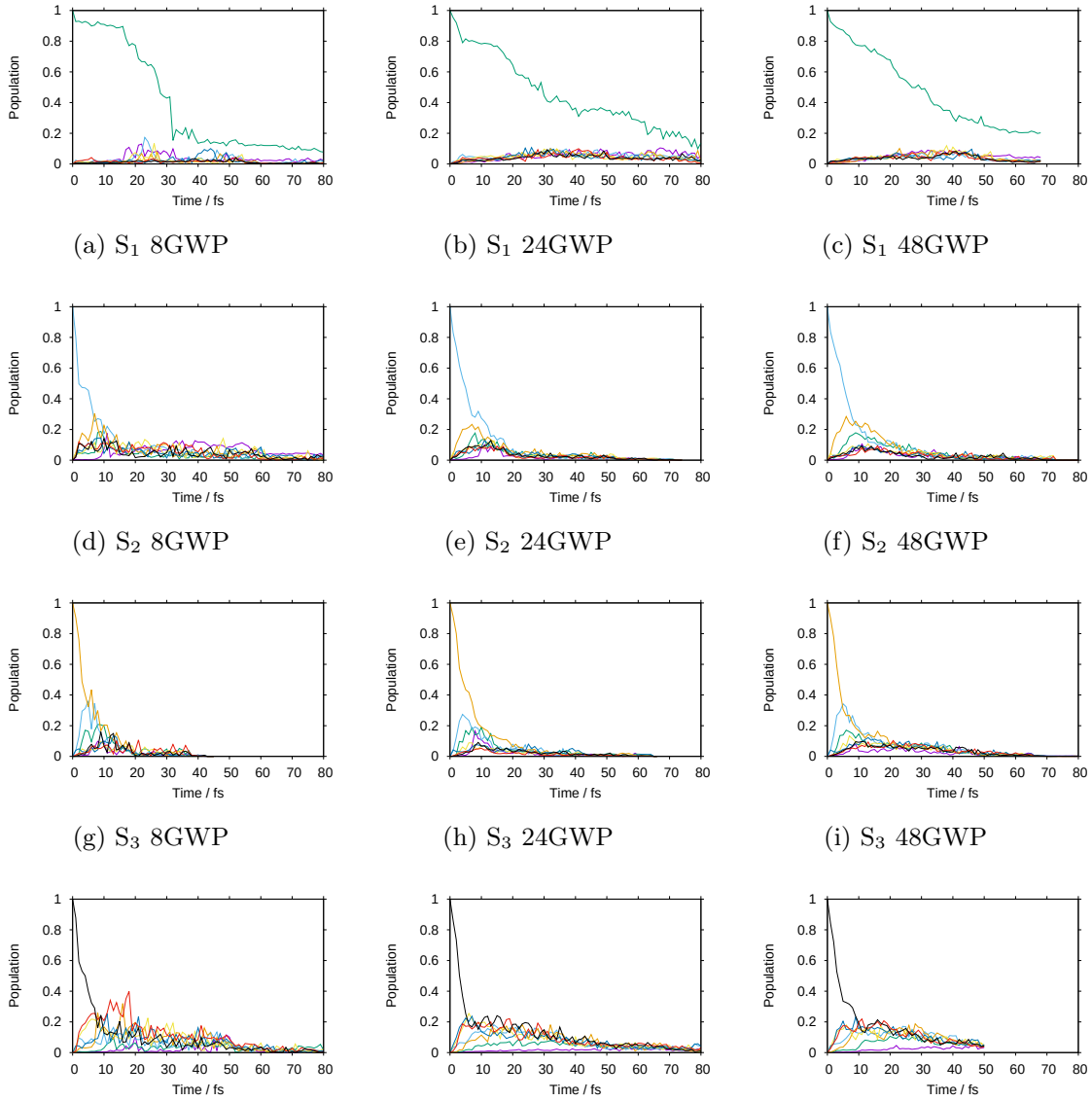
In the S_2 state it can be seen that after approximately 6 fs about 25% of the population has been transferred to the S_3 (orange) state, with around 18% of the population being transferred to the S_1 (green) state by 9 fs. The reverse of this fast population transfer is also seen in the S_3 plot, where about 32% of the population is transferred to the S_2 (light blue) state with the next most significant population

transfer going into the S_1 state. In the S_2 state it can be seen that by about 11 fs the S_0(purple), S_5(dark blue), S_6(red) and S_7(black) states are equivalently populated. It can also be seen that a greater amount of population transfer takes place going from the S_3 to S_4 (yellow) state, a feature not apparent in the S_2 result.

In the S_7 state it can be seen that the population transfer initially populates the S_6, S_5 and S_4 states equivalently. At about 11 fs the population of the S_7 to S_2 are approximately equivalent. It can also be seen that while the S_0 state does not experience any fast population transfer, the population transferring to the S_0 state consistently increases across the full propagation.

In Figure 4.9 the "smoothing" effect of the increase in the number of GWPs is clearly seen in the S_1 and S_2 states of formimidic acid. In this case not only does the increase from 8 to 24 GWPs show an improvement, but the 24 to the 48 GWPs also shows an improvement, demonstrating that the inclusion of this higher number of basis functions representing the system has resulted in a better realisation of the results. Although the S_7 48 GWP calculation has not reached full propagation it can be seen in the 8 and 24 GWP results that the majority of the dynamical behaviour occurs in the first 8 fs. This rapid decrease in dynamical behaviour is also observed in the S_2 and S_3 states. As is seen in the state population data for FAM the S_1 state of FIM displays a much longer timescale for the decay of population in comparison to the excited states of the system, though in comparison to the decay of the S_1 state of FAM the rate of decay is faster.

It should also be noted that the timescales for population of the other states in the S_1 propagation also changes with increasing number of GWPs. Between 15 and 25 fs in the 8 GWP calculation the majority of the other states in the calculation become populated, while in the 24 GWP calculation the other states become populated between 10 and 15 fs, with the principal population transfer into the S_2 state. In the 48 GWP calculation the most significant population



(a) S_1 8GWP (b) S_1 24GWP (c) S_1 48GWP

(d) S_2 8GWP (e) S_2 24GWP (f) S_2 48GWP

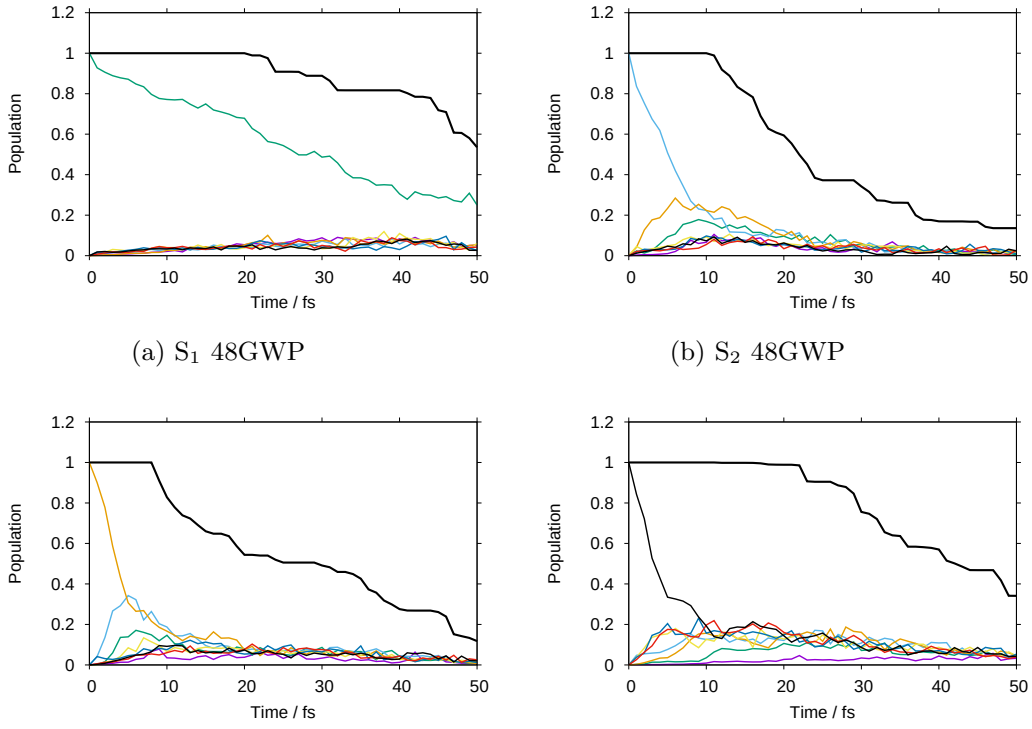
(g) S_3 8GWP (h) S_3 24GWP (i) S_3 48GWP

(j) S_7 8GWP (k) S_7 24GWP (l) S_7 48GWP

Fig. 4.7: Diabatic state populations from DD-vMCG simulations of formamide starting with a vertical excitation to various states and using various numbers of GWPs.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black.

transfer is into the S_0 state in two stages at about 12 and 15 fs with the increase in the population of the other states occurring more gradually over the course of the propagation.

In Figure 4.10 the first 30 fs of the 48 GWP propagations are shown with the addition of the thick black line representing the total density of the system. It

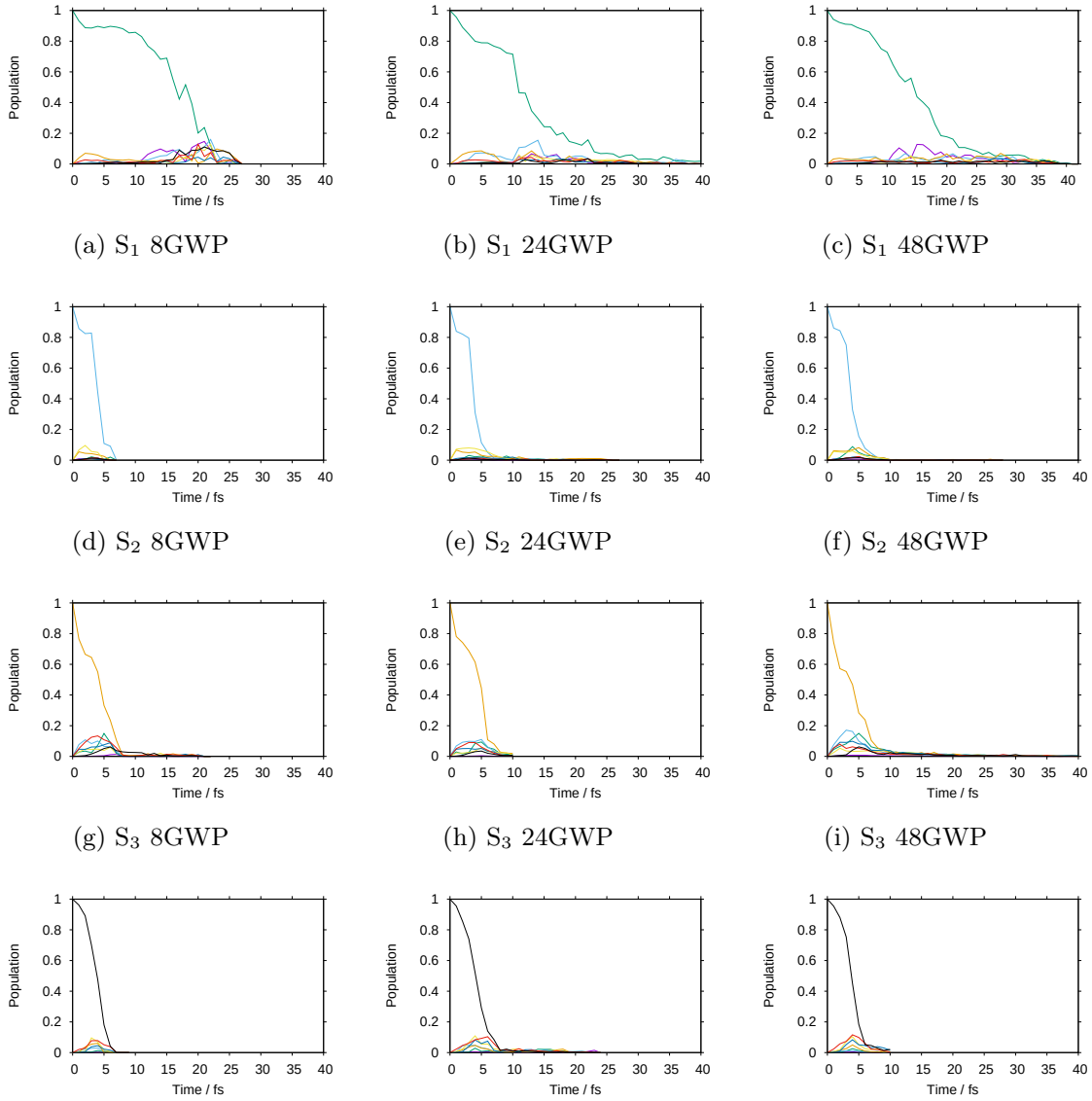


(a) S_1 48GWP (b) S_2 48GWP

(c) S_3 48GWP (d) S_7 48GWP

Fig. 4.8: Diabatic state populations from DD-vMCG simulations of formamide starting with a vertical excitation to various states. Final Results.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (norm^2): Thick black.

can be seen that in the S_1 state the decrease in total density follows the trend in the decreasing population of the S_1 state suggesting that the S_1 state is displaying some dissociative behaviour. In the S_2 state the decrease in the total density occurs rapidly between 4 and 10 fs with relatively little population transfer occurring. This rapid decrease coincides with a decrease in population of the S_3 and S_4 and an increase in population of the S_1 state, though due to the small increase it is not possible to state that it is the population of these states that led to the rapid decrease in the density. In the S_3 state it can be seen that there is a fast population transfer mostly to the S_2, but also to the S_5, S_6 and a small amount to



(a) S_1 8GWP (b) S_1 24GWP (c) S_1 48GWP

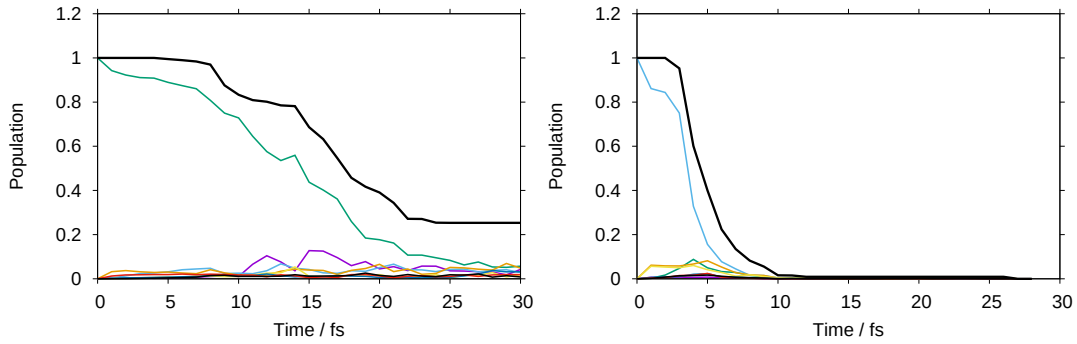
(d) S_2 8GWP (e) S_2 24GWP (f) S_2 48GWP

(g) S_3 8GWP (h) S_3 24GWP (i) S_3 48GWP

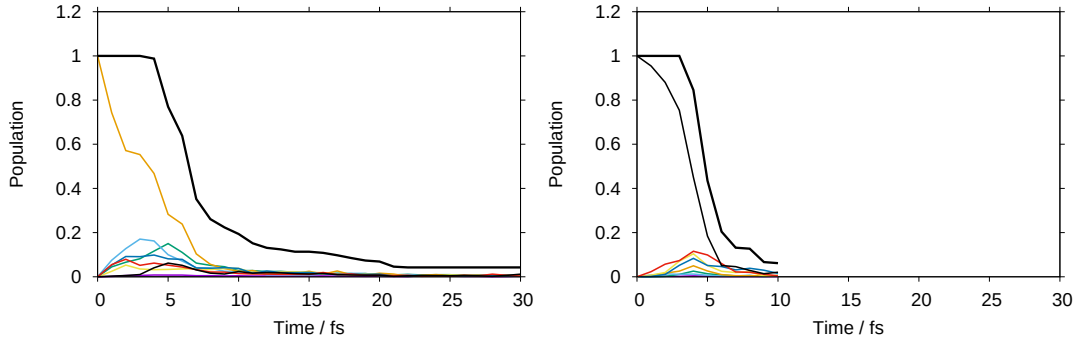
(j) S_7 8GWP (k) S_7 24GWP (l) S_7 48GWP

Fig. 4.9: Diabatic state populations from DD-vMCG simulations of formimidicacid starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black.

the S_4 states, and the total density does not decrease over this short timescale. It appears that in this S_3 calculation when the S_1 and S_7 states become populated, the total density proceeds to decrease rapidly, supporting the suggestion that the S_1, and perhaps the S_7, states are dissociative. Although the propagation on the S_7 state is curtailed at 10 fs it can be seen that there is also a rapid decrease in



(a) S_1 48GWP (b) S_2 48GWP



(c) S_3 48GWP (d) S_7 48GWP

Fig. 4.10: Diabatic state populations from DD-vMCG simulations of formimidicacid starting with a vertical excitation to various states. Final Results.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; S_7: black. Total density (norm^2): Thick black.

the total density starting at about 4 fs. At this point, however, the S_2 to S_6 states have become populated and it is not possible to resolve the behaviour further.

As a result of this analysis it can be seen that while dissociative behaviour can be seen in all sets of results, it is not possible from this analysis to state conclusively that particular states exhibit dissociative behaviours. Consequently further analysis is required in order to resolve these observed behaviours.

4.3.3 Gross Gaussian Populations

Following from the observation, in the previous section, of the density into the manifold of states, a useful description can be obtained if the distribution of the

products of the calculation are quantified. Consequently there is a requirement for the definition of the importance, or weight, a GWP has in the description of a total wavepacket. If an orthonormal basis is used,

$$\Psi(r,t) = \sum_i A_i \phi_i \tag{4.4}$$

an expression for the density is given as

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \sum_{ij} A_i^* \langle \phi_i | \phi_j \rangle A_j & (4.5) \\ &= \sum_i |A_i|^2 & (4.6) \end{aligned}$$

so the "population" for each basis function can be defined as

$$\mathrm{pop}_i = |A_i|^2 \tag{4.7}$$

If the basis functions are normalised but not orthogonal,

$$\Psi(r,t) = \sum_i A_i g_i \tag{4.8}$$

the overlaps of the basis functions now play a role

$$\begin{aligned} \langle \Psi | \Psi \rangle &= \sum_{ij} A_i^* \langle g_i | g_j \rangle A_j & (4.9) \\ &= \sum_{ij} A_i^* S_{ij} A_j & (4.10) \\ &= \sum_i |A_i|^2 + \sum_{i<j} 2 Re S_{ij} & (4.11) \end{aligned}$$

When the overlaps between the functions are divided evenly, the *Gross Gaussian Population* [122], or GGP, is hence defined where

$$\begin{aligned} \mathrm{GGP}_i &= |A_i|^2 + \sum_{j \neq i} Re A_i^* S_{ij} A_j & (4.12) \\ &= \sum_j Re A_i^* S_{ij} A_j & (4.13) \end{aligned}$$

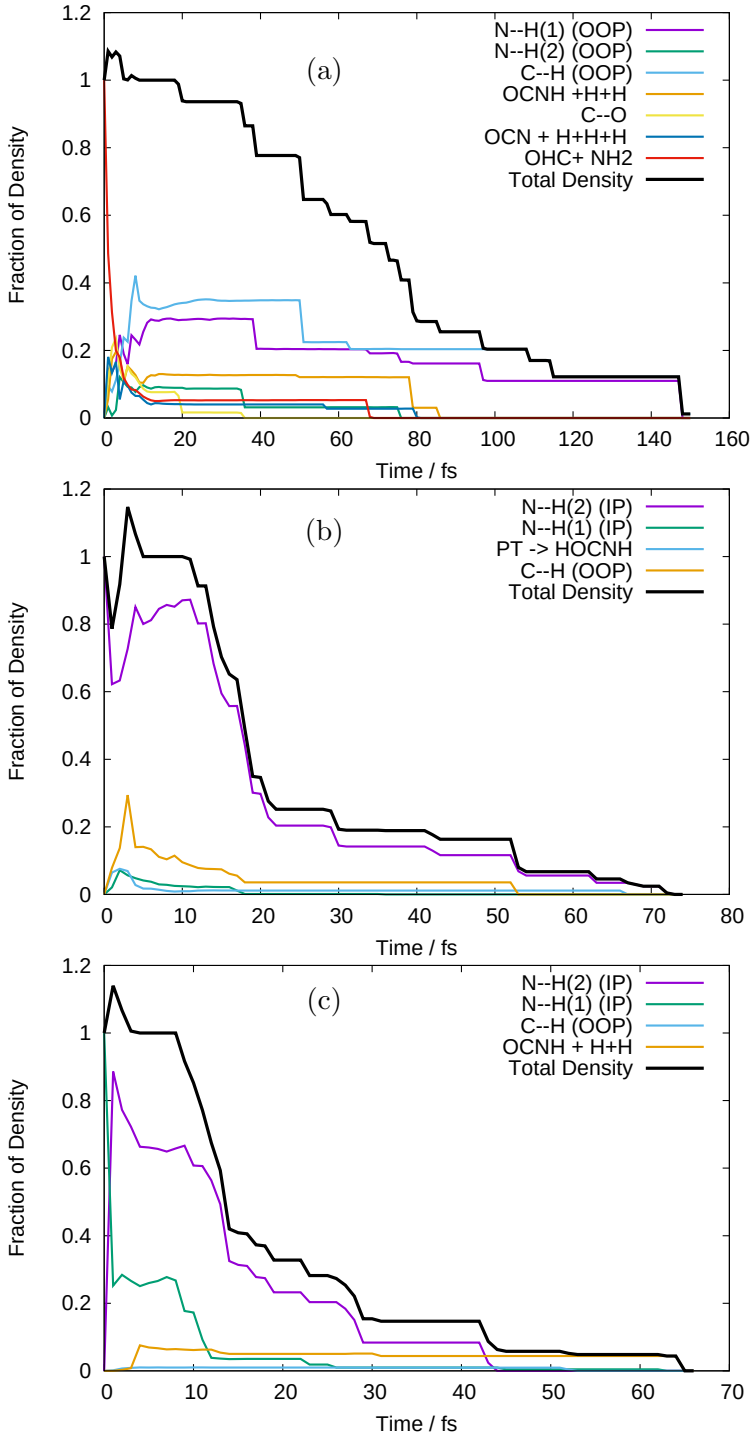


Fig. 4.11: The fraction of density going into different product channels from DD-vMCG simulations of formamide with SA8-CAS(10,8)/6-31G∗ following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3. Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies, IP and OOP signify if the dissociation occurs of in- or out-of-plane.

In order to use these GGPs it is first necessary to perform a visual inspection of the geometries defined by the trajectory of the centre of the GWPs. The fraction of the density, defined by the GGP, along each of the product channels were then categorised allowing analysis of the product distribution. The results of this analysis of the trajectories of the GWPs for formamide and formimidic acid can be seen in Figures 4.11 and 4.12 respectively. It was found that the product distribution of the S_7 state for both FAM and FIM displayed little behaviour of interest, in part due to fast decay, so only the results of the S_1, S_2 and S_3 are included in this analysis. It should be noted that the > 1 population in the first few femtoseconds of the propagation is likely due to numerical instabilities in the method.

In Figure 4.11, representing the product distribution of FAM, upon cursory inspection it can be seen that while the product distribution in the S_2 and S_3 states is dominated by one behaviour, the number of product channels, and hence the product distribution in the S_1 state, is more evenly distributed between the possible pathways. In the S_1 state, it can be seen that dissociation in the first 70 to 80 fs of the propagation occurs along a multitude of pathways, the principal of which represent the N-H(1) and C-H oop motions, which correspond to the ν_1, ν_2 and ν_3 vibrational modes of the system. It is in the movement in these channels that the density continues after 80 fs, until all of the Gaussians have been absorbed by the CAPs. In both the S_2 and the S_3 states the principal product channel is defined by the N-H(2) bond stretching. In the S_2 state it can be seen that a very small amount of proton transfer occurs. In this process, the H(1) dissociates and the H(3) subsequently transfers to the oxygen. Although the proton transfer involving the H(1) atom was expected, this mechanism of proton transfer was somewhat unexpected. In the S_3 state it can be seen that a secondary product channel to the nuclear motion in the first 12 fs is observed along a N-H(1)

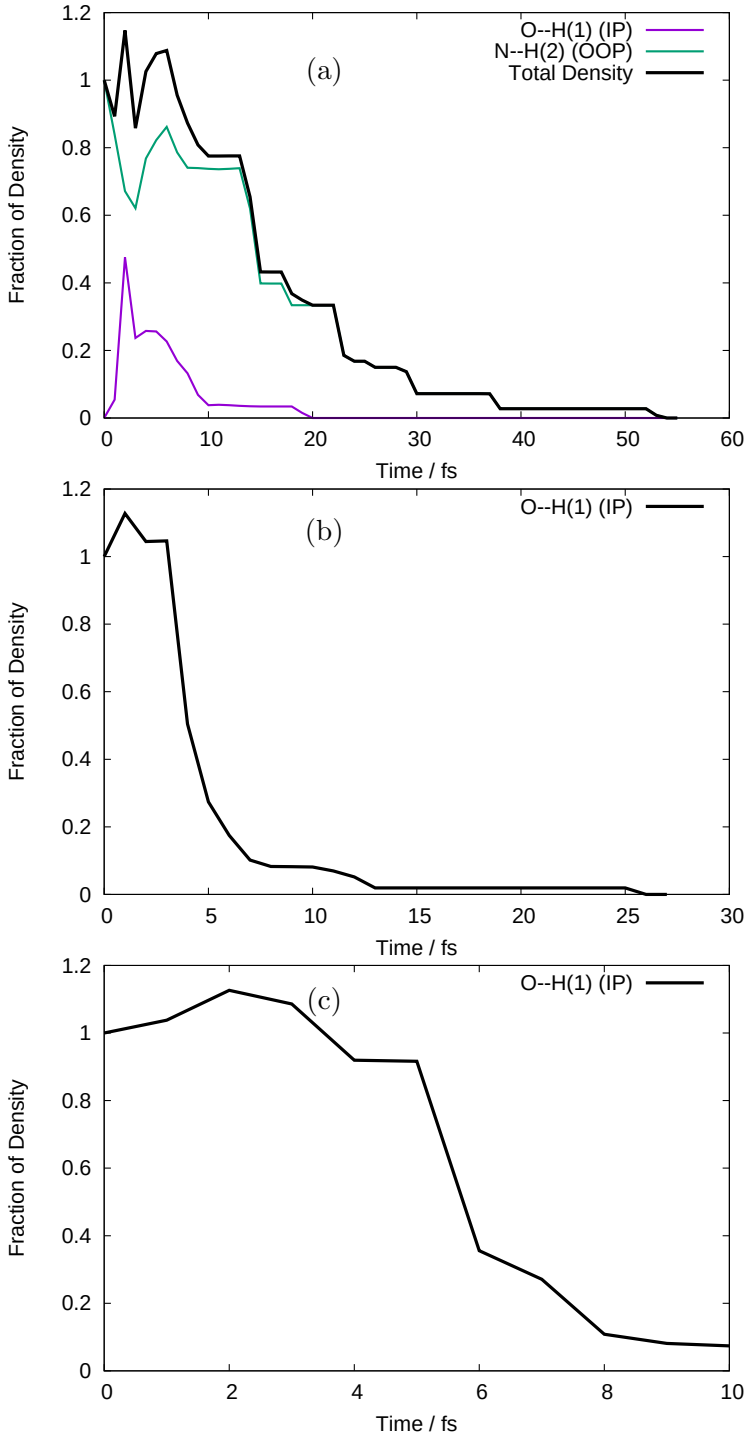


Fig. 4.12: The fraction of density going into different product channels from DD-vMCG simulations of formimidic acid with SA8-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3 Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies, IP and OOP signify if the dissociation occurs of in- or out-of-plane.

stretching mode. It should be noted that while the C-H out-of-plane motion is the only motion observed in all three of these states, significant NH_2 character is also observed in all three states, both in the in-plane and out-of-plane motions.

In Figure 4.12, representing the product distribution of FIM, upon cursory inspection it can be seen that while the number of product channels for the S_1 is greater than in the S_2 and S_3 states, as was the case with FAM, only two product channels are seen. In all three states the O-H(1) in-plane stretching motion, characterised as the ν_{11} mode, to the point of dissociation is seen, while in the S_1 the N-H(2) out-of-plane motion, characterised in a combination of the ν_3 and ν_4 modes, is also seen. In the S_1 state it is the N-H(2) out-of-plane motion that is dominant, though the timescale over which the O-H(1) stretching motion is seen is comparable to the timescale in the S_2 state.

As a result of this investigation it can be seen that a number of different product channels are available in the dynamics of formamide, whereas the OH stretching motion in formimidic acid is dominant.

4.3.4 Potential Energy Surfaces

During the course of a Direct Dynamics propagation the energies are calculated and stored in the database. As a result 1- and 2-dimensional cuts of the potential energy surfaces can be made along the coordinates of the input. As the LHA is used the surfaces should be symmetrical around the central point of the mode. By looking at the adiabatic and diabatic representation of the same mode, or modes, regions displaying non-adiabatic features such as avoided crossings, or conical intersections, can be determined. From a practical perspective these surfaces can also be used to determine if the Direct Dynamics calculations have been propagated for sufficient time periods, if the diabatisation scheme has been successful and other technical faults which are displayed as discontinuities in the surfaces. Using the observations of the GGPs the surface cuts were selected along the modes

which were most clearly represented by the density of the system.

In Figure 4.13 the adiabatic and diabatic cuts of the potential energy surface along the ν_{11}(a) and b)) and ν_{12}(c) and d)) modes of formamide, characterised as the N-H_2 symmetric and anti-symmetric stretching modes, are shown. In a 2-dimensional representation, or combination, the NH dissociations occur at 45^o to the plane of these modes. It is immediately clear that the potential energy surfaces are smooth, with the diabatic states correctly representing crossings between states.

In Figures 4.13 a) and b) there appears to be a dissociative state in the negative displacement direction shown as the green, S_1 state in the adiabatic representation, which in the diabatic representation becomes the orange, previously S_3, state. However, in both the diabatic and adiabatic representations it can be seen that there exists a kink in the surfaces, approximately located at -2 displacement. From the shape of the curves at this point it is likely that there is a higher order state which cuts down through most of the states, and is not represented in this picture.

In Figures 4.13 c) and d) at the Frank-Condon point it can be seen that the blue, S_2, and orange, S_3, states are very close in energy and crossing between these states occur very close to the Franck-Condon point. It is this orange state, in the diabatic picture, that represents the NH-stretching motion.

In Figure 4.14 the adiabatic and diabatic cuts of the potential energy surface along the ν_{10}(a) and b)) and ν_{11}(c) and d)) modes of formimidic acid, characterised as the N-H_2 and the OH stretching modes, are shown. These modes represent the equivalent modes to the ν_{11} and ν_{12} of FAM. In Figures 4.14 a) and b) there appears to a major failure in the calculation in the negative direction along the mode, at about -4. This has been identified as a failure in the electronic structure calculations. Consequently the only comments that can be made about this representation are that the dark blue (S_5) and yellow (S_4) states are very close in energy

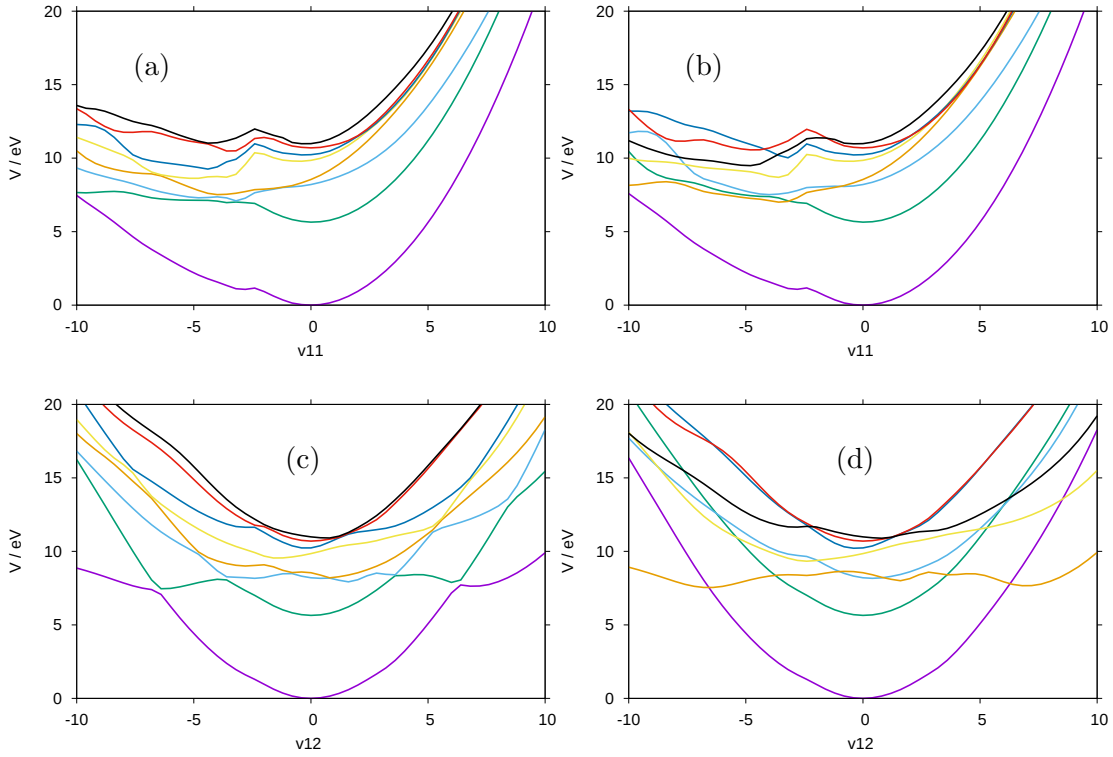


Fig. 4.13: Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formamide from DD-vMCG simulations . (a) ν_{11} (N-H_2 symmetric stretch) adiabatic (b) ν_{11} diabatic (c) ν_{12} (N-H_2 antisymmetric stretch) adiabatic (d) ν_{12} diabatic.

along the positive direction of the mode and that the crossings between these and the red (S_6) and black (S_7) states are well resolved. It appears as if there may be a dissociative state in the negative direction along the mode though it is unclear as to whether this state originates from a state within the precalculated manifold.

In Figures 4.14 c) and d) it can be seen that the surfaces are much more smooth and they clearly show a dissociative state which is the green, S_1, state in the adiabatic representation, or the light blue, S_2, state in the diabatic representation. As the states are well resolved in this mode it is clear that the OH dissociative pathway was open to the dynamics of the molecule. In the diabatic representation there is a good representation of the crossings between the states, although a region approximately located at 2, in the positive direction shows a significant convergence of the yellow, blue, red and black, S_4 to S_7 , states and it is unclear if this region has been correctly resolved.

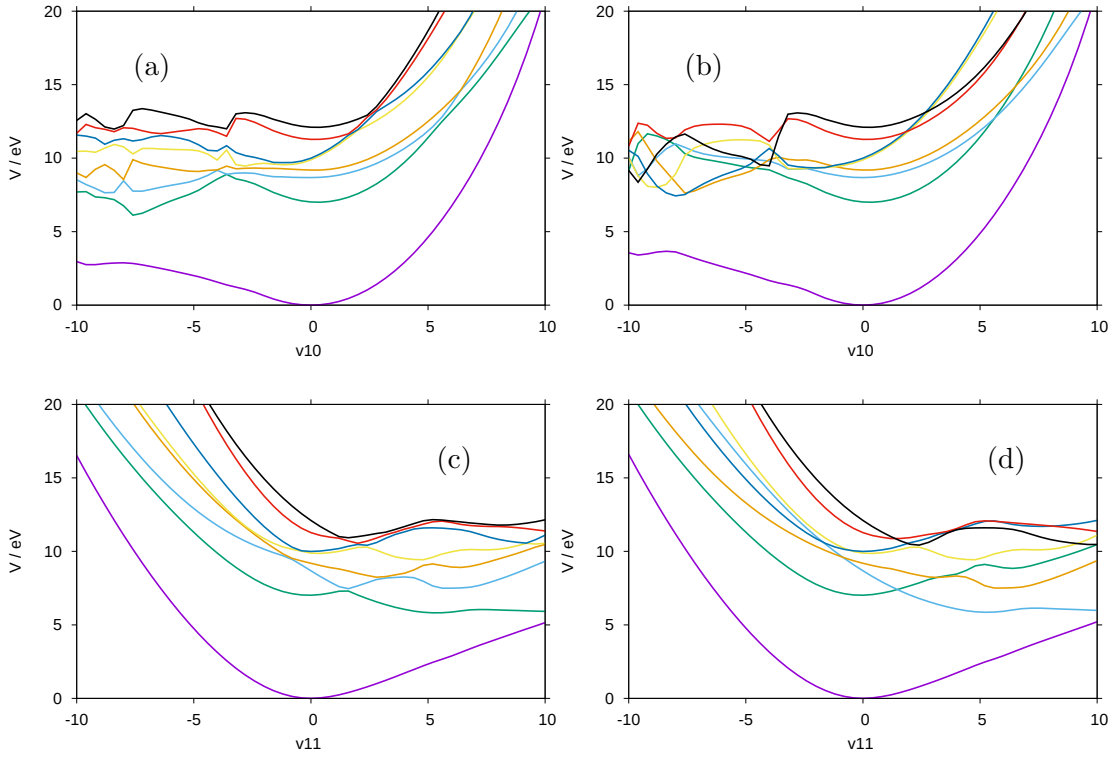


Fig. 4.14: Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of formimidicacid from DD-vMCG simulations . (a) ν_{10} (N-H(2) stretch) adiabatic (b) ν_{10} diabatic (c) ν_{11} (O-H(1) stretch) adiabatic (d) ν_{11} diabatic.

These potential energy surfaces provide useful information regarding the nature of the global potential along particular coordinates. In the case of FAM a combination of modes provide a relaxation pathway to the dissociation of the NH(1) bond. In the case of FIM it can be seen that a relaxation pathway to dissociation of the OH-bond crosses all but the S_0 state. Consequently this gives a direct example of where the efficient relaxation, leading to a dissociation is simplified when only 1 mode, as opposed to a combination of modes, exists.

4.4 Summary and Conclusions

In this Chapter quantum chemistry and quantum dynamics calculations were performed in order to study the excited-state dynamics of formamide and formimidic acid. Initial considerations for the representation of these two systems focused on testing the limitations of the DD-vMCG method, whilst maintaining a balance

between computational and time expense with accuracy and stability.

An initial ground state quantum chemistry study was carried out investigating the effect of the choice of basis set on the time taken for calculations in addition to an analysis of the choice of a planar versus a non-planar starting geometry. As expected, it was found that while an increase in accuracy of basis set (greater split valence, inclusion of diffuse and polarisation functions) results in more accurate results, the calculation times became significantly increased. In order to maintain resonable calulation times in the dynamics a compromise would have to be made, dependent upon the orbitals and excited states required. It was also found that the ground state, optimised geometry of formimidic acid and the transition state were energetically almost identical in the planar and non-planar geometries. While the difference in energy between the planar and non-planar optimised geometries of formamide was more significant it was still a negligible difference and it was decided that the planar geometry would be used.

An extensive investigation was then carried out analysing the choice of CAS size and composition, as well as the effect of the inclusion of varying numbers of excited states. Utilising a compromise 6-31G* basis set, it was found that the use of a CAS(10,8) over 8 averaged states yielded the best description of the system and the greatest stability for both formamide and formimidic acid. It was also found that at this level of theory, and basis set, optically bright π-π^* transitions were present in the main configurations of both formamide (S_6 principle, S_3 and S_7 secondary) and formimidic acid (S_3 principle, S_6 secondary). Additionally, a lone-pair - π^* transition is seen as the main configuration for the S_1 state of both systems.

As the DD-vMCG calculations were performed using the normal modes of the systems, an analysis of the character and numerical values of the vibrational frequencies of both systems was then carried out in order to identify significant

modes (such as those leading to fragmentation or potential proton transfer) with a comparison to higher and lower accuracy calculations. It was found that the 6-31G* SA8 CAS(10,8) calculations resulted in good agreement with the higher order (6-311+G* CCSD) results, especially in the case of the imaginary mode of formamide, ν_1.

The Direct Dynamics protocol was then established, including the initial population and numbers of the GWPs, the application of CAPs on particular long-range-motion modes, and the propagation time. Of particular importance is the process whereby the Direct Dynamics propagations were carried out in multiple iterations, each time constructing an increasing sized database of points for use in the following calculations.

Three types of analysis were carried out on the results of the final sets of Direct Dynamics calculations. An analysis of the state populations resulting from calculations utilising 8, 24 and 48 GWPs propagated on the S_1, S_2, S_3 and S_7 states over 150 fs for both formamide and formimidic acid was carried out. The results for both systems showed that an increase in the number of GWPs propagated, resulted in smoother curves and hence a better realisation of population transfer processes. A more detailed representation of the 48 GWP results, showing the total density as well as the state populations, were then analysed. The decrease in population of the S_1 state for propagations on the S_1 state, in both systems, exhibited much longer timescales than the propagations on the higher states. The results of the formamide calculations show that fast population transfer occurs to all states, particularly between the S_2 and S_3 states, though in the S_7 propagation the degree of population transfer to the S_0 state occurs more slowly. The reduction in the total density in the formamide results occurs after 15 fs in the S_2 and S_3, while in the S_1 and S_7 occurs after 20 fs, indicating that dissociative processes are less dominant in these simulations. The results of the formimidic acid calculations

generally show that the degree of population transfer is significantly less than in the formamide calculations. The propagations on the S_2, $S_3$and S_7 states display a significant degree of dissociation occurring between 5 and 10 fs. These results indicate that formamide is significantly more photostable in the S_2 and higher states than formimidic acid.

As a result of the reduction of the total density observed in the state population analysis, a visual inspection of the geometries defined by the trajectory at the centre of each GWP was carried out showing the distribution of the total density into differing product channels. The GWPs were then weighted according to their importance to the description of the total wavepacket.

This gross Gaussian population analysis showed that in the S_1 propagation of formamide the reduction of the total density was distributed into seven product channels representing a variety of fragmentations, the most significant of which represented the breaking of the N-H bonds. While the N-H(2) dissociation was dominant in both the S_2 and S_3 propagations, in the S_2 propagation a small degree of population went into a proton transfer of H(1) occurring from the nitrogen to the oxygen via the carbon. The gross Gaussian population analysis for the S_2 and S_3 propagations of formimidic acid shows only the breaking of the O-H bond, while in the S_1 state the dissociation of the N-H(2) dominates, with a lesser degree of dissociation of the O-H.

During the course of the Direct Dynamics propagation, the energies of the system in each of the states are stored and can hence be visualised by taking cuts along the global potential in either the diabatic or adiabatic representation along selected normal modes. The surfaces should be smooth, in that any discontinuities (kinks) in the surface may represent the failure of the DD-vMCG code (e.g. failure in the diabatisation), that other high lying states not included in the quantum chemistry calculations are crossing calculated states or that the quantum chemistry

calculations are failing and extrapolation is occurring.

The results of the gross Gaussian population analysis of formamide showed that the dominant motion was along the N-H stretching modes, ν_{11} and ν_{12}. The potential surfaces of ν_{12} (N-H_2 antisymmetric stretch) are symmetric and show that the S_3 (diabatic, S_1 adiabatic) state is dissociative. As expected, the surfaces along ν_{11} (N-H_2 symmetric stretch) are not symmetric around Q_0. Additionally, along ν_{11} mode, a discontinuity in the surface (around -2) indicates a higher lying, uncalculated, state may be crossing with the calculated states and the quantum chemistry calculations begin to fail. It should be noted that the channel for N-H dissociation occurs at 45^o to the 2-dimensional surface of these combined ν_{11} and ν_{12} modes. The potential energy surfaces of formimidic acid along the ν_{10} (N-H(2) stretch) and ν_{12} (O-H stretch) modes, are the only modes that exhibit the most dynamical behaviour and represent the analogue to the ν_{11} and ν_{12} of formamide. While neither of these cuts are symmetric, the surfaces along ν_{11} surfaces are relatively smooth and a relaxation pathway to dissociation crosses all but the S_0 state. The surfaces representing the ν_{10} mode are extremely discontinuous in the negative region. This level of disorder implies that there was a severe breakdown of the quantum chemistry calculations in this region, resulting in extrapolated points.

It can clearly be seen in both the process by which the CAS composition, size and number of states is chosen, as well as in the potential energy surfaces that the excited state dynamics of formamide and formimidic acid is difficult to resolve due to the number of states required and the close proximity of multiple states in large regions of the global potential. However, a variety of interesting features can be resolved from these results. While formamide has a relatively high degree of photostability in higher lying states, it also posseses a large number of fragmentation pathways (particularly in the S_1 propagation) including an indirect

pathway for proton transfer, though higher level calculations are required in order to be able elucidate the significance of this pathway. Formamidic acid, conversely, demonstrates a high degree photoactivity above the ground state, predominantly resulting in O-H dissociation due to efficient relaxation pathway into the dissociative state.

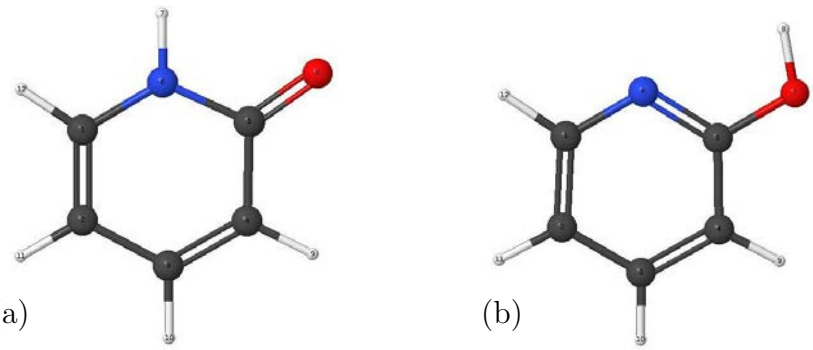
Chapter 5

2-Pyridone/2-Hydroxypyridine Proton Transfer

5.1 Introduction

The structural characteristics and photochemical behaviours of aromatic heterocycles and hydrogen-bonded complexes have been the subject of extensive chemical and biological research. One example, 2-pyridone, has been widely investigated as it possesses the same hydrogen-bonding sites as cytosine, thymine and uracil, [179–181] and exhibits a second stable tautomeric form, 2-hydroxypyridine, (Figure 5.1) reached via an enol-keto tautomerisation. [182–184]

2-pyridone belongs to a series of cyclic amides commonly referred to as lactams, the acid analogues of which are cyclic carboximidic acids commonly referred to as lactims. The series is defined by increasing ring size where a 4 membered ring



(a) (b)

Fig. 5.1: The ground state optimised structures of a) 2-hydroxypyridine (TIM) and b) 2-pyridone (TAM).

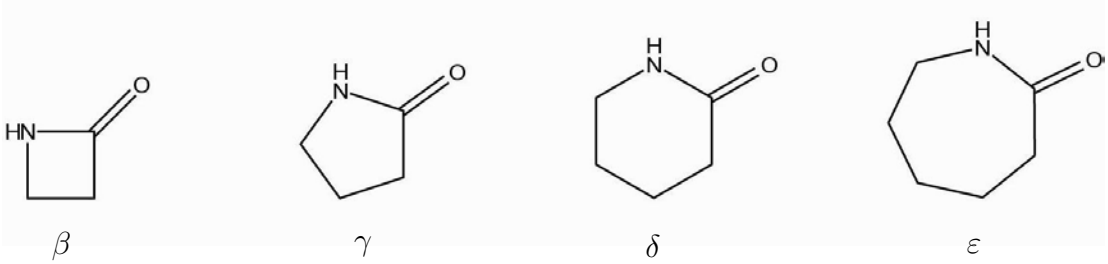


Fig. 5.2: The general structures of the lactam series ordered in increasing ring size, from β-lactam (4 membered ring) to ϵ-lactam (7 membered ring). 2-pyridone is a δ-lactam where the ring is conjugated.

is β-lactam, a 5 membered ring is γ-lactam, and so on, as seen in Figure 5.2. Lactams are a relatively versatile series of compounds involved in a wide variety of reaction processes in the form of catalysts, intermediates, and reactants. β-lactams have been widely used in antibiotics, though a number of these products have been phased out due to undesirable side effects and as many pathogens have become resistant to these medicines. [185] ε-lactam, more commonly referred to a caprolactam or naphtholactam, was used to produce the first spinnable polymer achieved through a polycondensation reaction [186] and has also been used in the development of dyes for cotton and wool, including lightfast violets and blues. [187, 188]

As only the δ-lactam, 6 membered ring, is to be investigated here, 2-pyridone will be referred to as lactam, or TAM, and its acid analogue 2-hydroxypyridine will be referred to as lactim, or TIM. The study of TIM/TAM and its analogous behaviour to that of Watson-Crick base-pairs, may provide essential information as to the causes of mutagenesis in DNA/RNA, whether as a result of mutagens, such as UV radiation and reactive oxygen species (ROS), [189] or by spontaneous occurrence. [179, 190, 191] The TIM/TAM tautomeric equilibrium at the ground state in gas phase has been shown to favour the TIM form by approximately 3 kJ mol^{-1}. [192, 193] Low level quantum chemistry calculations indicate that the barrier to proton transfer from TIM to TAM is between 142-159 kJ mol^{-1} [183,194], thus it can be stated that spontaneous proton transfer is unlikely at the ground

state.

At the S_1 electronic state TAM is reported to be more stable and the endothermicity of TAM (78 kJ mol^{-1}) [184, 194] suggests that after photoexcitation the barrier height to the enol-keto tautomerisation is much lower than in the ground, S_0, state. Consequently, it is thought that an excited state intramolecular proton transfer, ESIPT, may be possible in this state. [189] Although ESIPT processes are observed in many other DNA/base-pair analogous systems, low level calculations at the S_1 state suggest that the activation energy of the TAM to TIM tautomerism is 215 kJ mol^{-1} [195] or 148 kJ mol^{-1} [194], consequently conflicting reasoning and data make the previous statement inconclusive. It could be suggested that the mechanism for the TIM to TAM conversion may follow a roaming wavepacket dynamics schema similar to that of Poisson et al. [196].

In nature tautomeric equilibrium takes place in the presence of water molecules, or in the from of hydrogen bonded dimers. It is reported that a water solvent results in the S_0 TAM being stabilised by a further 12 kJ mol^{-1} relative to TIM, likely due to the polar solvent stabilising the zwitterion resonance structure of TAM, which does not exist for TIM. [180] Two reaction schemes for the tautomerisation of TIM/TAM in a water solvent have been proposed, either by a concerted mechanism involving a neutral cyclic intermediate, or via a cyclic ionic intermediate involving a proton transfer from TIM to the solvent which is subsequently transferred back. [184] These effects have been modelled spectroscopically and theoretically in three systems, monosolvated, disolvated and in solution. In the monosolvated system it has been found that the S_0 proton transfer barrier is reduced from the isolated molecule barrier height by 95 kJ mol^{-1} and in the disolvated by 35 kJ mol^{-1}. However, excited state experiments involving the singly and doubly hydrated TAM vary in the degree of stabilisation reported. Disolvated TAM is reported to be more stable in the S_1 state by between 78-82 kJ mol^{-1}, whereas the monosolvated TAM

is reported to be between 43-68 kJ mol^{-1} more stable than TIM. The barrier heights to the proton transfer in these systems have been calculated to be 66.9 kJ mol^{-1} for the monosolvated system and 109 kJ mol^{-1} for disolvated system, though the accuracy of these calculations is limited [184, 194]. Conclusive data as to the overall solvent effect of the water has yet to be obtained.

Although limited wavepacket dynamics calculations have been carried out in the study of TIM and TAM [196, 197], to date no studies have investigated the role of the S_3 and higher states. Due to the inconclusive data available, there is also a requirement for high level quantum chemical calculations to be carried out investigating the excited states of TIM and TAM.

It is clear, at this point, that a quantum dynamical study of TIM and TAM presents an ideal extension from the work of Chapter 4 not only due to the matching active sites, and limited number of previous dynamical studies, but also, due to the higher number of degrees of freedom, as these systems present a more demanding challenge for the DD-vMCG method

5.2 Electronic Structure Calculations

An excited state dynamics calculation of TAM and TIM presents a greater challenge to the limitations of both the DD-vMCG method and the external quantum chemistry calculations in comparison to the study of FAM and FIM. In addition to the difficulties presented due to the presence of the nitrogen atom, the large number of degrees of freedom, and the unknown number of excited states required in the calculations, a major limitation to the study of these systems is the amount of time required for the external quantum chemistry calculations to run. As a demonstration, in Table 5.1 the times taken for a selection of optimisations of TAM, carried out using various levels of theory and basis set, are shown accompanied by the equivalent calculation times for the optimisations of FAM.

	HF		MP2		CCSD	
	FAM	TAM	FAM	TAM	FAM	TAM
3-21G∗	12.6	24.4	23.3		65.8	1783.2
6-31G∗	26.9		44.8	187.3	287.1	
6-311+G∗	43.4		73.7		966.8	53583.0
cc-pVDZ	31.9		65.3		418.1	9941.2
cc-pVTZ	330.8	1147.9	1483.2		18771.3	

Table 5.1: A comparison of the CPU times taken (in seconds) for equivalent optimisations of formamide and 2-pyridone showing selection from levels of theory and basis sets. Gaussian 09 [46] was used on a Xeon ES-2640 8-core workstation.

It can easily be seen that the use of the Dunning basis sets and the Pople triple-ζ basis set would be unsuitable as the Direct Dynamics calculation time would be dominated by the time taken for the electronic structure calculations.

Consequently, it was decided that, as with the FAM calculations, a large active space would be chosen for the CAS calculations while a compromise basis-set size, 6-31G∗, would be chosen.

5.2.1 Excited State Calculations

State-Averaged CAS Investigation

Following the philosophy behind the selection of requirements for the active space as described in Chapter 4.2.2, it was decided that as with the study of FIM and FAM, emphasis would be placed on the possible description of a proton transfer between the isomers TIM and TAM. Accordingly, in order to adequately describe this potential proton transfer the molecular orbitals involving the N-H σ- and σ^*-bonding as well as the oxygen lone pair in TAM are required, while in TIM the corresponding O-H σ- and σ^*-bonding character along with the nitrogen lone pair must be described. Due to the conjugated ring in TAM and TIM it is even more important for adequate description of the π-bonding network to be included in the active space.

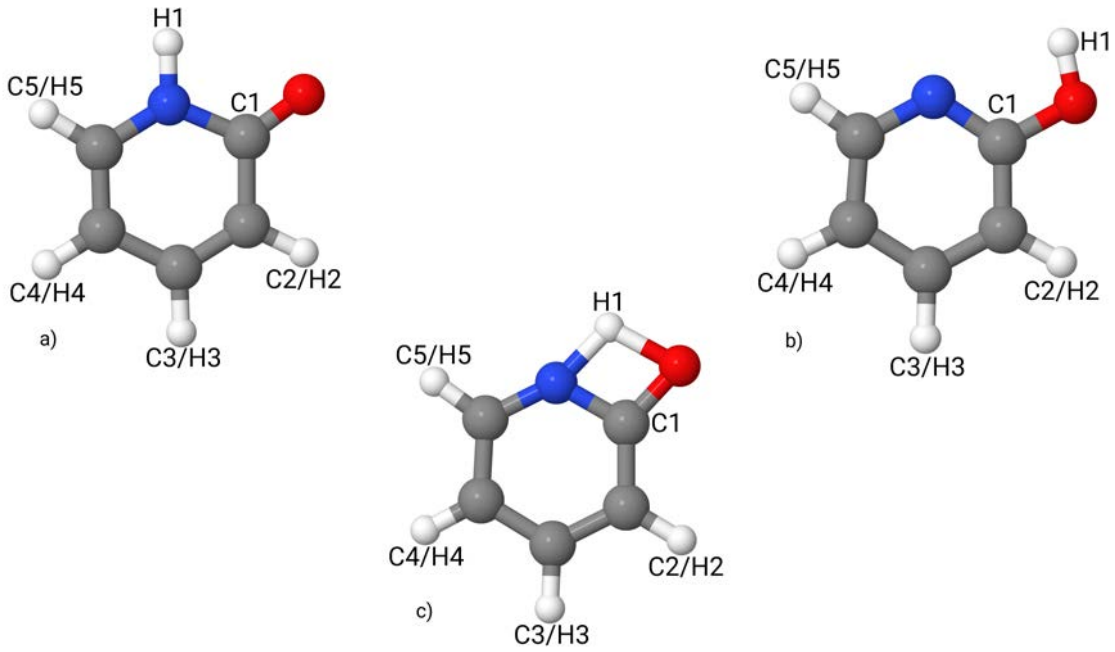


Fig. 5.3: The 6-31G∗ MP2 optimised geometries of a) 2-pyridone, b) 2-hydroxypyridine and c) The transition state. The labels for the carbon atoms and hydrogen atoms are for clarity in the interpretation of the data in Tables 5.4 and 5.6.

It should be noted, at this point, that if too large an active space is chosen, the quantum chemistry software would not be able to calculate higher lying excited sates. After an inspection of the molecular orbitals of both TAM and TIM calculated at the Hartree-Fock level it was decided that either a CAS(12,10), a CAS(10,9) or a CAS(10,8) would be required. However, when initial state averaged calculations using a 12 electron, 10 orbital size active space were carried out it was found that the calculations could not converge if the number of excited states included in the calculation was higher than 3. As a result, the state averaged CAS investigation was carried out using the 10 electron, 9 and 8 orbital active spaces, the results of which can be seen in Figure 5.4 for 2-pyridone and Figure 5.5 for 2-hydroxypyridine. It should also be noted that in these Figures for simplicity the NH(1) and OH(1) bonding labels have been abbreviated to NH and OH.

Fig. 5.4: The characterisation of the molecular orbitals and excited states for 2-pyridone using CAS(10,8) (left) and CAS(10,9) (right) sized active spaces. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. The key to the colour scheme is on the right hand side. The list of orbitals above the cells list those in the CAS space. The numbers in red are for states with significant oscillator strength (> 0.01).

In Figure 5.4 it can be seen that while calculations for state averaging over a total of 5 to 9 states were performed, only the SA5 and SA6 calculations utilising the CAS(10,9) were successful, while at the CAS(10,8) level of theory the SA5 to SA8 calculations were successful. This was seen as an indicator that the use of the CAS(10,8) calculations was likely to result in increased stability in the calculations. Upon inspection of the ASC for the CAS(10,9) calculations it is immediately clear that while the two successful calculations contained the prerequisite oxygen lone pair, π and π^* molecular orbitals, neither the N-H(1)-bonding nor -antibonding character is present in the active space. However, upon inspection of the ASC for the CAS(10,8) calculations it can be seen that only the SA5 calculation failed to include the prerequisite N-H(1) σ and σ^* orbitals, while all states also include the oxygen lone pair, π and π^* bonding character.

The XSC of both the CAS(10,8) and CAS(10,9) calculations show that aside from the SA5 calculation, all S_1 and S_3 states are characterised by an OLP to π^* transition, while all S_2 states are characterised by a π to π^* transition. In fact it can be seen that the XSC of the CAS(10,8) SA6 to SA8 calculations is consistent across all states. Despite the dominance of the π to π^* transitions, the S_5 is characterised as a π to NH^* transition. It can also be seen that the calculated excitation energies are less consistent than in the FAM and FIM calculations, though with the limited amount of data it is difficult to draw any conclusions from this. An additional observation to be made is that while the character of the excited states is consistent, the SA8 calculation found the S_3 state to be optically bright as opposed to the S_4 in the preceeding calculations.

From the analysis of this set of results it can be seen that for 2-pyridone the CAS(10,8) sized calculation utilising between 5 and 7 excited states would be the most suitable for the Direct Dynamics calculations.

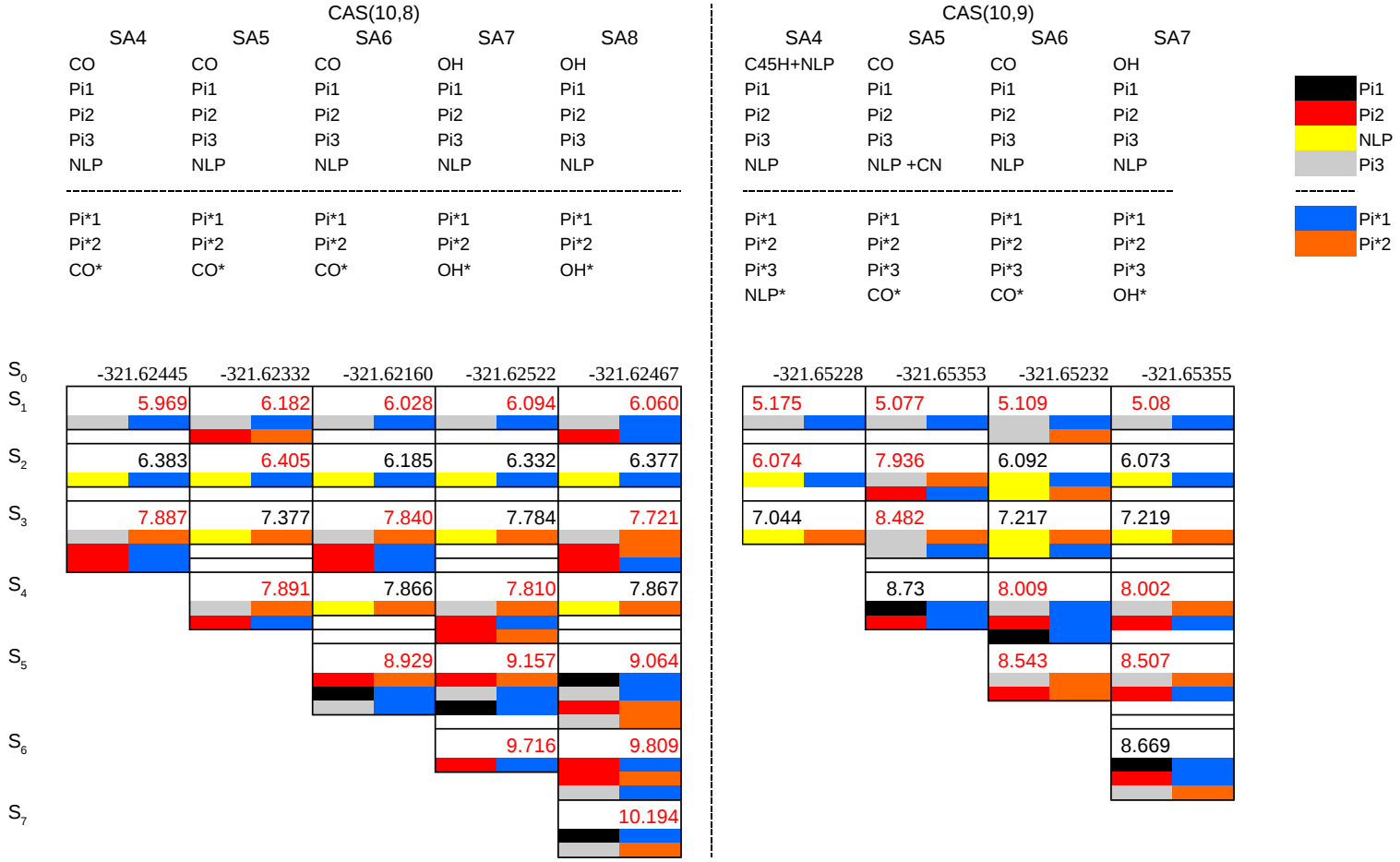


Fig. 5.5: The characterisation of the molecular orbitals and excited states for 2-hydroxypyridine using CAS(10,8) (left) and CAS(10,9) (right) sized active spaces. Each cell represents a state with excitation energy (in eV) relative to the ground-state (energy in Hartree). The coloured boxes relate to the character of the main configurations with an electron being promoted from the left to the right box. The key to the colour scheme is on the right hand side. The list of orbitals above the cells list those in the CAS space. The numbers in red are for states with significant oscillator strength (> 0.01).

In Figure 5.5 it can be seen that while state averaged calculations were carried out over a total of 4 to 9 states at the CAS(10,9) level of theory the SA8 and SA9 calculations were unsuccessful while at the CAS(10,8) level of theory only the SA9 calculation was unsuccessful. Upon inspection of the ASC for the CAS(10,9) calculations it can be seen that while the prerequisite nitrogen lone pair, π and π^* character is seen in the active space of the SA4 to SA7 calculations, only the SA7 includes the OH(1) σ- and σ^*- bonding orbitals. In the ASC of the CAS(10,8) calculations it can also be seen that the nitrogen lone pair, π and π^* character molecular orbitals are present in the SA4 to SA8 calculations, the OH(1) σ- and σ^* orbitals are present only in the SA7 and SA8 calculations.

As with the results of the calculations for 2-pyridone, a consistent trend of character in the XSC is seen for the S_1 state, which is characterised as a π to π^* transition, and with the exception of the CAS(10,9) SA5 calculation the S_2 is characterised as a NLP to π^* transition. This represents a reverse of the character of the first two states in comparison to the 2-pyridone results. Unlike the TIM results, the S_3 state does not exhibit a clear trend of character. Upon closer inspection of the S_3 and S_4 states in the CAS(10,8) calculations it can be seen that the character of the states has swapped in the SA5 and SA7 by comparison to the SA4, SA6 and SA8, which is likely due to the states being very close in energy. This same "swap" is also observed in the identification of the states with significant oscillator strengths. Although the character of the S_5 state in the CAS(10,8) calculations is not consistent, combinations of the same three principal transitions are seen, all of which represent π to π^* transitions, albeit from π molecular orbitals of differing character. It can also be seen that with the exception of the SA6 S_2 and the SA5 S_3 the calculated excitation energies are relatively consistent across all states in the CAS(10,8) calculations.

Consequently it can be seen that for 2-hydroxypyridine the CAS(10,8) sized

calculations are more stable with the SA7 and SA8 being the most suitable for the Direct Dynamics calculations.

	Energy (eV)	Oscillator Strength (au)	Main configurations
S_1	4.821	0.0006	0.64(OLP – Pi*1)
S_2	4.914	0.1082	0.61(Pi3 – Pi*1)
S_3	6.126	0.0000	0.64(OLP – Pi*2)
S_4	7.610	0.1924	0.42(Pi3 – Pi*2) + 0.38(OLP – NH*)
S_5	7.758	0.0000	0.64(Pi3 – NH*)
S_6	8.422	0.2273	0.52(Pi2 – Pi*1)

Table 5.2: 2-pyridone energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA7-CAS(10,8)/6-31G* calculation

As a result of this investigation, following from the requirement that the maximum number of excited states should be included in the calculation, it is clear that the CAS(10,8) including 7 excited states, the SA8 calculation, is the obvious choice. However, the initial point of the Direct Dynamics calculations require a calculation of the energy, gradient and Hessian, along with the non-adiabatic couplings between all of the states. This calculation is performed as a standalone calculation providing both an estimate of the time required for the quantum chemistry calculations as well as presenting a final test to the stability of the calculation before moving on to the Direct Dynamics. Although this step proved to be merely a formality for the formamide and formimidic acid calculations, this calculation for the TAM repeatedly failed to complete. Consequently, following the results of the state average characterisations, the next largest CAS(10,8) calculation, state averaged over a total of 7 states, was then tested and used in the Direct Dynamics study.

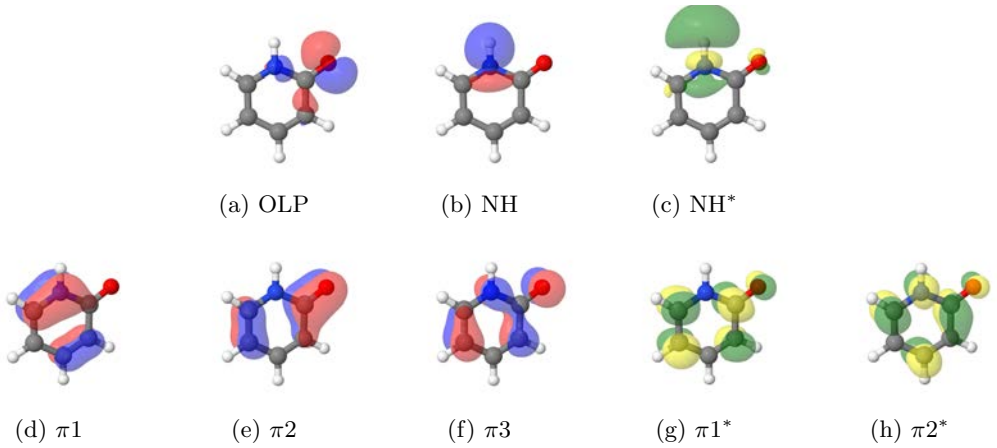
It should be noted that while the 2-hydroxypyridine CAS(10,8) SA8 calculation was successful, the time taken for this calculation was 21615.69 seconds whereas the CAS(10,8) SA7 calculation took 19656.85 seconds As this represents an approximate 10% reduction in the calculation time, given the aforementioned

requirements for efficiency in the calculations, the reduction of the number of excited states presented an improvement in the overall time-expense for the Direct Dynamics calculation.

	Energy (eV)	Oscillator Strength (au)	Main configurations
S_1	6.094	0.0514	0.45(Pi3 – Pi*1) - 0.28 (Pi3 – Pi*2) + 0.26(Pi2 – Pi*2)
S_2	6.331	0.0092	0.59(NLP – Pi*1) +0.28(NLP – Pi*2)
S_3	7.783	0.0006	0.59(NLP – Pi*2) -0.29(NLP – Pi*1)
S_4	7.810	0.0251	0.45(Pi3 – Pi*2) + 0.29(Pi2 – Pi*1) + 0.28(Pi2 – Pi*2)
S_5	9.157	0.6450	0.39(Pi2 – Pi*2) + 0.39(Pi3 – Pi*1) + 0.26(Pi1 – Pi*1)
S_6	9.715	0.7825	0.45(Pi1 – Pi*1)

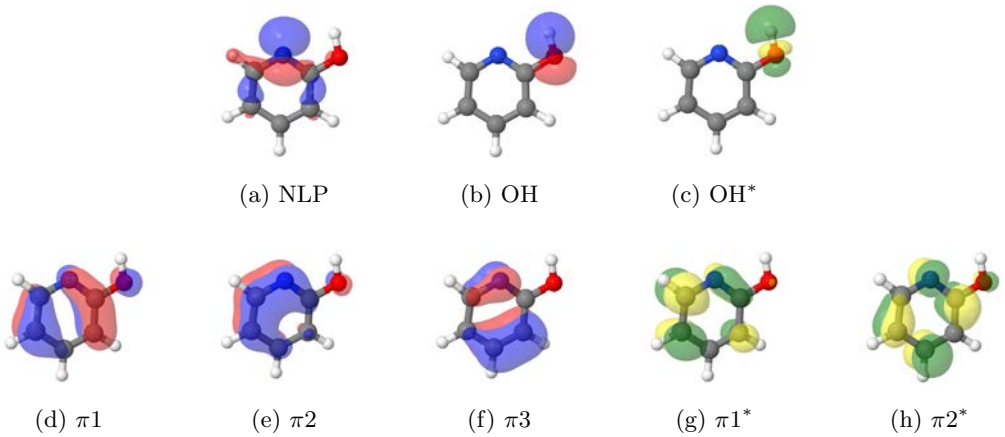
Table 5.3: 2-hydroxypyridine energies, oscillator strengths and coefficients of main configurations (values > 0.15) from a SA7-CAS(10,8)/6-31G* calculation

The final results of the excitation energies, oscillator strengths and characterisations from the SA7-CAS(10,8)/6-31G* calculations of TAM and TIM are summarised in Tables 5.2 and 5.3 respectively. In Table 5.2 it can be seen that the states principally characterised by a π to π^* transition are optically bright where the S_6 $\pi 2$ to π^*1 is the brightest, followed by the S_4 $\pi 3$ to π^*2 then the S_2 $\pi 3$ to π^*1. In Table 5.3 it can also be seen that the states principally characterised by a π to π^* transition are optically bright. In this case the S_5 $\pi 1$ to π^*1 and S_6 $\pi 2$ to $\pi^*2/\pi 3$ to π^*1 states exhibit relatively high oscillator strengths whilst the S_1 $\pi 3$ to π^*1 and S_4 $\pi 3$ to π^*2 are comparatively weaker in strength, though still optically bright. The molecular orbitals of TAM and TIM can be seen in Figures 5.6 and 5.7, respectively, from which the full characterisations of the π orbitals can be interpreted.



(a) OLP (b) NH (c) NH* (d) $\pi 1$ (e) $\pi 2$ (f) $\pi 3$ (g) $\pi 1^*$ (h) $\pi 2^*$

Fig. 5.6: CAS molecular orbitals of 2-pyridone from the CAS(10,8) SA7 calculation.



(a) NLP (b) OH (c) OH* (d) $\pi 1$ (e) $\pi 2$ (f) $\pi 3$ (g) $\pi 1^*$ (h) $\pi 2^*$

Fig. 5.7: CAS molecular orbitals of 2-hydroxypyridine from the CAS(10,8) SA7 calculation.

Vibrational Frequencies

As stated in Chapters 3.3.1 and 4.2.2 the coordinates, defined either in Cartesian coordinates or normal modes, are required so as to provide the ground state potential surface and width of the total wavepacket. It was decided that the normal modes of TIM and TAM would be used and hence the characterisation of the normal modes is necessary in order to carry out analysis of the Direct Dynamics results. The notation for characterisation of the modes follows the geometry labels outlined in Figure 5.3. As both the TAM and TIM have C_s symmetry the modes are defined only as A' and A'' indicating in-plane (ip) and out-of-plane (oop) mo-

tion. Generally, in the characterisation of the modes the atoms grouped together exhibit motion in-phase with each other, while the motions of atoms either side of the hyphen are generally out-of-phase (oo-phase) or opposed to each other. An example of this can be seen in Table 5.4, ν_{12} where the H(2) and H(3) are moving in-phase with each other, as are the H(4) and H(5), while the combined motion of the H(2) and H(3) is out-of-phase with the motion of the H(4) and H(5) combined motion.

In Tables 5.4 and 5.5 the normal modes of the 6-31G$*$ SA7 CAS(10,8) optimised structure of 2-pyridone are characterised, alongside the frequencies calculated at the 6-31G$*$ MP2 and 6-311+G$*$ CCSD levels of theory. Comparing the frequencies of the MP2 and CCSD calculations it can be seen that the majority of the lower frequency modes, ν_1 to ν_{13}, are predominantly lower at the CCSD level of theory, while the CCSD calculated frequencies in the mid-range frequency modes, ν_{14} to ν_{25}, are predominantly higher than those calculated at the MP2 level of theory. It can also be seen that while the high-frequency C-H stretching modes are at lower calculated frequencies in the CCSD calculations, the N-H stretching mode is higher frequency, implying that at the CCSD level of theory the NH bond is stronger than at the MP2 level of theory.

Comparing the calculated frequencies resulting from the CAS(10,8) SA7 calculation, to the MP2 and CCSD calculated frequencies it can be seen that apart from the ν_1 mode, all of the CAS modes are at a higher frequency.

Label	Frequency (cm^{-1})			Character
	MP2 6-31G*	CCSD 6-311+G*	CAS(10,8) 6-31G* SA7	
$\nu_1(A'')$	201.69	116.46	160.69	OC(4)-NC(2) oop rocking
$\nu_2(A'')$	387.21	352.92	421.10	C(2)C(5)-NC(3) oop twisting
$\nu_3(A')$	456.77	392.21	485.15	O ip wagging
$\nu_4(A'')$	467.04	457.72	542.70	H(1)H(2)H(3)H(5)-C(1)C(4) oop oo-phase wagging
$\nu_5(A')$	539.84	550.05	575.12	C(1)-C(4) ip ring breathing
$\nu_6(A')$	608.46	601.89	653.54	C(3)N-C(2)C(5) ip twisting
$\nu_7(A'')$	654.12	618.26	754.47	H(1)H(2)H(4) oop in-phase wagging
$\nu_8(A'')$	703.68	634.18	777.92	H(1) oop wagging
$\nu_9(A'')$	758.14	736.45	806.24	H(1)H(2)H(3) oop in-phase wagging
$\nu_{10}(A')$	800.94	807.79	880.85	C(1)-N C(1)-C(2) ip in-phase stretching
$\nu_{11}(A'')$	834.05	832.57	923.91	H(2)H(3) oop in-phase wagging
$\nu_{12}(A')$	873.39	867.00	978.50	H(2)H(3) oop oo-phase wagging
$\nu_{13}(A'')$	946.02	930.81	1027.32	C(3)-C(4) ip stretching
$\nu_{14}(A')$	984.69	1003.07	1077.71	C(2)C(4)N-C(5)C(3) ip twisting
$\nu_{15}(A')$	997.76	1018.40	1148.81	H(3)H(2) oop oo-phase wagging
$\nu_{16}(A'')$	1101.20	1119.10	1182.00	H(4)H(2)-H(1) ip oo-phase wagging
$\nu_{17}(A')$	1182.33	1181.62	1263.75	H(2)-H(3)-H(4) ip wagging
$\nu_{18}(A')$	1240.68	1242.32	1315.91	H(1)-H(5)-H(4) ip wagging
$\nu_{19}(A')$	1261.69	1292.87	1377.71	H(1)H(2)-C(1)H(4) ip in-phase wagging
$\nu_{20}(A')$	1402.76	1407.33	1519.07	H(1)-H(2)-H(3) ip wagging
$\nu_{21}(A')$	1442.43	1474.93	1562.03	H(3)-H(1)-H(2) ip wagging
$\nu_{22}(A')$	1485.58	1514.24	1642.36	H(5)-H(4)-H(2) ip wagging
$\nu_{23}(A')$	1594.98	1619.40	1677.77	H(5)-H(3) ip oo-phase wagging
$\nu_{24}(A')$	1682.41	1698.31	1765.59	H(5)-H(4) ip oo-phase wagging
$\nu_{25}(A')$	1747.64	1792.54	2011.64	H(1) ip wagging C(1)O stretching
$\nu_{26}(A')$	3210.96	3183.35	3367.16	C(3)-H(3) stretching
$\nu_{27}(A')$	3238.29	3214.01	3409.18	C(2)-H(2) stretching
$\nu_{28}(A')$	3254.06	3221.55	3417.53	C(4)-H(4) stretching
$\nu_{29}(A')$	3263.03	3235.57	3449.89	C(5)-H(5) stretching
$\nu_{30}(A')$	3590.05	3640.81	3706.24	NH(1) stretching

Table 5.4: The normal mode frequencies of lactam calculated at different levels of theory at the optimised C_s structure where the numbering of the carbon atoms and hydrogen atoms is as in Figure 5.3a).

Table 5.5: Normal modes of lactam calculated at the CAS(10,8)/6-31G* SA7 level of theory at the optimised C_s structure, where "Freq." is an abbreviation of frequency.

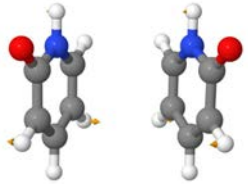
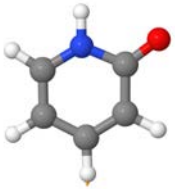
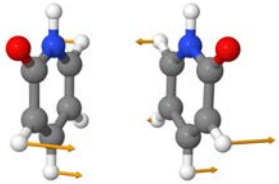
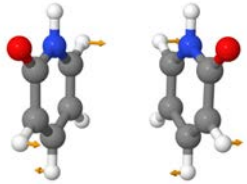
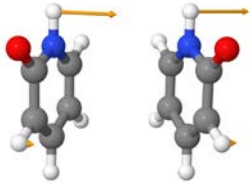
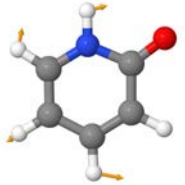
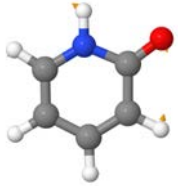
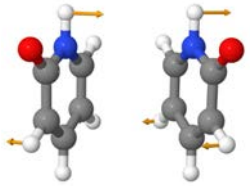
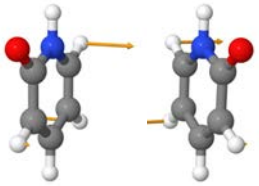
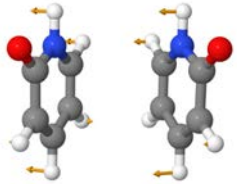
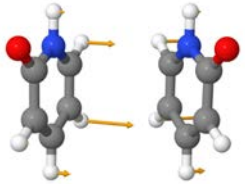
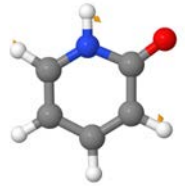
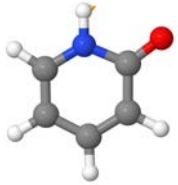
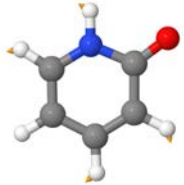
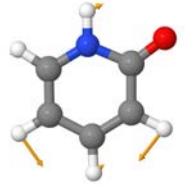
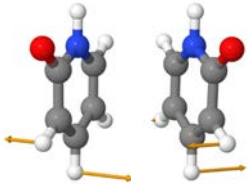
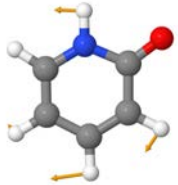
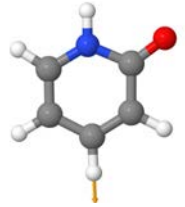
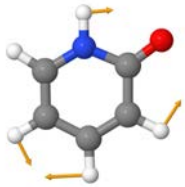
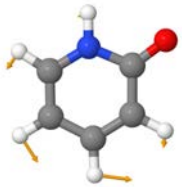
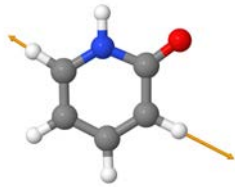
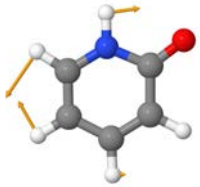
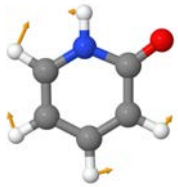
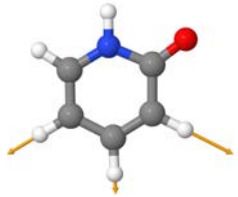
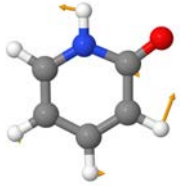
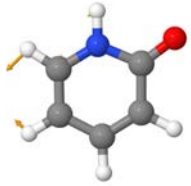
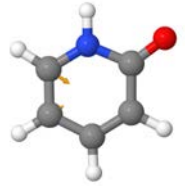
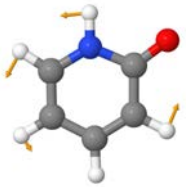
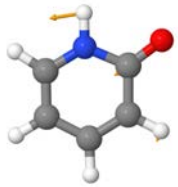
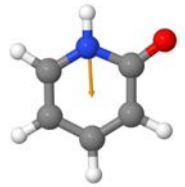
Label	Freq. (cm^{-1})	Label	Freq. (cm^{-1})	Label	Freq. (cm^{-1})
$\nu_1(A'')$ 	160.69	$\nu_6(A')$ 	653.54	$\nu_{11}(A'')$ 	923.91
$\nu_2(A'')$ 	421.10	$\nu_7(A'')$ 	754.47	$\nu_{12}(A')$ 	978.50
$\nu_3(A')$ 	485.15	$\nu_8(A'')$ 	777.92	$\nu_{13}(A'')$ 	1027.32
$\nu_4(A'')$ 	542.70	$\nu_9(A'')$ 	806.24	$\nu_{14}(A')$ 	1077.71
$\nu_5(A')$ 	575.12	$\nu_{10}(A')$ 	880.85	$\nu_{15}(A')$ 	1148.81

Table 5.5: Continued

Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})
$\nu_{16}(A'')$		1182.00	$\nu_{21}(A')$		1562.03	$\nu_{26}(A')$		3367.16
$\nu_{17}(A')$		1263.75	$\nu_{22}(A')$		1642.36	$\nu_{27}(A')$		3409.18
$\nu_{18}(A')$		1315.91	$\nu_{23}(A')$		1677.77	$\nu_{28}(A')$		3417.53
$\nu_{19}(A')$		1377.71	$\nu_{24}(A')$		1765.59	$\nu_{29}(A')$		3449.89
$\nu_{20}(A')$		1519.07	$\nu_{25}(A')$		2011.64	$\nu_{30}(A')$		3706.24

In Tables 5.6 and 5.7 the normal modes of the 6-31G∗ SA7 CAS(10,8) optimised structure of 2-hydroxypyridine are characterised, alongside the frequencies calculated at the 6-31G∗ MP2 and 6-311+G∗ CCSD levels of theory. Comparing the frequencies of the MP2 and CCSD calculations it can be seen that in the mid- to high-frequency range, ν_{15} to ν_{30}, the calculated frequencies at the CCSD are predominantly lower than those of the MP2 calculations, with the exception of the OH-stretching mode, ν_{30}. As with the results of the TAM calculations, the high-frequency TIM C-H stretching modes are at lower calculated frequencies in the CCSD calculations than at the MP2 level of theory, while the O-H stretching mode is higher frequency, implying that at the CCSD level of theory the OH bond is calculated to be stronger than at the MP2 level of theory.

It can be seen that the 6-31G∗ SA7 CAS(10,8) calculated frequencies are lower than the frequencies calculated at the other levels of theory only in the ν_4 vibrational mode.

The inclusion of the pictographic representations of the normal modes are particularly useful in a system of this size when later attempting to identify and categorise the motion seen in the result of the Direct Dynamics calculations.

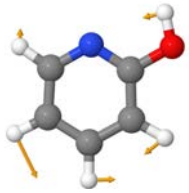
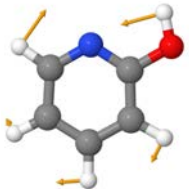
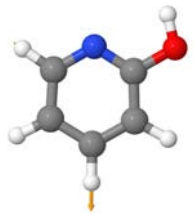
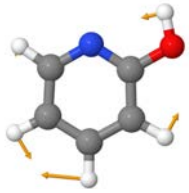
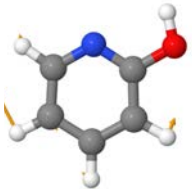
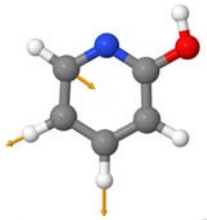
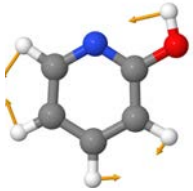
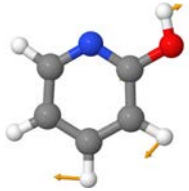
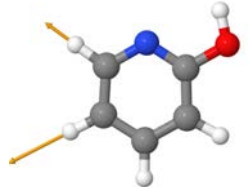
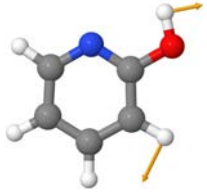
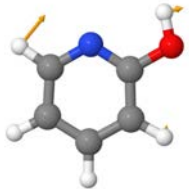
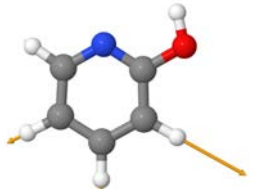
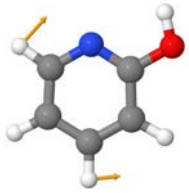
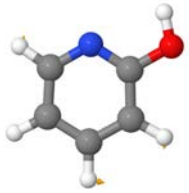
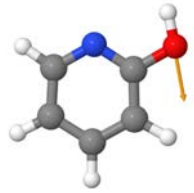
Label	Frequency (cm^{-1})			Character
	MP2 6-31G*	CCSD 6-311+G*	CAS(10,8) 6-31G* SA7	
$\nu_1(A'')$	215.99	212.17	248.63	O(H(1))-C(4) oop bending
$\nu_2(A'')$	407.42	400.82	439.12	OH(1) ip wagging
$\nu_3(A')$	414.08	422.72	469.27	H(1)(-NC(3)) oop (in-phase) wagging
$\nu_4(A'')$	488.70	441.61	483.99	H(1)(-C(2)C(5)) oop (in-phase) wagging
$\nu_5(A')$	544.17	516.21	596.80	H(1)(C(1)C(4))-H(2)H(3)H(5) oop oo-phase wagging
$\nu_6(A')$	566.78	570.76	606.56	C(1)-C(4) ring breathing
$\nu_7(A'')$	639.41	641.22	680.32	C(2)-C(5) ring breathing
$\nu_8(A'')$	672.51	686.41	731.12	C(1)-C(2) stretching
$\nu_9(A'')$	772.17	765.88	813.91	H(2)H(4)N-C(1)C(5)C(3) oop rocking
$\nu_{10}(A')$	859.85	844.66	865.19	H(2)H(3)H(4)H(5) oop in-phase wagging
$\nu_{11}(A'')$	865.87	865.48	980.16	C(3)-C(5) ring breathing
$\nu_{12}(A')$	919.58	917.94	980.81	H(2)H(3)-H(4)H(5) oop oo-phase wagging
$\nu_{13}(A'')$	929.73	944.83	1039.87	H(5)-H(4)oop oo-phase wagging
$\nu_{14}(A')$	1016.43	1017.02	1087.95	NC(4)C(2) ring breathing
$\nu_{15}(A')$	1083.08	1069.61	1138.80	H(3)-H(2)H(4) oop oo-phase wagging
$\nu_{16}(A'')$	1129.84	1121.54	1188.69	H(4)H(3)(H(1))-H(2) ip oo-phase wagging
$\nu_{17}(A')$	1199.55	1162.53	1241.55	H(4)H(2)H(1)-H(3) ip oo-phase wagging
$\nu_{18}(A')$	1231.47	1230.04	1279.11	H(1)H(5)H(3)-H(2)H(4) ip oo-phase wagging
$\nu_{19}(A')$	1347.60	1282.93	1383.34	H(2)-H(1) ip in-phase wagging
$\nu_{20}(A')$	1379.52	1346.04	1438.46	H(3)-H(5) ip oo-phase wagging
$\nu_{21}(A')$	1434.74	1387.06	1477.77	H(5)H(4)H(3)H(2)-H(1) ip oo-phase wagging
$\nu_{22}(A')$	1518.12	1505.52	1615.77	H(4) ip wagging
$\nu_{23}(A')$	1541.15	1533.84	1617.87	H(3)H(2)H(1) ip in-phase wagging
$\nu_{24}(A')$	1660.01	1660.16	1663.15	H(5)H(1) ip in-phase wagging
$\nu_{25}(A')$	1688.38	1686.16	1841.11	H(2)H(5)-H(3) ip oo-phase wagging
$\nu_{26}(A')$	3225.79	3184.02	3366.80	H(3) stretching
$\nu_{27}(A')$	3236.92	3191.27	3372.75	H(5) stretching
$\nu_{28}(A')$	3261.64	3215.44	3395.36	H(4) stretching
$\nu_{29}(A')$	3268.68	3223.66	3406.89	H(2) stretching
$\nu_{30}(A')$	3706.33	3806.50	3746.61	H(1) stretching

Table 5.6: The Normal mode frequencies of lactim calculated at different levels of theory at the optimised C_s structure where the numbering of the carbons and hydrogens is as in Figure 5.3b).

Table 5.7: Normal modes of lactim calculated at the CAS(10,8)/6-31G* SA7 level of theory at the optimised C_s structure, where "Freq." is an abbreviation of frequency.

Label	Freq. (cm^{-1})	Label	Freq. (cm^{-1})	Label	Freq. (cm^{-1})
$\nu_1(A'')$	248.63	$\nu_6(A')$	606.56	$\nu_{11}(A')$	980.16
$\nu_2(A')$	439.12	$\nu_7(A')$	680.32	$\nu_{12}(A'')$	980.81
$\nu_3(A'')$	469.27	$\nu_8(A')$	731.12	$\nu_{13}(A'')$	1039.87
$\nu_4(A'')$	483.99	$\nu_9(A'')$	813.91	$\nu_{14}(A')$	1087.95
$\nu_5(A'')$	596.80	$\nu_{10}(A'')$	865.19	$\nu_{15}(A'')$	1138.80

Table 5.7: Continued

Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})	Label		Freq. (cm^{-1})
$\nu_{16}(A'')$		1188.69	$\nu_{21}(A')$		1477.77	$\nu_{26}(A')$		3366.80
$\nu_{17}(A')$		1241.55	$\nu_{22}(A')$		1615.77	$\nu_{27}(A')$		3372.75
$\nu_{18}(A')$		1279.11	$\nu_{23}(A')$		1617.87	$\nu_{28}(A')$		3395.36
$\nu_{19}(A')$		1383.34	$\nu_{24}(A')$		1663.15	$\nu_{29}(A')$		3406.89
$\nu_{20}(A')$		1438.46	$\nu_{25}(A')$		1841.11	$\nu_{30}(A')$		3746.61

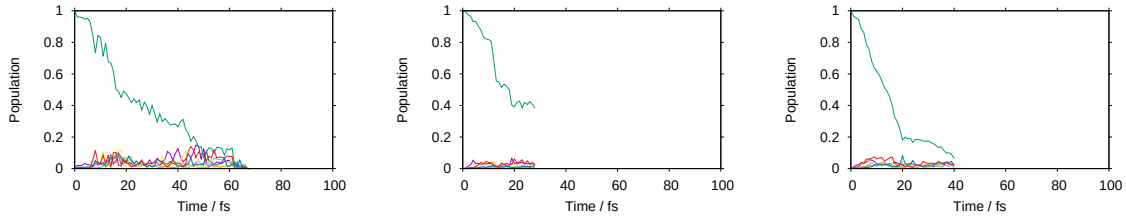
5.3 Direct Dynamics

Following the same protocol for the building of the database and the placement of the CAPs as described in Chapter 4.3.1 the Direct Dynamics Calculations for 2-pyridone and 2-hydroxypyridine were carried out and results obtained. It should be noted that the same predefined parameters for the propagation time, 150 fs, and dbmin, 0.25, were used in the Direct Dynamics calculations on 2-pyridone and 2-hydroxypyridine, as were used with the FAM and FIM propagations.

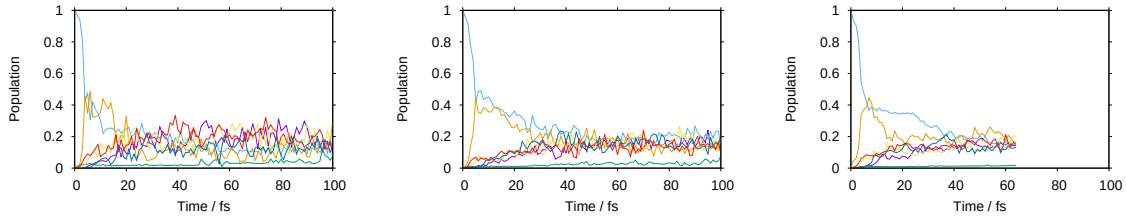
5.3.1 State Population Analysis

In Figures 5.8 and 5.10 the diabatic state population results of the 2-pyridone and 2-hydroxypyridine, respectively, for each of the 7, 21 and 42 GWP propagations on the S_1, S_2, S_3 and S_6 states are shown. Figures 5.9 and 5.11 detail the results of the first 80 fs of the 21 GWP propagations, with a comparison to the total density. These state population calculations, along with the detail of the total density, give vital information as to the behaviour of the system where the inclusion of the total density allows the population transfer and the dissociative behaviours to be resolved.

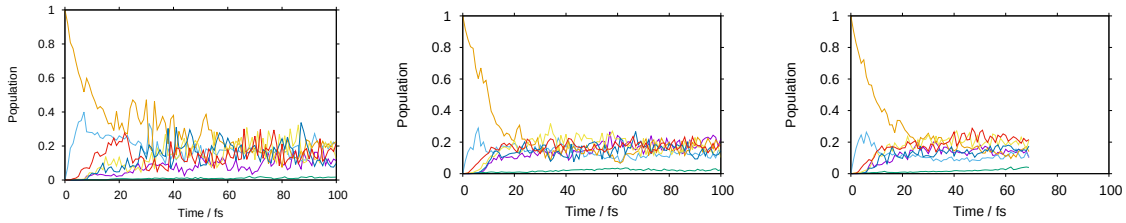
In Figure 5.8 it can be seen that the increase from 7 to 21 GWPs and from 21 to 42 GWPs result in a smoothing of the representation of the decay and population transfer and this convergence can clearly be seen in all but the S_6 state. The increase in the number of GWPs, or basis functions, means that the basis set for the dynamics has greater flexibility and can cover more of phase space. Following from this observation it can be stated that while the plots for the state populations on the S_6 state do not become more smooth, it can clearly be seen that while in the 7 GWP calculation the state populations appear to be dissipating, in the 21 and more so in the 42 GWP propagations it can be seen that the density appears to be staying in the upper states and not decaying.



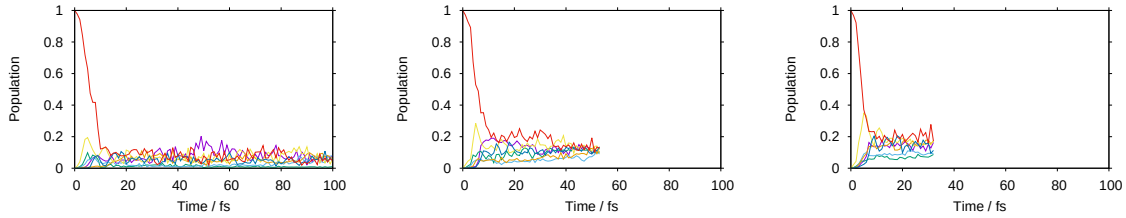
(a) S_1 7GWP (b) S_1 21GWP (c) S_1 42GWP



(d) S_2 7GWP (e) S_2 21GWP (f) S_2 42GWP



(g) S_3 7GWP (h) S_3 21GWP (i) S_3 42GWP



(j) S_6 7GWP (k) S_6 21GWP (l) S_6 42GWP

Fig. 5.8: Diabatic state populations from DD-vMCG simulations of 2-pyridone starting with a vertical excitation to various states and using various numbers of GWPs.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red.

In the S_1 calculations it can be seen that the timescale for the decay of population in the S_1 state occurs initially over the first 20 fs where it can be seen that population appears to be being transferred into all of the other states. This decrease in the population appears to be increasingly long lived upon increase of the number of basis functions included in the calculation.

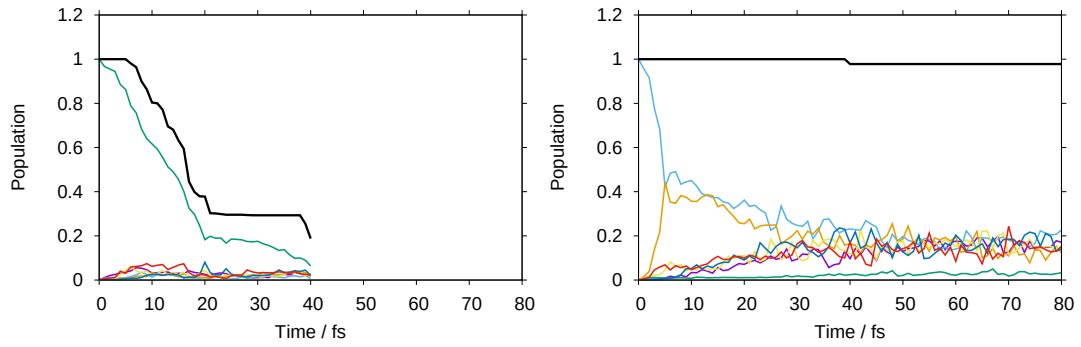
In the S_2 calculations it can be seen that there is a very fast population transfer from the S_2 (light blue) to the S_3 (orange) state, occurring within the first 10 fs. With increasing number of GWPs the degree to which this population transfer occurs is reduced, though still significant. The increased number of basis functions representing the system allows for the resolution of the population transfer to the other states to be increased and hence can be seen more clearly.

In the S_3 7 GWP calculation it appears as if there are two points of significant population transfer, the first at about 10 fs to the S_2 state, and the second at 20 fs to the S_6 (red) state. However, these features are diminished with increasing number of basis functions to the point where there appears to be no particular significance to the population transfer into the S_5 state. One feature that has remained constant across this set of calculations is the the slow increase in the population of the S_1 (green) state showing that a slow decay into this state is occurring.

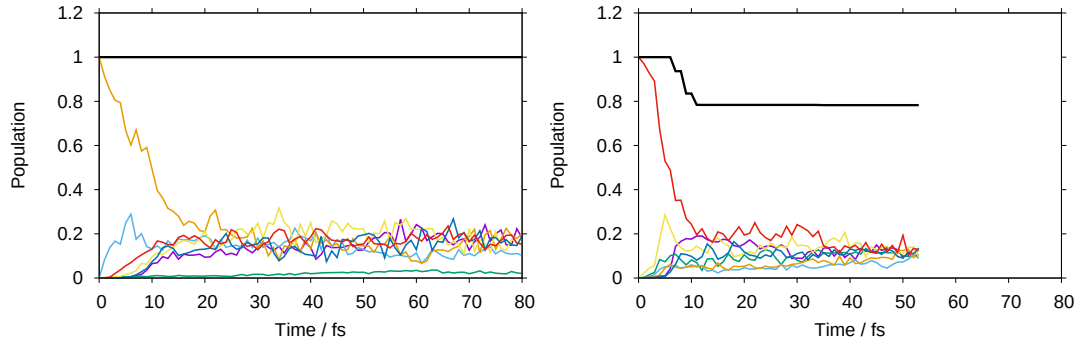
In the S_6 calculations the rapid initial decay is also observed in the first 10 fs, the degree of which is greater in the 7 GWP calculation than the 21 GWP calculation. The degree of this fast decay is apparently maintained between the 21 and 42 GWP calculations. However, despite this lower percentage of decay in the 21 and 42 GWP calculations, it can be seen that the initial population transfer to the S_4 state increases with increasing number of basis functions while the overall population in each of the states appears to be averaging to between 10 and 20%.

In Figure 5.9 it can be seen that the total density has decreased by about 70% during the first 20 fs of the S_1 propagation whereas the total density in the in the S_6 calculation decreases only by 20% in the first 10 fs of the calculation and then is maintained. While the total density in the S_2 state exhibits a slight decrease at 40 fs the most significant calculation with respect to the total density is the S_3 propagation where none of the density is lost over the entire calculation. This minimal,

or non-existent, change in the total density supports the previous statements that the density is being efficiently distributed among the excited states. It should also be noted that minimal population is being transferred into the S_0 state in the S_1 calculation, while the population is relaxing to the ground state to a much more significant degree in the S_2, S_3 and S_6 calculations. As stated previously, minimal population is being transferred into the S_1 state in the other propagations.



(a) S_1 21GWP (b) S_2 21GWP



(c) S_3 21GWP (d) S_6 21GWP

Fig. 5.9: Diabatic state populations from DD-vMCG simulations of 2-pyridone starting with a vertical excitation to various states. Final Results.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; Total density (norm^2): Thick black.

In the S_2, S_3 and S_6 calculations is is difficult to resolve the behaviour of the system any further as the population is being transferred to and between states rapidly.

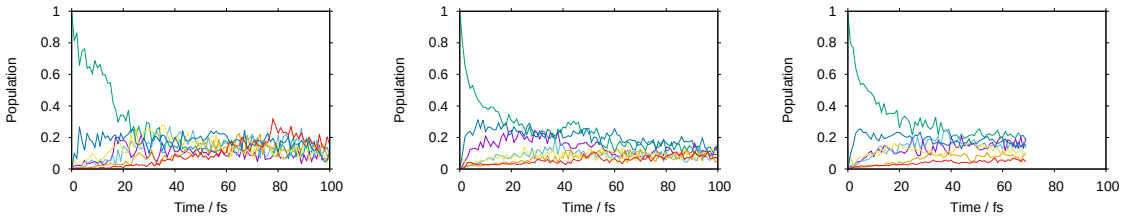
In Figure 5.10 it can be seen that while the increase in the number of GWPs reduced the overall noise in the state population transfer the improvement is not as marked as in the previous state population analyses partly due to the limited 42 GWP results available. However, upon inspection it can be seen that the curves going from the 7 GWPs to the 21 GWPs exhibit less chaotic population transfer, the most significant being the large population transfer exhibited at approximately 55 fs in the S_6 7GWP is not seen in the 21 GWP calculation.

In the S_1 state the 7 GWP calculation exhibits and initial 60% population decay over the first 20 fs, principally characterised as population transfer to the S_5 (dark blue) state. Between 40 and 60 fs the S_5 state becomes the highest populated state after which, at around 80 fs, the S_6 (red state) becomes the most highly populated. However, in the 21 GWP calculation the S_0 (purple) state also exhibits a large population increase and between 30 and 40 fs it appears as if the population on the S_0, S_1 and S_5 is equivalent, while the large population transfer to the S_6 at later propagation times is not exhibited. In the 42 GWP calculation the S_0, S_4 (yellow) and S_2 (light blue) exhibit concurrent behaviour for the first 20 fs, after which, at 40 fs, it can be seen that the S_2 is competing with the S_1 state populations. In this set of calculations it is not conclusive as to the behaviour of the system with increasing numbers of GWPs suggesting that the calculation may not yet have converged.

In the S_2 state it can be seen that the role of the S_3 (orange) state is significant in the the 7 GWP calculation while it does not particularly feature in the 21 GWP calculation. A similar diminution in the prevalence of the S_4 state is seen going from the 7 GWP to the 21 GWP calculation, though to a lesser extent.

In the S_3 state the decay in the population of the propagating state reduces to 20% population in around 18 fs, the fastest decay of seen in this system, and while the significant noise in the calculations reduces from the 7 GWP calculation

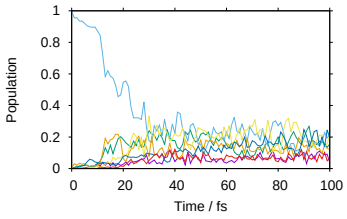
to the 21 GWP calculation, the prevalence of the population transfer to particular states is unchanged.



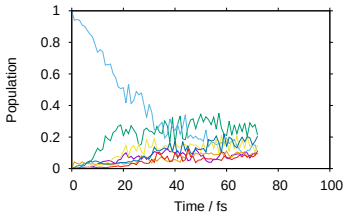
(a) S_1 7GWP

(b) S_1 21GWP

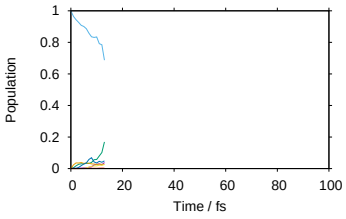
(c) S_1 42GWP



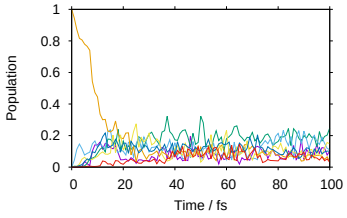
(d) S_2 7GWP



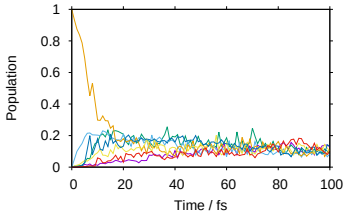
(e) S_2 21GWP



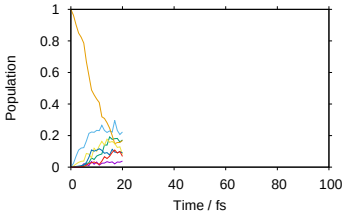
(f) S_2 42GWP



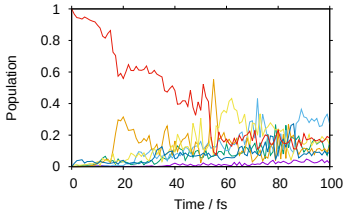
(g) S_3 7GWP



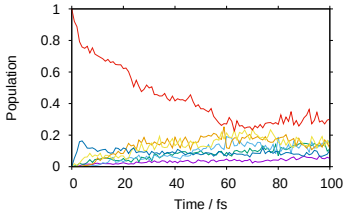
(h) S_3 21GWP



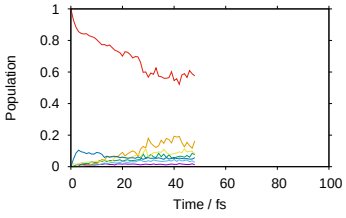
(i) S_3 42GWP



(j) S_6 7GWP



(k) S_6 21GWP



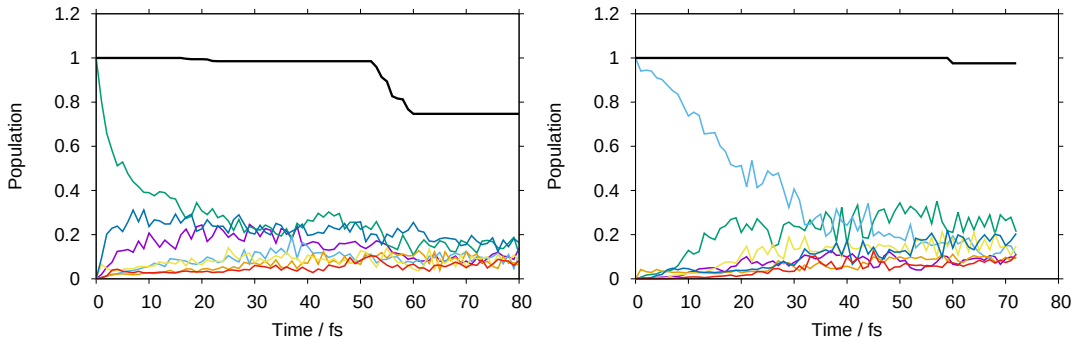
(l) S_6 42GWP

Fig. 5.10: Diabatic state populations from DD-vMCG simulations of 2-hydroxypyridine starting with a vertical excitation to various states and using various numbers of GWPs. Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red.

The S_6 set of calculations exhibits both the greatest improvement upon increase of the number of GWPs and the slowest decay in population of the propagating state, 20% population remaining past 60 fs in the 21 GWP calculation. In the

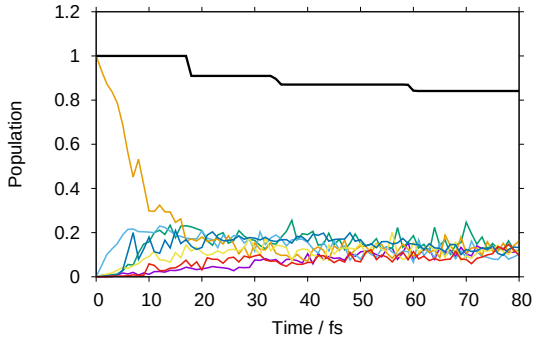
7 GWP calculation the most significant population transfer, to the S_3 state, at 20 fs is no longer seen in the 21 and 42 GWP results. In the 21 and 42 GWP result it can be seen that a small amount of population transfer to the S_5 state occurs within the first 5 fs, with no appreciable decay for the remainder of the propagation.

It should also be noted that every state in this set of calculations exhibits a small but appreciable amount of population transfer to the S_0 state.

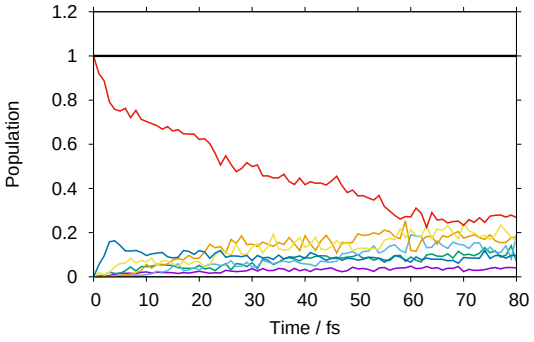


(a) S_1 21GWP

(b) S_2 21GWP



(c) S_3 21GWP



(d) S_6 21GWP

Fig. 5.11: Diabatic state populations from DD-vMCG simulations of 2-hydroxypyridine starting with a vertical excitation to various states. Final Results.
Key: S_0: purple; S_1: green; S_2: light blue; S_3: orange; S_4: yellow; S_5: dark blue; S_6: red; Total density (norm^2): Thick black.

In Figure 5.11 the first 80 fs of the 21 GWP calculations are shown, with the total density plotted for comparison. It can be seen that unlike in the previous

sets of results none of the states exhibit a large decrease in the density, indicating that relatively little to, in the case of S_6, no dissociation is occurring in these calculations. Consequently it is reasonable to state that 2-hydroxypyridine is relatively stable upon photoexcitation. The S_1 calculation exhibits the highest amount of population transfer to the S_0 ground state. In the S_2 and S_3 calculations the S_1 state presents significant behaviour while in the S_6 calculation relatively little transfer to the S_1 state occurs. It can also be seen that the S_5 state plays a significant role in the dynamics of the S_1 and S_2 calculations while in the S_6 calculation the S_5 state is significant only in the first 5 fs. It appears as if, in the S_1 and S_3, the populations on each of the states are converging, or equilibrating, further suggesting the photostability of the molecule.

As a result of this analysis it can be seen that aside from the TAM S_1 state 2-pyridone and its isomer 2-hydroxypyridine are relatively photostable, exhibiting a combination of short and long timescale behaviours, resulting in the distribution of population across all of the states of the system. Following the observations made in this section the further elucidation of the dissociation processes available to the dynamics of these molecules is vital to the understanding of the cases and modes which present molecular dynamical photoactivity.

5.3.2 Gross Gaussian Populations

Following a visual inspection of the geometries defined by the trajectory of the centre of the GWPs, the fraction of the density, defined by the gross Gaussian populations, along each of the product channels were then categorised to allow analysis of the product distribution. The results of this analysis of the trajectories of the GWPs for 2-pyridone and 2-hydroxypyridine can be seen in Figures 5.12 and 5.13 respectively. It was found, upon analysis of the S_7 state, that although the N-H dissociated in the dynamics of 2-pyridone, the motion did not make it as far as the CAP and hence no data were available for the analysis.

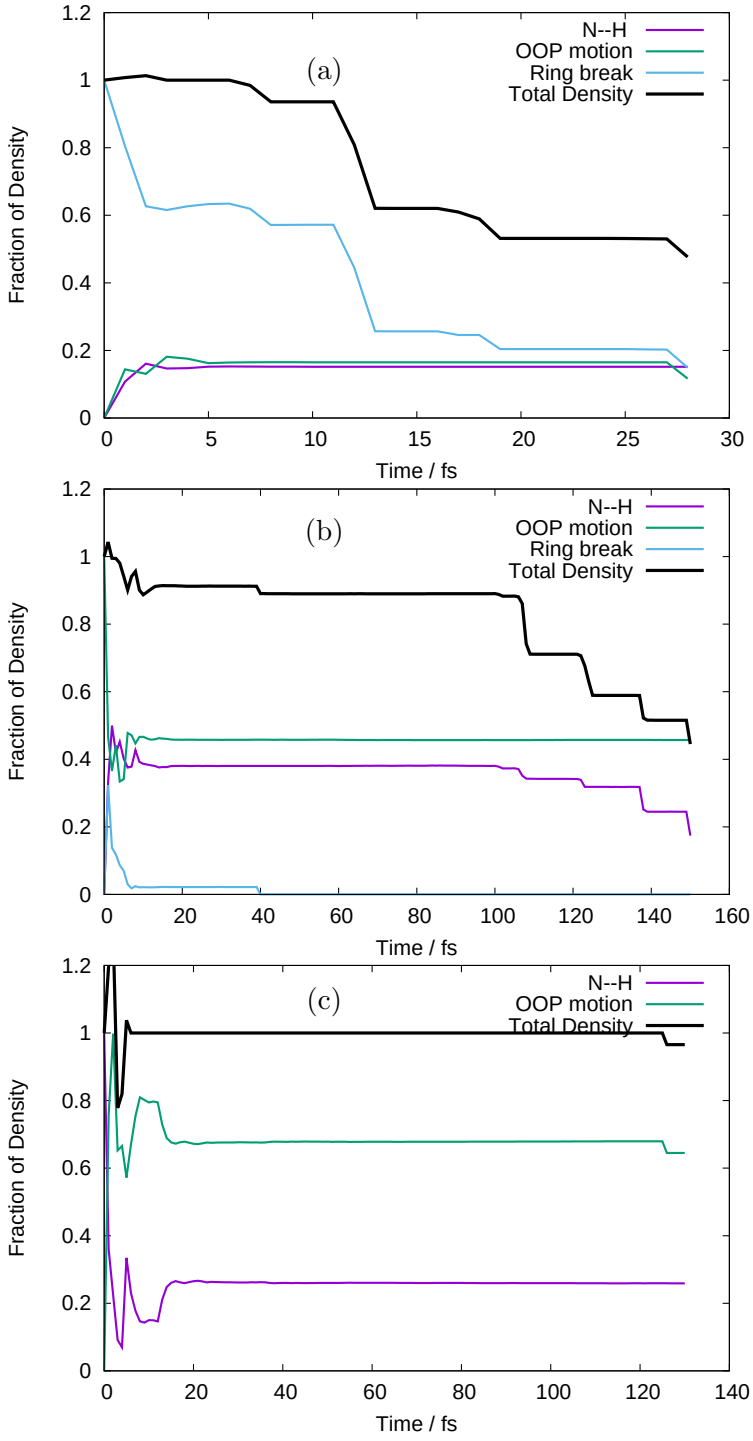


Fig. 5.12: The fraction of density going into different product channels from DD-vMCG simulations of 2-pyridone with SA7-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3. Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies OOP signifies out-of-plane ring vibration.

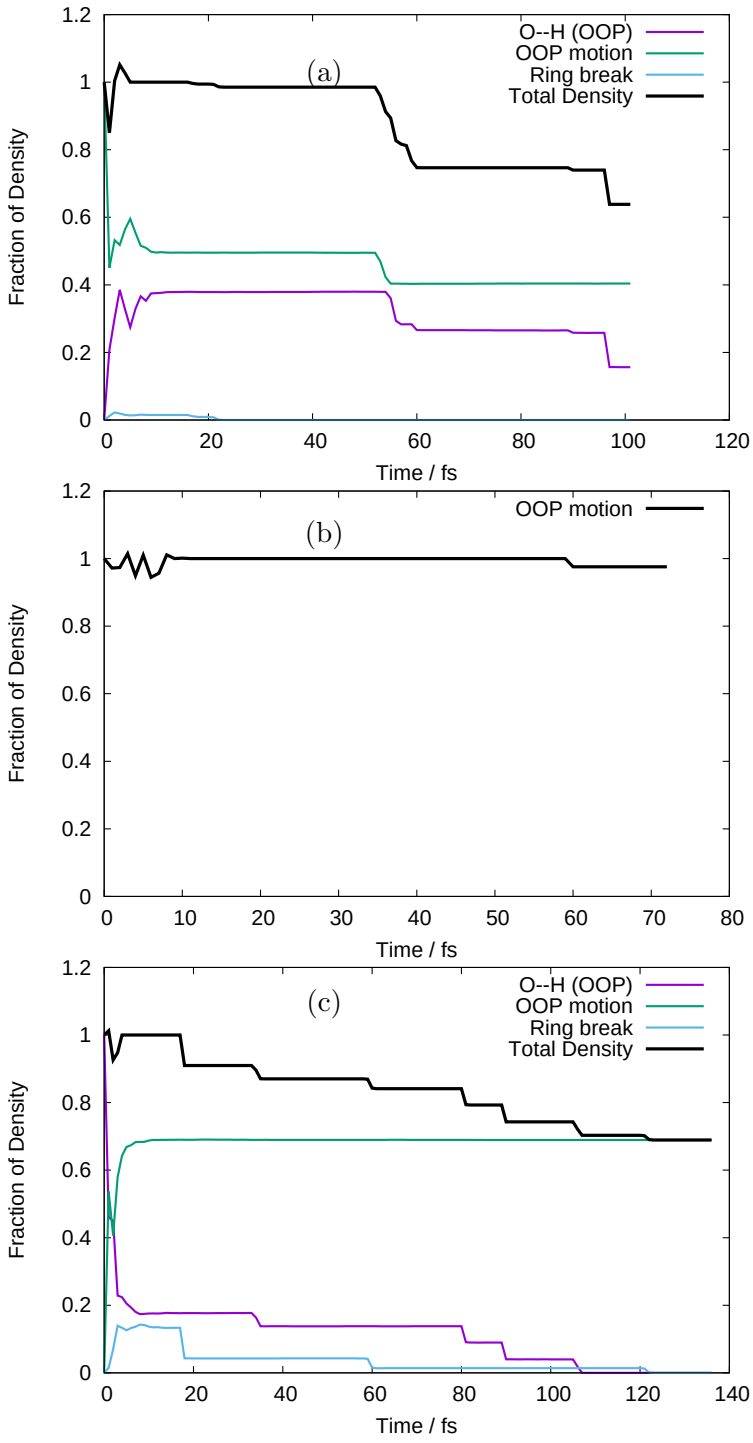


Fig. 5.13: The fraction of density going into different product channels from DD-vMCG simulations of 2-hydroxypyridine with SA7-CAS(10,8)/6-31G* following the potential surfaces starting in different states, where (a) S_1 (b) S_2 (c) S_3 Each line, or series, represents the different products defined either by the bond that breaks or by the products formed. As in the characterisation of the vibrational frequencies OOP signifies out-of-plane ring vibration.

In the S_7 state of 2-hydroxypyridine as none of the trajectories made it to the CAP, no data were available for the analysis of this state.

In Figure 5.12 the representation of N-H dissociation and out-of-plane ring vibration can be seen in all states, while the ring breaking does not occur in the S_3 state. It should be noted that the out-of-plane ring vibration did not result in dissociation, hence the separate classification of "ring breaking." In the S_1 state it can be seen that the ring breaking motion dominates for the duration of the propagation while the fraction of the density representing the N-H stretch and the out-of-plane vibration remains constant from around 2.5 fs until the end of the propagation. In the S_2 state the out-of-plane vibration dominates, with almost 50% of the density going into this motion. From the shape of the curve representing the N-H breaking motion it can be seen that this motion took 100 fs before it reached the CAP. The observed timescale for this motion in the S_2 may explain why the N-H dissociation is not observed in the S_1 state. In the S_3 state the out-of-plane vibration continued for around 125 fs before any observed loss of density, while the N-H stretching behaviour continued, without loss of density for the duration of the propagation. Consequently it may be that the NH dissociation takes over 100 fs to occur, but only if this motion is coupled to a ring-breaking mode, and hence the TAM is relatively stable to photoexcitation.

In Figure 5.13, representing the product distribution of TIM, the out-of-plane ring vibration is observed and represents the dominant feature in all three states. In the S_2 state, this out-of-plane ring vibration is representative of 100% of the density for around 60 fs, at which point it drops by 5%. In the S_1 and S_3 states in addition to the out-of-plane ring vibration, an OH out-of plane wagging motion is observed, as well as a ring breaking motion. Although this ring breaking motion is relatively long lived, approximately 120 fs, in the S_3 state, it represents a small fraction of the total density. In the S_2 state the ring breaking motion is represented

by an even smaller fraction of the total density and appears only for the first 20 fs. The OH out-of-plane motion represents between 30 and 40% of the total density in the S_1 state, while it represents less than 20% of the total density in the S_3 state. Consequently it can be said that while dissociation of TIM occurs in the excited states, it represents a relatively small fraction of the behaviour of the system upon photoexcitation.

As a result of this investigation it can be seen that 2-hydroxypyridine is relatively limited in the nuclear motion exhibited upon photoexcitation and while 2-pyridone is also relatively stable, more significant dissociative behaviour is seen around the 100 fs timescale.

5.3.3 Potential Energy Surfaces

Following from the behaviours seen in the Gross Gaussian populations it was decided that the surface cuts were selected along the modes which were most clearly represented by the density of the system. As the ring breaking and out-of-plane motion is the result of relatively large number of modes only the significant NH stretching, ν_{30}, mode was selected for 2-pyridone, while the three modes representing the out-of-plane and in-plane OH motion, ν_3 and ν_4 for out-of-plane, ν_{30} for the in-plane, were selected along which cuts of the potential energy surface were taken.

In Figure 5.14 the adiabatic and diabatic surfaces for the 2-pyridone calculation are shown. These surfaces, while there are no obvious discontinuities, display the expected behaviour in the 0 to +10 region. However, in the 0 to -10 region the surfaces have an unexpected form, the interpretation of which is too difficult at this time. It is possible to state that there appears to be at least one other significant state not included in the calculation (particularly at -2) and as the negative direction corresponds to the N-H bond lengthening, at least one dissociative state, though these comments are conjecture.

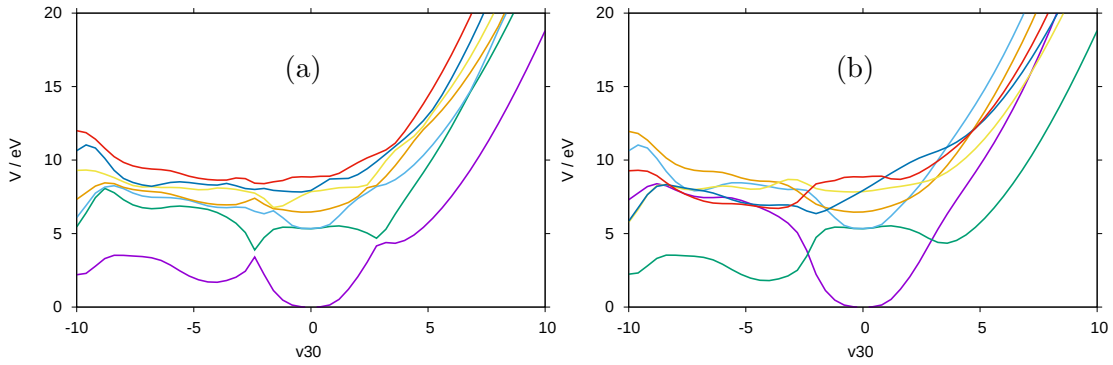


Fig. 5.14: Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-pyridone from DD-vMCG simulations . (a) ν_{30} (N-H stretch) adiabatic (b) ν_{30} diabatic.

In Figure 5.15 the adiabatic (left) and diabatic (right) surfaces representing the motion of the OH have been shown. In Figure 5.15 a) and b) the OH out-of-plane, ν_3, surfaces are shown. It is clear, in the diabatic surfaces, that the dark blue, S_5, state is strongly dissociative. It can also be seen that the energy difference between the S_1 and S_2, and between the S_3 and S_4 at the Franck-Condon point is very small in this mode.

In Figure 5.15 c) and d) the OH out-of-plane, ν_4, surfaces are shown. Although the S_5 in this state exhibits similar behaviour to that of the ν_3 calculation, this state is not as clearly dissociative, as can be seen in the upturn in energy of the state at around -7.5. However, as evidenced in the higher energy states at this (-7.5) point there is some further unexpected behaviour which may indicate the presence of another, previously uncalculated higher order state.

In Figure 5.15 e) and f) the adiabatic and diabatic cuts of the potential energy surface along the OH stretching, ν_{30}, mode are shown. With the exception of the dark blue, S_5, state, this mode exhibits behaviour similar to those of a harmonic oscillator. This nicely bound state clearly demonstrates the convergence of the DD-vMCG method in a multi-state problem.

In Figure 5.16 a 2-dimensional cut through the adiabatic potential energy surfaces of the S_0 and S_1 state along the OH out-of-plane, ν_3, and OH stretching, ν_{30},

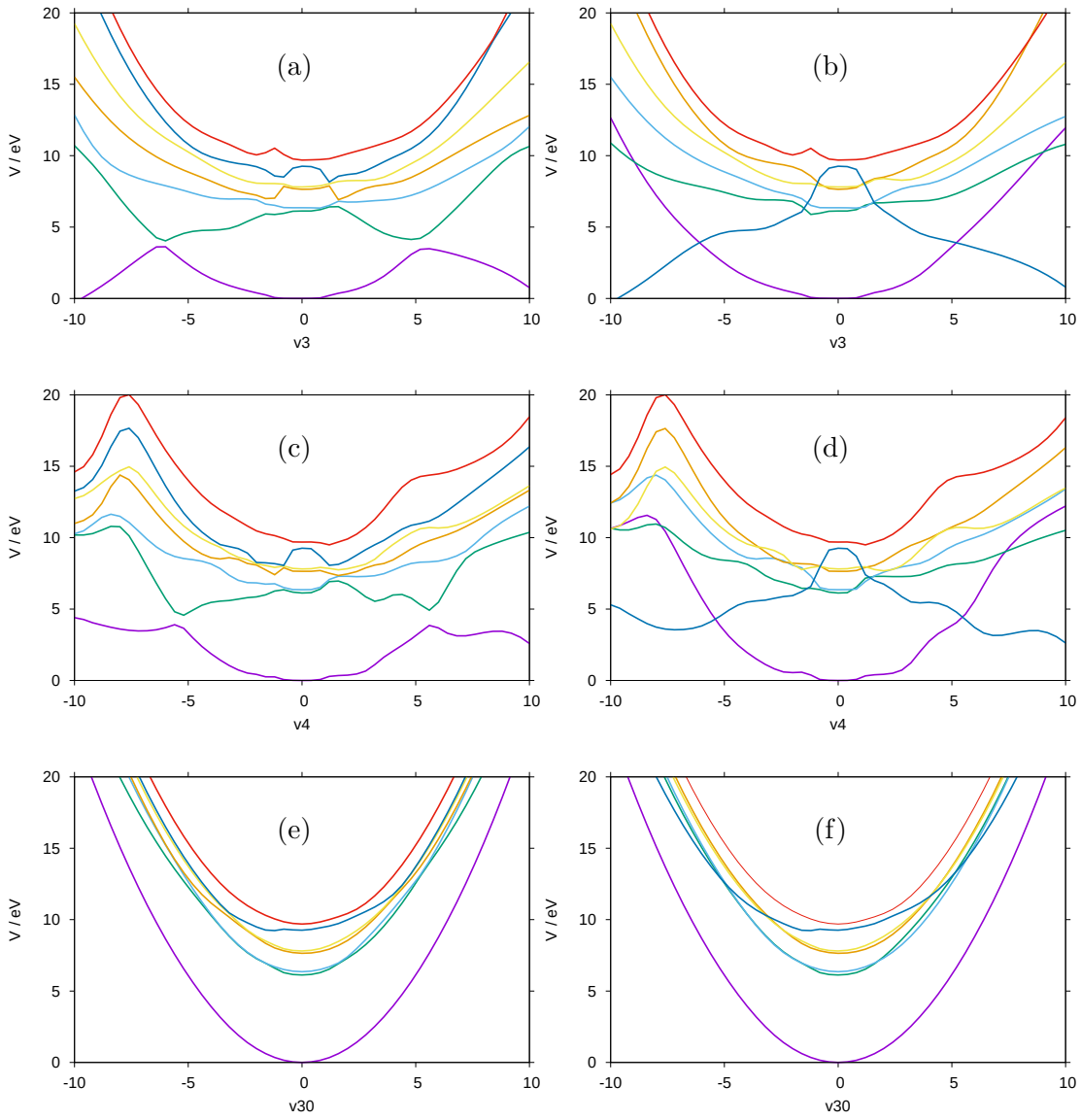


Fig. 5.15: Cuts through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-hydroxypyridine from DD-vMCG simulations . (a) ν_3 (O-H OOP1) adiabatic (b) ν_3 diabatic (c) ν_4 (O-H OOP2) adiabatic (d) ν_4 diabatic (a) ν_{30} (O-H stretch) adiabatic (b) ν_{30} diabatic.

modes can be seen. By inspection of the contours of this hypersurface a dissociation channel can be seen in the range of ν_3 0 to +10 and ν_{30} 0 to +5. This shows that the dissociation of the OH is driven not only by the in-plane, but also the out-of-plane motion.

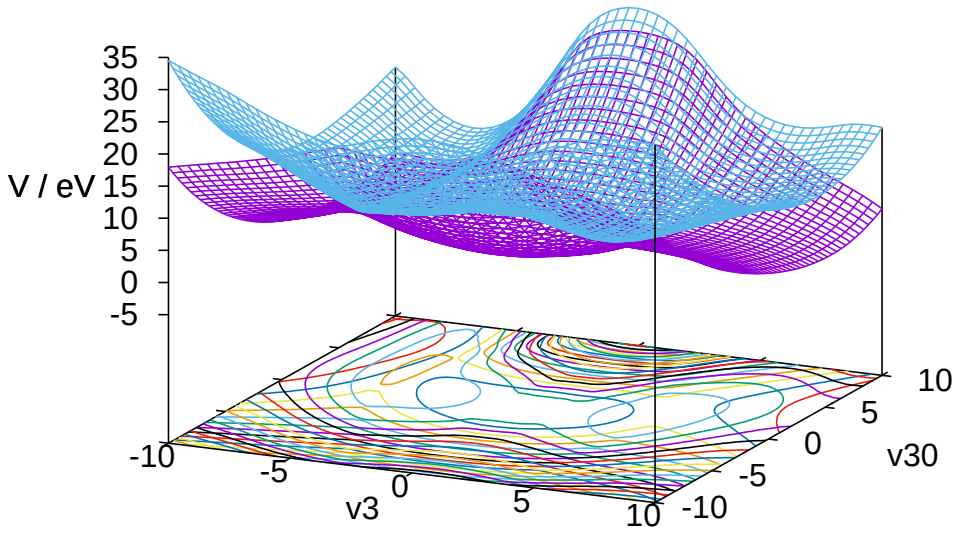


Fig. 5.16: Cut through the SA8-CAS(10,8)/6-31G* potential energy surfaces of 2-hydroxypyridine from DD-vMCG simulations. The adiabatic S_0 and S_1 in the coordinates ν_3 (O-H OOP1) and ν_{30} (O-H stretch).

5.4 Summary and Conclusions

In this Chapter, quantum chemistry and quantum dynamics calculations were performed in order to study the excited-state dynamics of 2-pyridone and 2-hydroxypyridine. As with the previous Chapter, the considerations for the representation of these two systems focussed on testing the limitations of the DD-vMCG method whilst maintaining the balance between computational and time expense with accuracy and stability. Due to the larger system size, and hence larger number of degrees of freedom, this represented a significant challenge to both the quantum chemistry calculations as well as the DD-vMCG method.

A brief study as to the timescales for the optimisations of the two sytems was carried out culminating in the conclusion (as with the formamide study) that the basis set size would be sacrificed in order to maintain feasible calculation times with a larger CAS size. Following the philosophy behind the selection requirements for the active orbitals of formamide (and formimidic acid), with the caveat that there exists a limit to the number of calculable excited states, an investigation

was carried out analysing the choice of CAS size and composition, as well as the effect of the inclusion of varying numbers of excited states. It was found that while a CAS(12,10) or CAS(10,9) captured the desired detail of the bonding character of the both systems, due to instabilities with the inclusion of increasing excited states a compromise would be made, in that a CAS(10,8) size would be used. Due to the stability and saving in computational time, a 6-31G* basis set with a CAS(10,8) over 7 averaged excited states was used for both 2-pyridone and 2-hydroxypyridine. It was found that π-π^* transitions were the dominant configurations for 2-hydroxypyridine and (to a lesser extent) 2-pyridone, with the S_6 state, in both cases, as the brightest state.

As with the formamide study, the DD-vMCG calculations were performed using the normal modes of the systems. An analysis of the character and numerical values of the vibrational frequencies of both systems was then carried out in order to identify significant modes (such as those leading to fragmentation or potential proton transfer) with a comparison to higher and lower accuracy calculations. Due to the planar equilibrium geometries of both molecules, unlike formamide, no imaginary modes are observed. It was found that the calculated frequencies of both molecules were higher than those calculated at a higher level of theory. This due the size of the CAS being relatively small in comparison to the size of the molecule and relative to the more complete description of formamide, hence providing a more limited description of the system.

Following the Direct Dynamics protocol outlined in Chapter 4.3.1, final Dynamics calculations were propagated on the S_1, S_2, S_3 and S_6 states, using 7, 21 and 42 GWPs over 150 fs. Three types of analysis were carried out on the results of the final sets of Direct Dynamics calculations.

The initial diabatic state population analysis of shows, in all but the 2-pyridone S_6 state, that the increase in GWPs results in a smoothing of the curves. It

should be noted that while the smoothing effect with increased GWPs is not seen in the 2-pyridone S_6 calculation, the 21 and 42 GWP calculations show density staying in the upper states, with little dissipation, whereas the 7 GWP result shows significant dissipation. Due to the significant increase in the calculation time for the quantum chemistry calculations, a number of the 21 and 42 GWP calculations, however, did not propagate for the full 150 fs. As a result, more detailed analysis (state populations, gross Gaussian populations and potential energy surfaces) is performed on the results of the 21 GWP calculations.

A detailed analysis of the state populations show that the only significant loss of total density is seen in the S_1 propagation of 2-pyridone. The results of the 2-pyridone calculations show that while relatively little population transfer occurs between states in the S_1 propagations, a significant amount of population transfer occurs in the other propagations, particularly between the S_2 and S_3 states in the S_2 propagation. In the S_2 and S_3 propagations that there is almost no population transfer to the S_1 state. It can also be seen that in the S_1 propagation there is little population transferred to to S_0 state, in the S_2 and S_3 propagations a steady increase of S_0 population is seen while significant population transfer to the S_0 states occurs at around 7 fs in the S_6 propagation. The results of the 2-hydroxypyridine calculations show that there is a large amount of population transfer to the S_0 state in the S_1 calculation and to a significantly lesser degree in the S_2, S_3 and S_6 propagations. Additionally, the S_1 state only plays a significant role in the S_2 and S_3 calculations while the S_5 state features significantly in only the S_1 and S_2 propagations. The results of this state population analysis indicate that 2-pyridone is photostable in the S_2 and higher states while 2-hydroxypyridine is photostable in all states.

The analysis of the gross Gaussian populations for 2-pyridone showed that in the S_1 propagation, the primary product channel (also accounting for the loss of

total density) was ring breaking mechanisms while in the S_2 and S_3 the primary product channel was out-of-plane motion with N-H bond breaking to a lesser extent. In the S_2 propagation the N-H bond breaking was more significant than in the S_3. It also appears to account for loss of total density in the S_2 state, while the out-of-plane motion in the S_3 accounts for the loss of total density. The gross Gaussian populations curves of 2-hydroxypyridine show that in the S_2 state, only out-of-plane motion is seen, while in the S_1 and S_3 states it is the dominant behaviour. In the S_1 state a significant amount of O-H out-of-plane is motion is seen whereas in the S_3 state this O-H out-of-plane motion is seen, but to a lesser degree. It should also be noted that in the S_2 and S_3 states, the out-of-plane ring vibration does not result in a loss of the total density which implies that the O-H out-of-plane motion is a much longer range motion. The results of this state population analysis further reinforce the observation that while 2-pyridone is relatively photostable, 2-hydroxypyridine is significantly more photostable.

As a result of the gross Gaussian population analysis the potential energy surfaces along the ν_{30} (N-H stretching) mode for 2-pyridone and the ν_3, ν_4 and ν_{30} (O-H out-of-plane (x 2) and O-H in-plane motion respectively) of 2-hydroxypyridine were analysed. The adiabatic and diabatic surfaces of 2-pyridone in the positive (0 to +10) direction are relatively smooth, whereas in the negative direction (0 to -10) there are a number discontinuities rendering further interpretation difficult. However the negative direction corresponds to the N-H bond lengthening and while a number of significant discontinuities exist rendering further interpretation difficult, it appears as if there may be a dissociative state. The adiabatic and diabatic surfaces of 2-hydroxypyridine, particularly the ν_{30} mode, are much smoother than the surface for 2-pyridone. The surfaces representing the ν_{30} mode are nicely bound, with some minor features, demonstrating good convergence of the DD-vMCG method. The surfaces representing the ν_3 and ν_4 modes appear to

show dissociative behaviour (the former more clearly than the latter) though the discontinuities in these surfaces again imply at least one other uncalculated state plays a significant role. The 2-dimensional surface of the S_0 and S_1 states along the ν_{30} and ν_3 modes show that a dissociation channel appears when these two motions are coupled. The potential energy surfaces of these two systems show that in the immediate region of Q_0, due to a number of close lying states, the quantum chemistry calculations require a higher degree of accuracy or a larger number of states to be able to resolve the features of the surface.

As a result of this study it can be seen that the excited state dynamics of 2-pyridone and 2-hydroxypyridine is not straightforward to resolve due to the number of states required and the close proximity of multiple states. These occurrences appear to be in specific regions of the global potential, in contrast to the large regions of difficulty in the formamide system. Nevertheless, as there is a limited number of product channels available to the system, resolution of the surfaces is simplified. It can also be stated that while 2-pyridone is relatively unstable in the S_1 state, the higher lying states of 2-pyridone and all states of 2-hydroxypyridine are highly photostable. It can also be stated that, as a result of this study, the N-H bond in 2-pyridone is more likely to experience long range (bond-breaking) motion than the O-H in 2-hydroxypyridine. It is, however, important to take into account the limiting factors of the size of the CAS, the significant amount of time required for the quatum chemistry calculations at every point in the calculation, and, more fundamentally, the limitations of the quantum chemistry programs with accurate CAS routines available for systems of this size when a large number of excited states are also required.

Chapter 6

The Electronic Wavepacket

6.1 Introduction

As outlined in previous chapters, the DD-vMCG method circumvents the requirement for pre-calculated potential energy surfaces, utilising coupled Gaussian wavepackets in order to explore configuration space even in systems with strong vibronic coupling. This is useful for the interpretation, and perhaps even prediction, of femto- to pico-second laser induced processes. However, a problem in this method arises when attempting to carry out quantum dynamics calculations modelling attosecond laser spectroscopic experiments. At the attosecond timescale it is crucial for there to be a description of the electron dynamics as it is the motions of the electrons that can be seen to be driving reactive processes. [41, 42] The movement of electrons across a molecule, driven by the nuclear motion, is called *charge transfer* whereas this flow of electrons across a molecule, not driven by the nuclear motion, is called *charge migration*. [198, 199]

A feature of these attosecond studies is that in order for the laser pulse to, for example, ionise a molecule, the band-width of the laser pulse must be wide - of the order of a few eV. This represents a major challenge to the computational simulation of these experiments as the inclusion of a large number of excited states, with couplings between these states, is necessary for the calculation to be able to describe the system to a sufficient degree of accuracy.

In the Ehrenfest approach the non-adiabatic couplings along a nuclear trajectory are calculated, resulting in a single mean-field potential energy surface. Ehrenfest trajectories are then propagated on a single potential surface which is constructed from a weighted average of the manifold of the electronic surfaces, the weighting defined by the electronic population of the states. Trajectory calculations have been carried out using this method in the study of charge migration in benzene, [200] glycine [201] and toluene. [202] However, the coupling of the trajectories in this method is ignored, resulting in a loss of coherence in the nuclear wavepacket. A method for combating this issue of decoherence, the Multi-Configurational Ehrenfest (MCE) of Shalashilin, is to include this lost coupling by the use of coupled coherent states [203, 204], which can be shown to be related to GWPs.

An alternative Ehrenfest method will be presented here, following and continuing from References [205] and [44]. Calculations testing this new Ehrenfest approach in comparison to the DD-vMCG will be presented, as an extension to the work of Reference [44]. The selected test system, allene (Figure 6.1), is a useful benchmark for this new method as the twisting motion of the molecule acts as a switch, allowing charge migration to occur, flowing from one end of the molecule to the other. It is also a useful example as pre-computed potential energy surfaces are available in the MCTDH software package, allowing grid-based benchmarking calculations to be carried out.

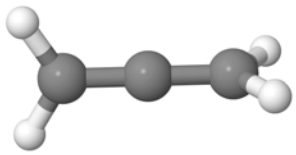


Fig. 6.1: The allene molecule

It should be noted that the text and figures of Sections 6.2, 6.2.2 and 6.3.1 - 6.3.2 are taken in part directly from Reference [44]. With respect to the work

taken from this paper (Reference [44]), I performed all of the calculations and was involved in the theory development, in particular the construction of the hierarchy of methods required to benchmark the new Ehrenfest approach.

6.2 Theoretical Framework of the Ehrenfest Approach

An extensive discussion of the Ehrenfest method has been published in the literature [206–211] However, in order to fully appreciate the theoretical implications of the Ehrenfest approach its underlying approximations must first be explored. [205] If the TDSE is written, explicitly expressing the nuclear, R, and electronic, r, coordinates

$$i\hbar\frac{\partial}{\partial t}\Phi(r,R,t) = \hat{H}\Phi(r,R,t) \tag{6.1}$$

the wavefunction ansatz for the Ehrenfest method can be written as

$$\Phi(r,R,t) = \Psi(r,t)\cdot\chi(R,t) \tag{6.2}$$

where the total wavefunction has been separated into the nuclear, $\chi(R,t)$, and electronic, $\Psi(r,t)$. In this form it can be seen that the total wavefunction is defined where the populated electronic states share the same nuclear wavepacket. Consequently the Ehrenfest wavefunction in this ansatz does not allow decoherence, though when applied to attosecond timescales this does not prove to be an issue as electronic decoherence occurs on a much longer timescale.

In order to proceed from this point it is useful to reformulate Equation 6.1 by inserting the ansatz Equation 6.2 where a phase factor in the total, nuclear and electronic wavefunctions has been introduced. Multiplying on the left by the complex conjugates of the nuclear and electronic wavefunctions ($\chi^*(R,t)$ and

$\Psi^*(r,t)$) and integrating over the nuclear and electronic coordinates, respectively, yields

$$i\hbar\frac{\partial\Psi(r,t)}{\partial t} = -\sum_i \frac{\hbar}{2m_e}\nabla_i^2\,\Psi(r,t) + \langle\chi(R,t)|\hat{V}_{n-e}(r,R)|\Psi(R,t)\rangle_R \cdot \Psi(r,t) \quad (6.3)$$

$$i\hbar\frac{\partial\chi(R,t)}{\partial t} = -\sum_I \frac{\hbar}{2M_I}\nabla_I^2\,\chi(R,t) + \langle\Psi(r,t)|\hat{H}(r;R)|\Psi(r,t)\rangle_r \cdot \chi(R,t) \quad (6.4)$$

Here, the lower case indices, i and m_e, refer to the electrons and electronic mass respectively, and the upper case indices, I and M_I, refer respectively to the nuclei and nuclear mass of I. The $\hat{V}_{n-e}(r,R)$ term in Equation 6.3 allows the inclusion of inter-particle interactions, while the $\hat{H}(r;R)$ in Equation 6.4 is the electronic Hamiltonian where the nuclei are at fixed positions R. By inspection of the second term on the right-hand-side of both equations it can be seen that for Equation 6.3, the interaction between the electrons at r and nuclei at R is weighted by the probability the nuclei are at point R while for Equation 6.4 the corresponding interaction is weighted by the probability the electrons are at point r. These coupled equations are the effective potentials experienced by one type of particle due to the other and hence are described in a mean-field manner. Consequently it can be stated that the appropriate expectation values of the electronic and nuclear wavefunctions determine the time-dependent effective potentials in which the nuclei and electrons can move.

The Ehrenfest approach is to take the classical limit of the coupled Equations 6.3 and 6.4. Writing the nuclear wavefunction in polar coordinates in terms of a real and positive amplitude A and phase S has the exact form

$$\chi(R,t) = A(R,t)\cdot\exp\left(\frac{i}{\hbar}S(R,t)\right) \quad (6.5)$$

where the inclusion of Planck's constant in the exponent ensures the exponential is unitless. Inserting this into Equation 6.4, allowing the real parts on both sides to be equal

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S)^2 + \langle \Psi(r,t)|\hat{H}_e(r;R)|\Psi(r,t)\rangle_r = \hbar^2 \sum_I \frac{1}{2M_I}\frac{\nabla_I^2 A}{A} \tag{6.6}$$

Taking the classical limit to be where $\hbar \to 0$ consequently reduces the right hand side of this to zero, which explicitly is

$$\frac{\partial S}{\partial t} + \sum_I \frac{1}{2M_I}(\nabla_I S)^2 + \langle \Psi(r,t)|\hat{H}_e(r;R)|\Psi(r,t)\rangle_r = 0 \tag{6.7}$$

where the Schrödinger equation, in this form, can be seen to be identical to a variant of the Hamilton-Jacobi equation. Consequently Equation 6.7 is equivalent to the Newtonian equations of motion where $P_I = \nabla_I S$ is the classically defined momentum of the nucleus I

$$\frac{\partial P_I}{\partial t} = -\nabla_I \langle \Psi(r,t)|\hat{H}_e(r;R)|\Psi(r,t)\rangle_r \tag{6.8}$$

As a result, the classical limit of Equation 6.3 can be taken by replacing $\chi(R,t)$ with a delta function in the classical trajectory of $R(t)$ as

$$i\hbar\frac{\partial \Psi(r,t;R)}{\partial t} = \left(-\sum_i \frac{\hbar^2}{2m_e}\nabla_i^2 + \hat{V}_{n-e}(r,R(t))\right) \tag{6.9}$$

which allows the electronic wavefunction to be written in the form

$$\Psi(r,t;R) = \hat{H}_e(r;R(t)) \cdot \Psi(r,t;R) \tag{6.10}$$

This results in an electronic wavefunction Ψ now parametrically dependent on the nuclear geometry at time t through $\hat{V}_{n-e}(r,R(t))$ and consequently $\hat{H}_e(r;R(t))$. As the nuclear motion is treated classically, the spacial delocalisation of the nuclei is lost and hence the nuclear motion is treated as a classical trajectory.

The Ehrenfest method is hence defined by Equations 6.8 and 6.10. It should be noted that nuclear motion is only successfully described if the energy and topology of the potential energy surfaces of the electronic states included in the calculation are similar. It should also be noted that when the electronic states are weakly coupled the potential of a highly populated electronic state will dominate the nuclear motion. This may result in regions of space that are left unexplored by the trajectories as they are accessible only on the sparsely populated electronic states. An advantageous feature of the Ehrenfest method is that if a complete set of basis functions is used, its applications and results are not dependent on the choice of basis functions. Consequently, in principle, the Ehrenfest method can be applied without choosing basis functions by numerical integration of Equation 6.10. [205]

The electronic wavefunction can be expanded in a basis set of orthonormal configurations $\{\phi_l\}$ or eigenstates $\{\tilde{\phi}_l\}$ of the form

$$\Psi(R,t) = \sum_l a_l(t)\phi_l(r;R) = \sum_l c_l(t)\tilde{\phi}_l(r;R) \tag{6.11}$$

It is convenient to use the eigenstate expansion in order to prove that non-adiabatic couplings are present in the Ehrenfest method. By substituting the expansion of Equation 6.11 into the expression of the electronic wavefunction in Equation 6.10, multiplying by a complex conjugate of the eigenstates ($\tilde{\phi}_k^*(r;R)$) and integrating out the electronic coordinate r gives

$$\begin{aligned} i\hbar\frac{\partial c_k(t)}{\partial t} &= c_k(t)E_k(R) - i\hbar\sum_l c_l(t)\left\langle \tilde{\phi}_k|\frac{\partial}{\partial t}\tilde{\phi}_l\right\rangle \\ &= c_k(t)E_k(R) - i\hbar\sum_{l,J} c_l(t)d_{lk}^J(R)\cdot\dot{R}_J \end{aligned} \tag{6.12}$$

where E_k are the eigenvalues and $d_{lk}^J(R) = \int \tilde{\phi}_k^*(r;R)\nabla_{R_J}\tilde{\phi}_l(r;R)dr$ which are the nonadiabatic couplings. This equation gives the time dependent variations of the amplitudes along an Ehrenfest classical trajectory, if the electronic wavefunc-

tion is expanded in the adiabatic basis. It is these equations that also define the time-dependence of the amplitudes of the electronic basis in the surface hopping methods introduced in Chapter 3.3. It should be noted that if the diabatic expansion of the electronic wavefunction is used, a similar expression can be obtained.

6.2.1 The Ehrenfest Multi Configurational Gaussian Method

As discussed in Chapter 3.3.1, the vMCG method is a variant of the MCTDH method, where all of the single particle functions of the MCTDH method have been replaced with Gaussian functions. The vMCG method has been further extended from a grid-based method to the DD-vMCG method whereby potential energy surfaces of a system are calculated on-the-fly utilising an external quantum chemistry software to provide the relevant energies, gradients and couplings. However, as shown in Chapters 4 and 5, systems which require a large number of excited states in order to describe the behaviour of the system fully represent a major limitation to the method. Consequently it can be seen that an Ehrenfest-type approach to the description of a system would be desirable, especially in the representation of ultrafast (attosecond) dynamics.

In order to incorporate the Ehrenfest approach into the vMCG method it is necessary to reformulate the equations of motion for the coefficients and the Gaussian basis functions of the vMCG ansatz. By introducing electronic functions of the form

$$|\psi_j(x_{el}, t)\rangle = \sum_s c_{js}(t)|s\rangle \tag{6.13}$$

the vMCG ansatz can be rewritten as

$$\Psi(R, t) = \sum_j B_j(t) g_j(R, t)|\psi_j(r, t)\rangle \tag{6.14}$$

where the expansion coefficients, B, are related to the vMCG expansion coefficients, A, by

$$A_{js} = B_j c_{js} \tag{6.15}$$

Applying the Dirac-Frenkel variational principle (Equation 3.102) to the new ansatz equation (Equation 6.14) by making variations of a coefficient,δB_i, a linear parameter of a Gaussian basis function, $\delta\xi_{i\alpha}$, and an electronic function, $\delta\psi_i$, obtains the following set of coupled equations

$$\sum_j iS_{ij}\dot{B}_j = \sum_j \left(H_{ij} - iS^{el}_{ij}\tau^g_{ij} - iS^g_{ij}\tau^{el}_{ij}\right) B_j \tag{6.16}$$

$$\sum_{j\beta} i\rho_{ij}S^{el}_{ij}S^{(\alpha\beta)}_{ij}\dot{\xi}_{j\beta} = \sum_j \left(\rho_{ij}H^{(\alpha 0)}_{ij} - i\rho_{ij}\tau^{el}_{ij}S^{(\alpha 0)}_{ij} - iB^*_i S^{el}_{ij}S^{(\alpha 0)}_{ij}\dot{B}_j\right) \tag{6.17}$$

$$\sum_j i\rho_{ij}S^g_{ij}|\dot{\psi}_j\rangle = \sum_j \left(\rho_{ij}H^g_{ij} - i\rho_{ij}\tau^g_{ij} - iB^*_i S^g_{ij}\dot{B}_j\right)|\psi_j\rangle \tag{6.18}$$

where the notation for the matrix elements is defined as

$$\left.\begin{array}{lllllllll} S^{el}_{ij} &=& \langle\psi_i|\psi_j\rangle; & S^g_{ij} &=& \langle g_i|g_j\rangle; & S_{ij} &=& \langle\psi_i g_i|\psi_j g_j\rangle \\ \tau^{el}_{ij} &=& \langle\psi_i|\dot{\psi}_j\rangle; & \tau^g_{ij} &=& \langle g_i|\dot{g}_j\rangle; & \rho_{ij} &=& B^*_j B_j \\ H^{el}_{ij} &=& \langle\psi_i|H|\psi_j\rangle; & H^g_{ij} &=& \langle g_i|H|g_j\rangle; & H_{ij} &=& \langle\psi_i g_i|H|\psi_j g_j\rangle \end{array}\right\} \tag{6.19}$$

It should be noted, at this point, that both the electronic functions and the Gaussian basis functions are normalised but are non-orthogonal

$$\left.\begin{array}{ll} S^{el}_{ij} = \langle\psi_i|\psi_j\rangle = 1 & i = j \\ = \sum_s c^*_{is}c_{js} & i \neq j \end{array}\right\} \tag{6.20}$$

If the spirit of the vMCG method is followed, Equation 6.16 can thus define an equation of motion for the expansion coefficient, B, as

$$i\dot{B}_k = \sum_{ij} S^{-1}_{ki}\left(H_{ij} - iS^{el}_{ij}\tau^g_{ij} - iS^g_{ij}\tau^{el}_{ij}\right) B_j \tag{6.21}$$

Multiplying this expression by B_j^* and substituting into Equation 6.17 provides an equation of motion for the linear Gaussian basis function parameters

$$i\dot{\xi}_{k\beta} = \sum_{ij\alpha} C^{-1}_{k\beta,i\alpha} \left(Y_{i\alpha} - iS^{(\alpha 0)}_{ij}\tau^{el}_{ij} - i\sum_{lm} S^{el}_{il} S^{(\alpha 0)}_{il} S^{-1}_{lm} S^{g}_{mj} \tau^{el}_{mj} \right) \tag{6.22}$$

Here the C and Y tensor are altered from the standard vMCG as

$$C_{i\alpha,j\beta} = \rho_{ij} \left(S^{el}_{ij} S^{(\alpha\beta)}_{ij} - \sum_{kl} S^{el}_{ik} S^{(\beta 0)}_{ik} S^{-1}_{kl} S^{(0\alpha)}_{kj} S^{el}_{kj} \right) \tag{6.23}$$

$$Y_{i\alpha} = \sum_{j} \rho_{ij} \left(H^{(\alpha 0)}_{ij} - \sum_{kl} S^{el}_{ik} S^{(\alpha 0)}_{ik} S^{-1}_{kl} H_{lj} \right) \tag{6.24}$$

If Equation 6.21 is substituted into Equation 6.18

$$\begin{aligned} \sum_j i\rho_{ij} S^g_{ij} |\dot{\psi}_j\rangle = \sum_j \Bigg(& \rho_{ij} H^g_{ij} - i\rho_{ij}\tau^g_{ij} \\ & - \sum_{kl} \rho_{il} S^g_{ij} S^{-1}_{jk} (H_{kl} - iS^{el}_{kl}\tau^g_{kl} - iS^g_{kl}\tau^{el}_{kl}) \Bigg) |\psi_j\rangle \end{aligned} \tag{6.25}$$

and then rearranged

$$\begin{aligned} i|\dot{\psi}_k\rangle = \sum_{ij} (\rho_{ki} S^g_{ki})^{-1} \Bigg[& \rho_{ij} \left(\left(H^g_{ij} - i\tau^g_{ij} \right) - \sum_{lm} S^g_{il} |\psi_l\rangle S^{-1}_{lm} \langle\psi_m| \left(H^g_{mj} - i\tau^g_{mj} \right) \right) \\ & + \sum_{lm} i\rho_{im} S^g_{ij} S^{-1}_{jl} S^g_{lm} \tau^{el}_{lm} \Bigg] |\psi_j\rangle \end{aligned} \tag{6.26}$$

the result is an equation of motion for the electronic wavefunction. As the Equations 6.21, 6.22 and 6.26 are a reformulation of the vMCG equations of motion, they are consequently exact solutions to the TDSE and allow the connection of Ehrenfest dynamics to full quantum dynamics. [212]

In order to solve these Ehrenfest equations of motion directly would require the substitution of τ^g in Equation 6.26 by τ^{el} in Equation 6.22, with an appropriate

choice for expression of the matrix τ^{el}. However, if the off-diagonal elements of the matrices S^g_{ij}, τ^g_{ij} and H^g_{ij} are ignored, the electronic functions become decoupled, leading to an expression for the equation of motion for the electronic wavefunction in the form

$$i|\dot{\psi}_k\rangle = \left(H^g_{kk} - H_{kk} + i\tau^{el}_{kk}\right)|\psi_k\rangle \tag{6.27}$$

These off-diagonal terms can be ignored if the Gaussian basis functions are sufficiently narrow, as in this case the off-diagonal terms will be small. The introduction of this approximation results in equations of motion that are no longer variational. However, if a sufficiently large basis set is used these equations of motion are still a full solution to the TDSE.

By defining the centre of a Gaussian basis function g_k as q_k,

$$H^g_{kk} = \langle g_k|H|g_k\rangle = \langle T_N\rangle_{kk} + H_{el}(q_k) \tag{6.28}$$

where $\langle T_N\rangle_{kk}$ is the kinetic energy of the GWP and $H_{el}(q_k)$ is the electronic Hamiltonian at q_k, the expression for energy of configuration k can be split into two parts

$$\langle g_k\psi_k|H|g_k\psi_k\rangle = \langle T_N\rangle_{kk} + \langle\psi_k|H_{el}(q_k)|\psi_k\rangle \tag{6.29}$$

As the on-diagonal elements of the matrix τ can be defined by an real number, by choosing

$$i\tau^{el}_{kk} = \langle\psi_k|H_{el}(q_k)|\psi_k\rangle \tag{6.30}$$

a simplified form of the equation of motion for the electronic functions can be written

$$i|\dot{\psi}_k\rangle = H_{el}(q_k)|\psi_k\rangle \tag{6.31}$$

As the Hamiltonian follows the centres of the Gaussian basis functions, the electronic function therefore follows the Gaussians. It can easily be seen that as the equations of motion for the Gaussian basis functions are variational, the trajectories of the nuclei are not classical. Although the electronic functions are not variational, the expansion coefficients and the nuclear functions are, and hence this defines the variational Ehrenfest multi-configurational Gaussian, veMCG, method.

In order to provide a more practical computational scheme, the electronic and nuclear parts can be fully decoupled, resulting in the nuclear functions following classical trajectories, the equations of motion for which are defined as

$$i\dot{B}_k = \sum_{ij} S^{-1}_{ki} \left(H_{ij} - iS^{el}_{ij}\tau^g_{ij} - iS^g_{ij}\tau^{el}_{ij} \right) B_j \tag{6.32}$$

$$\dot{\xi}_{k\alpha} = -2\zeta_{k\alpha}\frac{p_{k\alpha}}{m} - i\frac{\partial V}{\partial x_\alpha}\,|_{R=\mathbf{q}} \tag{6.33}$$

$$i|\dot{\psi}_k\rangle = H_{el}(q_k)|\psi_k\rangle \tag{6.34}$$

The use of these decoupled equations of motion defines the Ehrenfest multi-configurational Gaussian, eMCG, method. Although the MCE method of Shalashilin and co-workers [209] was developed and implemented using a different approach, it can be shown that the eMCG and MCE equations of motion are equivalent.

If the off-diagonal elements of the equation of motion for the expansion coefficients are also ignored, it can hence be rewritten as

$$i\dot{B}_k = \left(H_{kk} - i\tau^g_{kk} - i\tau^{el}_{kk} \right) B_k \tag{6.35}$$

In this form, the expansion coefficients can be considered to be an associated phase to each classical trajectory and although the matrix elements of $\hat{H}_{kk}$ and τ^{el}_{kk} are given, the matrix elements τ^g_{kk} can be chosen. The selection of differing values of τ^g_{kk} defines whether the phase is propagated in the Gaussian basis functions or in

the expansion coefficients. An example of this choice, as seen in Reference [212], is

$$i\tau^g_{kk} = H_{kk} \tag{6.36}$$

while alternatively, the vMCG choice would be

$$i\tau^g_{kk} = 0 \tag{6.37}$$

which results in equations of motion for the B coefficients of

$$i\dot{B}_k = \langle g_k | T_N | q_k \rangle \tag{6.38}$$

In addition to the full vMCG equations, in Reference [44] a hierarchy of GWP methods is proposed. The *classical MCG* (clMCG) method is defined by allowing the GWP basis functions follow classical trajectories, ignoring the non-classical terms of Y_R in Equation 3.140. In the *independent* MCG (iMCG) method, coupling between the basis functions is ignored, which then follow independent trajectories. Equation 6.38 is thus an equivalent formulation of the iMCG, method. The expansion coefficients of the iMCG method. A_{ks}, are the electronic function coefficients, c_{ks}, where each trajectory has a time-dependent normalisation factor, B_{ks}.

In order to test the eMCG method calculations were first run on a simple model problem, charge transfer in allene after sudden ionisation.

6.2.2 The Model Hamiltonian

In Mahapatra et. al. [213], the allene radical cation was found to have five low lying states, of which the $\tilde{A}(^2E)/\tilde{B}(^2B_2)$ manifold will be of particular interest here. In this previous work, a vibronic-coupling model Hamiltonian was set up for this molecule, incorporating the coupling between these states, in order to simulate the photo-electron spectrum. The vibronic coupling model uses a diabatic electronic

basis, i.e. the kinetic energy operator is diagonal and coupling between the states appears as off-diagonal functions in the potential operator matrix (Equation 2.33). A low-order Taylor series is then used to express the potential functions

$$W_{ij} = \langle\psi_i|H_{el}|\psi_j\rangle \tag{6.39}$$

$$= C^{(0)} + \sum_a C_a^{(1)} q_a + \sum_{ab} C_{ab}^{(2)} q_a q_b + \ldots \tag{6.40}$$

The expansion coefficients are thus derivatives of electronic Hamiltonian matrix elements evaluated at a suitable molecular geometry, usually the Franck-Condon point. For example, the linear terms are the first derivatives with respect to the coordinates

$$C_a^{(1)} = \frac{\partial}{\partial q_a}\langle\psi_i|H_{el}|\psi_j\rangle \tag{6.41}$$

For these terms, symmetry considerations can be used to show that they will be zero unless the product of the symmetries of the electronic states and the nuclear coordinate contain the symmetric representation

$$\Gamma_i \otimes \Gamma_j \otimes \Gamma_a \supset A_i \tag{6.42}$$

For allene, the nuclear coordinates used are the neutral ground-state mass-frequency scaled normal mode vibrations. Neutral allene at the ground-state equilibrium geometry has D_{2d} symmetry. Thus if the symmetry of the system is considered, the product of the doubly degenerate irrep is $E \otimes E = A_1 \oplus B_1 \oplus B_2$, and only modes of these three symmetries can have non-zero linear terms in the Hamiltonian for a doubly degenerate state. In the $\tilde{A}$ state the three A_1 and three B_2 modes thus provide the linear on-diagonal coupling terms, and the one B_2 vibration the off-diagonal. To first-order, the Hamiltonian matrix can hence be written

$$\begin{aligned}\mathbf{H} &= \sum_{i=1}^{15} \frac{\omega_i}{2}\left(-\frac{\partial^2}{\partial Q_i^2} + Q_i^2\right)\mathbf{1} + \begin{pmatrix} E_E & 0 & 0 \\ 0 & E_E & 0 \\ 0 & 0 & E_{B_2} \end{pmatrix} \\ &+ \begin{pmatrix} \sum_{i=1}^{3} \kappa_i Q_i & 0 & 0 \\ 0 & \sum_{i=1}^{3} \kappa_i Q_i & 0 \\ 0 & 0 & \sum_{i=1}^{3} \kappa_i' Q_i \end{pmatrix} \\ &+ \begin{pmatrix} \sum_{i=5}^{7} \lambda_i Q_i & \lambda_4 Q_4 & \sum_{i=8}^{11} \lambda_i' Q_{ix} \\ \lambda_4 Q_4 & \sum_{i=5}^{7} -\lambda_i Q_i & \sum_{i=8}^{11} \lambda_i' Q_{iy} \\ \sum_{i=8}^{11} \lambda_i' Q_{ix} & \sum_{i=8}^{11} \lambda_i' Q_{ix} & 0 \end{pmatrix} \end{aligned} \tag{6.43}$$

where the normal mode Q_i corresponds to the vibration ν_i defined in Table 6.1. It should be noted that the full model is expanded to second order, the details and values of the parameters for which can be seen in Reference [213].

At the Franck-Condon point, $\mathbf{Q} = 0$ (termed $\mathbf{Q}_0$), the $\tilde{A}$ state is doubly-degenerate and forms a Jahn-Teller conical intersection, with the degeneracy lifted on the diagonal elements of the Hamiltonian matrix by the three B_2 modes. The degeneracy is lifted in the off-diagonal, coupling, Hamiltonian element by the B_1 mode. This is the torsional mode, with the end H–C–H units rotating towards a planar geometry, and the doubly degenerate $\tilde{A}$ state is subject to the rare $e \otimes B$ Jahn-Teller distortion, whereby the degeneracy is lifted by vibrations with B_1 and B_2 symmetry. This state is also further pseudo-Jahn-Teller coupled to the $\tilde{B}$ state via doubly-degenerate E vibrations.

The degenerate highest occupied molecular orbitals (HOMO) of the diabatic electronic basis functions for the neutral allene at the conical intersection are shown in the top panel of Figure 6.2. They are each localised on one H–C–H group. Ionisation of allene thus creates an electron "hole" in one end of the molecule. While the character of the MO is retained moving along the Jahn-Teller active B_2 modes, movement along the B_1 mode couples the two states and leads to population transfer between the states. This corresponds to charge being transferred between the ends of the molecule.

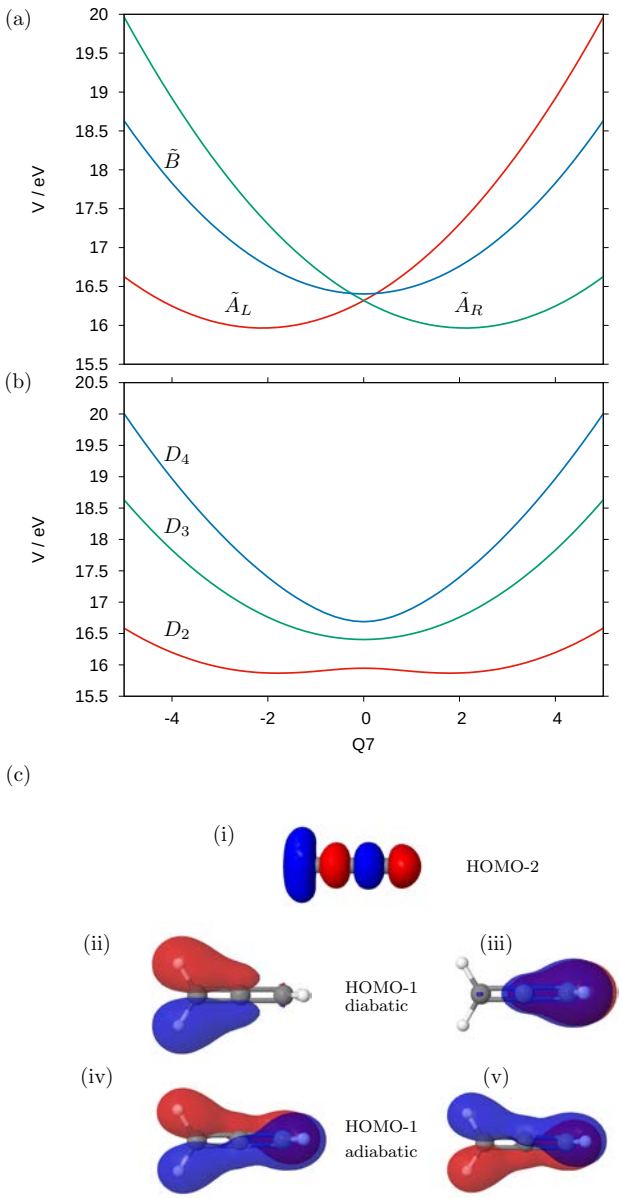


Fig. 6.2: The (a) diabatic and (b) adiabatic potential energy surfaces for the allene radical cation model Hamiltonian along the $Q_7(B_2)$ vibrational coordinate with a torsion angle kept at 45°. The molecular orbitals shown in the top panel define the diabatic surfaces, while those in the lower show the conjugation obtained at 45° due to the non-adiabatic coupling.

Charge migration takes place when the charge moves without nuclear motion. If the molecule remains at $\mathbf{Q}_0$, this does not occur as the coupling is zero. However if the molecule is twisted away from the equilibrium geometry along the B_1 torsion mode, the diabatic coupling means that the eigenfunctions of the ion are a superposition of the states populated by ionisation from the neutral molecule. Consequently, if the neutral degenerate states are labeled ψ_L and ψ_R, the ion state can be written

$$\psi_+ = c_L(t)\psi_L + c_R(t)\psi_R \tag{6.44}$$

If the ionisation removes the electron from ψ_L then at $t = 0$ the coefficients are $c_L(0) = 1$ and $c_R(0) = 0$. The coefficients evolve according to Rabi-like oscillations due to the coupling, and charge migration occurs. The MOs of the cation at a torsion angle of 45° are shown in the lower panel of Figure 6.2, along with the adiabatic surfaces, showing how the coupling leads to conjugation along the chain allowing the charge migration.

Label	Frequency (cm^{-1})	Description	MCTDH Basis Set N	n
$\nu_1(A_1)$	3015	HCH sym str	24	13
$\nu_2(A_1)$	1443	HCH in-phase bend	32	
$\nu_3(A_1)$	1072	CCC sym str	32	14
$\nu_4(B_1)$	865	HCCH torsion	32	
$\nu_5(B_2)$	3407	HCH oo-phase str	10	17
$\nu_6(B_2)$	1957	CCC anti-sym str	10	
$\nu_7(B_2)$	1398	HCH oo-phase bend	30	
$\nu_8(E)$	3486	HCH anti-sym str	10	1
$\nu_9(E)$	999	HCH rock	18	
$\nu_{10}(E)$	841	HCH wag	12	
$\nu_{11}(E)$	355	CCC bend	8	

Table 6.1: Definitions and descriptions of the 15 vibrational modes in the allene molecule. Frequencies are for the ground-state calculated at the MP2/cc-pVTZ level. N are the number of harmonic oscillator DVR functions used in the MCTDH calculations, and n the number of multi-dimensional single particle functions with the combination of modes included by the bracket.

By comparing the system dynamics of static nuclei to that of moving nuclei, the diabatic state populations of this simple model can thus be used to observe the signature of charge migration in charge transfer.

6.3 Results

6.3.1 Charge Migration versus Charge Transfer

Using the model Hamiltonian of allene described above, two different sets of calculations were performed using full grid-based quantum dynamics and the MCTDH method. At the equilibrium geometry the torsion angle between the H–C–H groups is 90°. By altering the torsion angle the initial wavepacket experiences coupling between the degenerate orbitals thus allowing charge migration to occur. The first set of calculations were started with a torsion angle of 75°, giving a small coupling between the degenerate orbitals, while the second set of calculations were started with a torsion angle of 45°, giving a larger coupling. As the model is set up in normal mode coordinates, a change in torsion angle requires changing not only the coupling mode but also the totally symmetric modes to keep the bond lengths correct. In the mass-frequency scaled normal mode coordinates of the model, an angle of 75° corresponds to a coordinate of $(Q_1, Q_2, Q_3, Q_4) = (-0.1308, -0.0556, 0.0100, 1.2543)$ and an angle of 45° corresponds to a coordinate of $(Q_1, Q_2, Q_3, Q_4) = (-1.1635, -0.4945, 0.0892, 3.6774)$ with all other coordinates having a value of 0.

The initial wavepacket is centred at the starting geometry for each model, i.e. displaced from the Franck-Condon point, $\mathbf{Q}_0$, with the width appropriate for the neutral ground-state vibrational frequencies. The simulation is initiated by placing this packet in the second diabatic state, corresponding to making a hole at the right-hand end of the twisted molecule upon ionisation. The population dynamics following this ionisation for both initial torsion angles are shown in Figure 6.3. It

should be noted that the coupling to the B-state was ignored as this plays no role in the charge migration.

Figure 6.3 (a) and (e) show the charge migration dynamics for the two angles calculated using the gMCG method with a single time-independent GWP, i.e. static nuclei. The state populations show a simple oscillation between the two ends of the molecule, with a period related to the coupling strength. The populations obtained from full quantum dynamics calculated using the MCTDH method, where the basis functions are the SPF, allowing the nuclei to move are shown in Figure 6.3 (b) and (f). Although the charge migration oscillations are damped, the signal is still visible in the first 10 fs, particularly in the $\Theta = 45°$ model.

Figure 6.3 (c) and (g) show the results of calculations using the vMCG method with 50 GWPs, where the dotted line represents the MCTDH result. The widths of the GWPs were taken as $1/\sqrt{2}$ in each direction which, in the mass-frequency scaled coordinate system of the Hamiltonian, corresponds to the width of the neutral ground-state wavepacket. It can clearly be seen that the vMCG method is able to describe the coupled nuclear and electronic dynamics of this system with a very small number of functions.

The final plots, Figure 6.3 (d) and (h), show the result with a single GWP following an Ehrenfest classical trajectory from the starting geometry. It is clear that the nuclear dynamics is not correctly described: the charge migration signal dominates and charge localisation occurs, with the dynamics ending at one end of the molecule rather than spread over the whole molecule as seen in the quantum dynamics. Interestingly the charge is at different ends depending on the initial torsion angle.

As the charge migration is damped in the model starting with a torsion angle of $\Theta = 45°$, in order to show the dynamics of the nuclei the expectation values of the three most active vibrational modes are plotted in Figure 6.4. These three

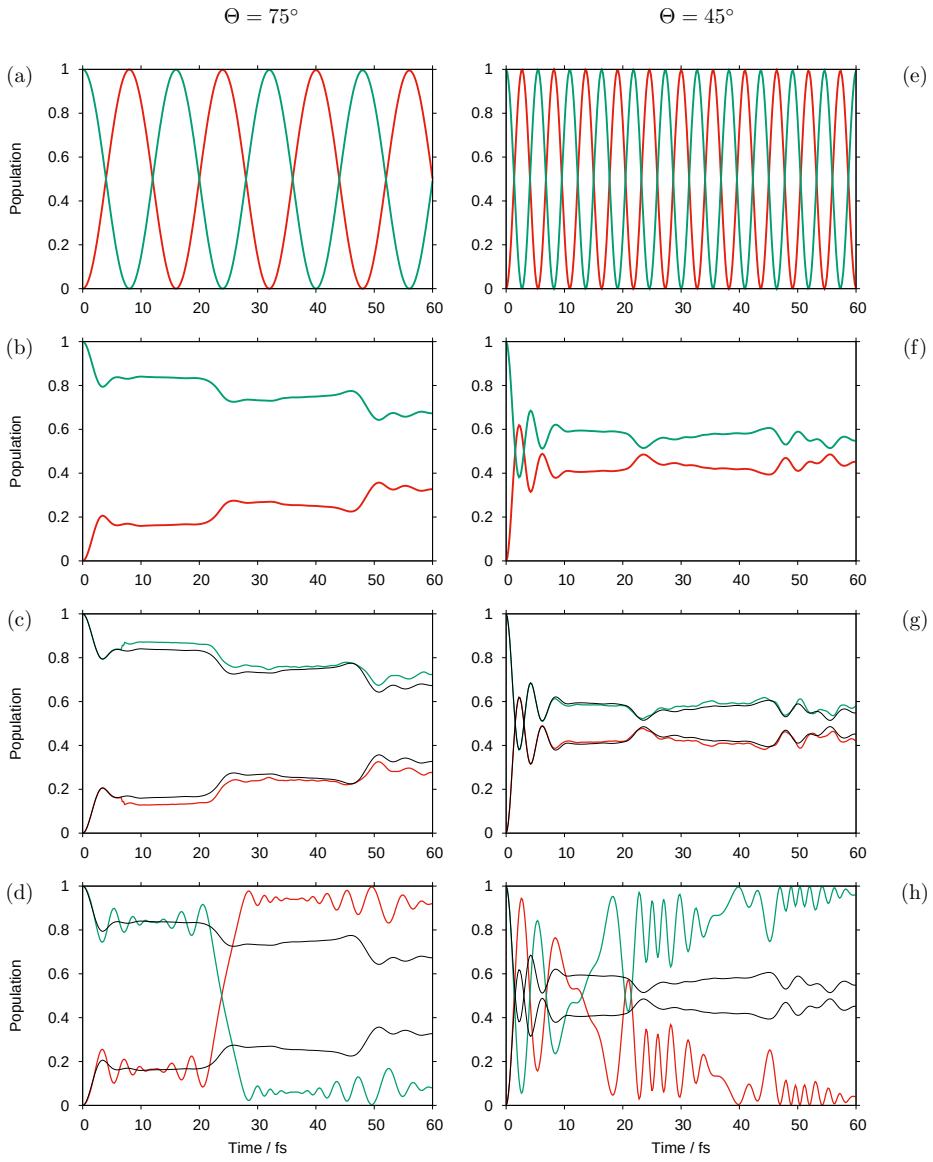


Fig. 6.3: Population dynamics of allene after ionisation on the right hand end of the molecule. In (a) - (d) the molecule had a torsional twist of $\Theta = 75°$ and in (e) - (h) the angle was $\Theta = 45°$. (a), (e) Charge migration with static nuclei. (b), (f) full quantum dynamics of charge migration and charge transfer using the MCTDH method. (c), (g) Full quantum dynamics using the vMCG method with 50 GWPs. The MCTDH result is the black dotted line. (d), (h) Population dynamics with a single GWP, i.e. classical nuclei.

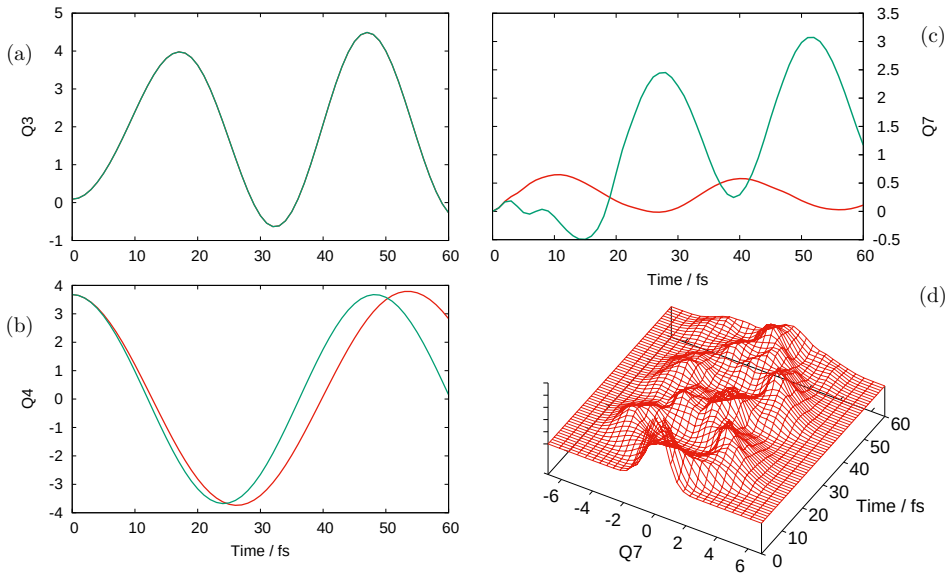


Fig. 6.4: Expectation Values of the main vibrational coordinates (a) Q_3 (b) Q_4 (c) Q_7 from simulations starting with a torsion angle of $\Theta = 45°$. In green are the results from the MCTDH calculation, in red the classical Ehrenfest trajectory from the Franck-Condon point. (d) The reduced density along Q_7 is shown as a function of time.

significant active modes are the B_1 coupling torsional vibration, Q_4, and the lowest frequency modes in the A_1 and B_2 symmetries, Q_3 and Q_7 respectively. Results from the full quantum dynamics simulations and the classical Ehrenfest trajectory are shown. The motion followed by the wavepacket along the symmetric mode Q_3 is classical, while the motion along the B_2 Jahn-Teller active mode is clearly very non-classical in nature. The reduced density along Q_7 is shown in Figure 6.4 (d), where a bifurcated waveform motion along this mode can be seen, something that a single classical trajectory cannot follow.

This model thus presents a significant challenge for a method in the description of coupled electronic-nuclear motion in a charge-transfer problem. The dynamics can be divided into the short-time (< 10 fs) during which charge migration dominates, and long-time (> 10 fs) when nuclear motion may lead to a permanent charge transfer.

6.3.2 Different Dynamics Models: Initial Conditions and Convergence

In a further set of calculations, the different levels of theory outlined above were used to see how well the coupled electronic-nuclear dynamics of the allene ionisation could be followed. When the non-variational GWPs of the clMCG, iMCG and eMCG methods are used, the initial positions of the functions play a key role in the dynamics, and for good results it is important to cover phase space. To enable easier convergence of the results, in the following only the three most important modes, Q_3, Q_4 and Q_7, will be included in the dynamics with the initial wavepacket kept stationary along the other modes. The initial wavepacket in all cases is that with $\Theta = 45°$.

The state populations as a function of time, calculated with the different levels of theory, are shown in Figure 6.5. In Figure 6.5(a) the vMCG result with 30 GWPs is compared to the full quantum dynamics result (dotted line). The charge-migration and charge-transfer dynamics are very similar to those seen in the 15-dimensional calculations above, showing the dominance of the three selected modes in the nuclear dynamics. The charge migration can clearly be seen at short timescales with a full oscillation in under 10 fs before the long-time damping and accompanying charge transfer. Again, the accuracy of the vMCG method is demonstrated, as fewer functions are needed than in the 15-dimensional calculation, indicating the smaller phase space that must be covered.

The best result obtained using classical GWPs at the clMCG level is shown, in Figure 6.5(b), where 150 functions were used. When the number of functions was increased, the result did not noticeably improve while with fewer basis functions the populations displayed stronger oscillations across the entire propagation. The short-time charge migration dynamical behaviour was reproduced well in every case. In these calculations the initial sampling of the GWPs were taken from a

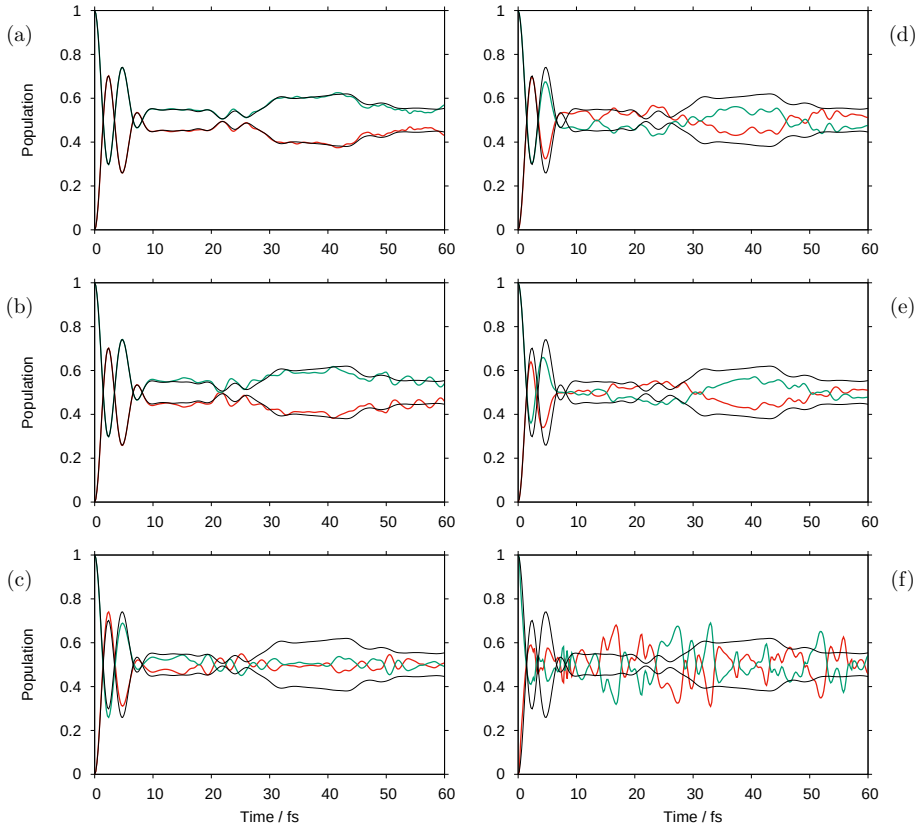


Fig. 6.5: Population dynamics of allene with a torsion angle of $\Theta = 45°$ after ionisation on the right hand end of the molecule calculated using different levels of theory and including only the main modes, Q_3, Q_4 and Q_7 in the dynamics. (a) The vMCG method with 30 variational GWPs with a width $1/\sqrt{2}$. (b) The clMCG method with 150 classical GWPs with a width $1/\sqrt{2}$ (c) The clMCG method with 200 classical GWPs with a width 0.4. (d) The iMCG method with 200 independent classical GWPs with a width 0.4 (e) The iMCG method with 200 independent classical GWPs with a width 0.4 and initial equal weights (f) The eMCG method with 150 classical GWPs with a width $1/\sqrt{2}$ In the iMCG, clMCG and eMCG calculations the initial GWPs were taken from a Wigner distribution. The full quantum result is shown in all plots as a black dotted line.

Wigner distribution of the ground-state neutral wavepacket. This should lead to faster convergence than simply taking random positions in configuration space, and indeed it produced better results with fewer functions.

In the vMCG and previous clMCG calculation, the widths of the GWPs were taken as $1/\sqrt{2}$ in each direction. Although it may be advantageous to take narrower GWPs in order to, for example minimise the error due to using a local harmonic approximation in the calculation of the integrals, when narrower GWPs are used a fit must be made of the initial wavepacket in the basis set. This was done by projecting the initial wavepacket on to the basis set, i.e. the initial coefficients were chosen by

$$A_{js} = \sum_b |g_j\rangle S^{-1}_{jb} \langle g_b|\psi(0)\rangle \tag{6.45}$$

with the index s set to be the initial state. When these narrower GWPs are used, more are required in order to cover phase space, and the quality of the result drops. A calculation using the clMCG level with 200 GWPs with a width of 0.4 is shown in Figure 6.5(c), where the convergence on the full result is poorer than the convergence displayed by the 150 wider GWPs. It should be noted that even with 200 GWPs the initial wavepacket is not exactly represented as the expectation values of the 3 modes are only approximately correct and the widths along all modes are less than $1/\sqrt{2}$ in coordinate space and greater than $1/\sqrt{2}$ in momentum space.

In Figure 6.5(d) the result from an iMCG calculation also with 200 GWPs is shown. Given the simplified representation of the evolving wavepacket the result is remarkably good, in that the shape of the population transfer is correct, and the short-time dynamics is reproduced well, though the oscillations are slightly damped compared to the full result. The long-time dynamics are not exactly reproduced due to the lack of nuclear coherence, which allows crossings of population not seen

in the full quantum result. The magnitude of the result is, however, reasonable. In Figure 6.5(e) the result from an iMCG calculation is shown where 200 GWPs starting with equal weights have been used, meaning that the initial wavepacket is defined by the Wigner distribution rather than by fitting. This choice makes little difference and shows that the representation of a wavepacket by many Ehrenfest trajectories selected utilising the Wigner function is a reasonable way to treat the system.

The final calculation, shown in Figure 6.5(f), is the result from an eMCG simulation with 150 GWPs using widths of $1/\sqrt{2}$. For this simulation, 200 classical trajectories, sampled using a Wigner distribution, were initially propagated and the coordinates and momenta along with the Ehrenfest potentials and state populations for each stored. The eMCG basis functions were then run along these trajectories using spline fits of the data and the evolution of the expansion coefficients calculated. This is the procedure that could be used to run an eMCG simulation along Ehrenfest trajectories calculated by a quantum chemistry program, such as Gaussian, which have been generated and used in a number of studies on charge migration [200–202].

It was hoped that this method would give a quality similar to the clMCG, but while the initial decay and the period of the charge migration oscillations in the short-time dynamics is reproduced, the oscillations are not correctly damped. This is likely, however, to be due to the implementation as, for example, the integrals used only the value of the potential at the centre rather than the full LHA.

6.3.3 Direct Dynamics Calculations

The allene $\tilde{A}/\tilde{B}$ manifold is unfortunately not suitable for direct dynamics calculations of charge migration, as in the model system used above the $\tilde{B}$-state was simply ignored, which is something that cannot be done in quantum chemistry simulations. Furthermore, it was found upon investigation of the electronic structure

of the allene cation, using CASSCF methods, that there is in fact another doubly degenerate state between the $\tilde{A}$ and $\tilde{B}$. This additional state is mostly characterised by double excitations so was not seen by the single-determinant methods used in the original study [213]. The inclusion of these additional states would require a coupled set of seven states, and no clear connection to charge migration. For this reason it was decided to instead study the doubly degenerate ground-state, $\tilde{X}(^2E)$ which can be described by orthogonal π-orbitals localised at opposite ends of the molecules. These are shown in Figure 6.6(c),(d). Figure 6.6(a),(b) show the conjugated orbitals that have D_{2d} symmetry and are formed from combinations of the localised π orbitals

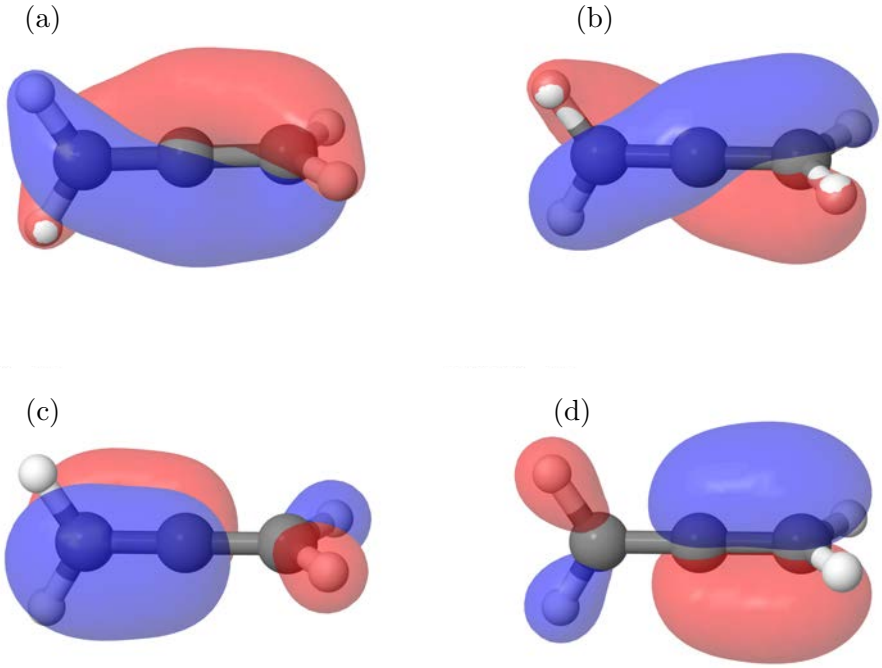


Fig. 6.6: Allene doubly degenerate HOMO that are ionised to form the $\tilde{X}$ cation state. (a),(b) Conjugated orbitals with D_{2d} symmetry (c),(d) Localised π orbitals

All calculations were performed using a 6-31G$*$ basis set at the CAS(4,3) level of theory, including the HOMO and LUMO π orbitals. Following the philosophy of the direct dynamics protocol in Chapter 4.3.1, in order to initially populate the database, a direct dynamics simulation was started on the first excited state of the

Mode	Symmetry	Frequency (cm^{-1})	Character
ν_{11}	E	369.95	CCC bend
ν_{11}	E	369.95	CCC bend
ν_{10}	E	877.43	HCH wag
ν_{10}	E	877.43	HCH wag
ν_4	B_1	908.85	HCCH torsion
ν_9	E	1052.00	HCH rock
ν_9	E	1052.00	HCH rock
ν_3	A_1	1113.63	CCC sym str
ν_7	B_2	1478.70	HCH oo-phase bend
ν_2	A_1	1526.86	HCH in-phase bend
ν_6	B_2	2064.70	CCC anti-sym str
ν_1	A_1	3206.49	HCH sym str
ν_5	B_2	3206.75	HCH oo-phase str
ν_8	E	3292.17	HCH anti-sym str
ν_8	E	3292.17	HCH anti-sym str

Table 6.2: Frequencies of the normal modes of the neutral allene molecule at the MP2/6-31G* level of theory. The labels are the Wilson nomenclature used in the model Hamiltonian. The symmetry labels are for the D_{2d} point group to which the molecule belongs. Note the accidental degeneracy of modes ν_2 and ν_5

cation using the normal modes and frequencies of the neutral ground-state to define the coordinates. These were calculated using a 6-31G$*$ basis set at the MP2 level of theory, the results of which are given in Table 6.2. The initial geometry was the neutral ground-state equilibrium geometry with a torsion angle of 75°. Due to the occurrence of a conical intersection located at the Franck-Condon point (a torsion angle of 90°), it is not possible to start at this point as the quantum chemistry cannot calculate the Hessian due to the cusp in the adiabatic surfaces. The initial propagation was for 60 fs with 5 GWPs, with widths given by the neutral ground state harmonic frequencies and displaced in momentum space around the initially populated basis function. This database was then used for a second propagation starting at 45° with 25 GWP in addition to a calculation starting at 75° also with 25 GWP.

Cuts through the potential surfaces are shown in Figure 6.7 and 6.8. In Figure

6.7 the diabatic couplings are shown. The cuts are taken keeping all coordinates not plotted at the equilibrium and the couplings should obey the same symmetry rules as used for building the vibronic coupling model of the $\tilde{A}/\tilde{B}$ manifold. Thus the coupling along the three totally symmetric modes ν_1, ν_2 and ν_3 should be zero as totally symmetric modes cannot couple different states. The plots in Figure 6.7(b), (d) and (f) show that the couplings are indeed small, but there is an unexpected peak at the origin, maybe due to the quantum chemistry breaking down at the conical intersection.

In the model, the B_1 mode provided the coupling. However, the torsional mode $\nu_4(B_1)$ shows no coupling (except for a spurious peak close to Q=0) and the coupling is in the B_2 modes, in particular ν_6 which exhibits a fairly smooth linear coupling. The different roles played by these modes is due to the properties of degenerate eigenstates as it is possible to rotate the electronic basis so that the B_1 and B_2 modes swap character and modes with B_2 symmetry provide the coupling instead.

The potential energy surfaces, both adiabatic and diabatic, are shown in Figure 6.8, with cuts shown along the $\nu_4(B_1)$, and $\nu_6(B_2)$ modes, the modes responsible for the Jahn-Teller coupling, as well as the low frequency symmetric mode $\nu_3(A_1)$. The surfaces along ν_6 show the correct behaviour - degenerate in the diabatic representation, but splitting in the adiabatic representation due to the coupling. Although degeneracy should also be represented in the $\nu_4(B_1)$, and $\nu_3(A_1)$ modes, the surfaces for these modes do not show this behaviour. This is presumably due to the spurious coupling at the Franck-Condon point. There is also a clear error in the adiabatic surfaces along both of these modes. Despite these errors, the diabatic surfaces on which the dynamics takes place are smooth due to the interpolation procedure.

State populations from these DD-vMCG simulations with a torsional twist and

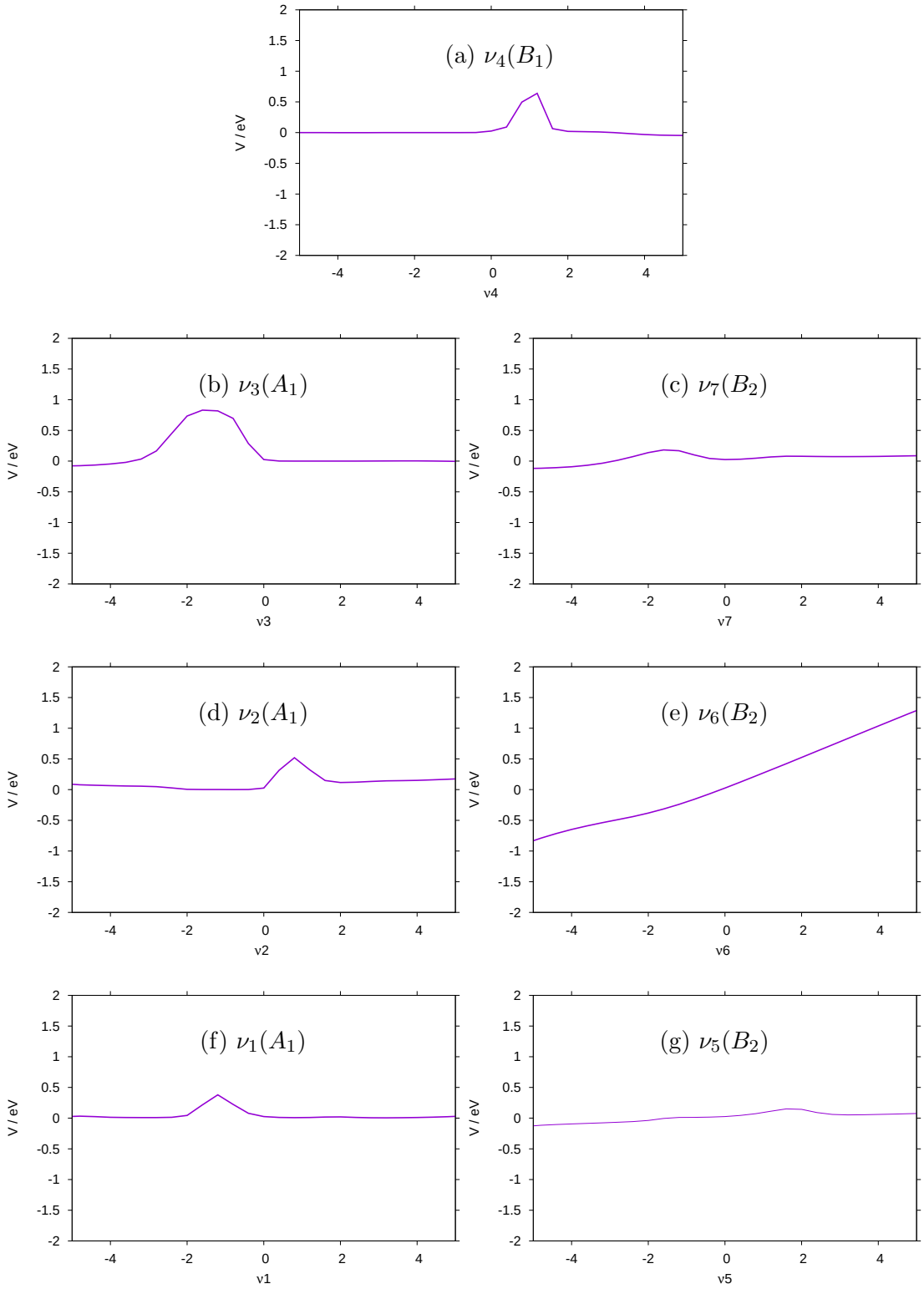


Fig. 6.7: Diabatic coupling from the SA2-CAS(4,3)/6-31G* potential energy surfaces of allene from DD-vMCG simulations. Cuts along the non-degenerate vibrations with all other coordinates kept at the equilibrium geometry.

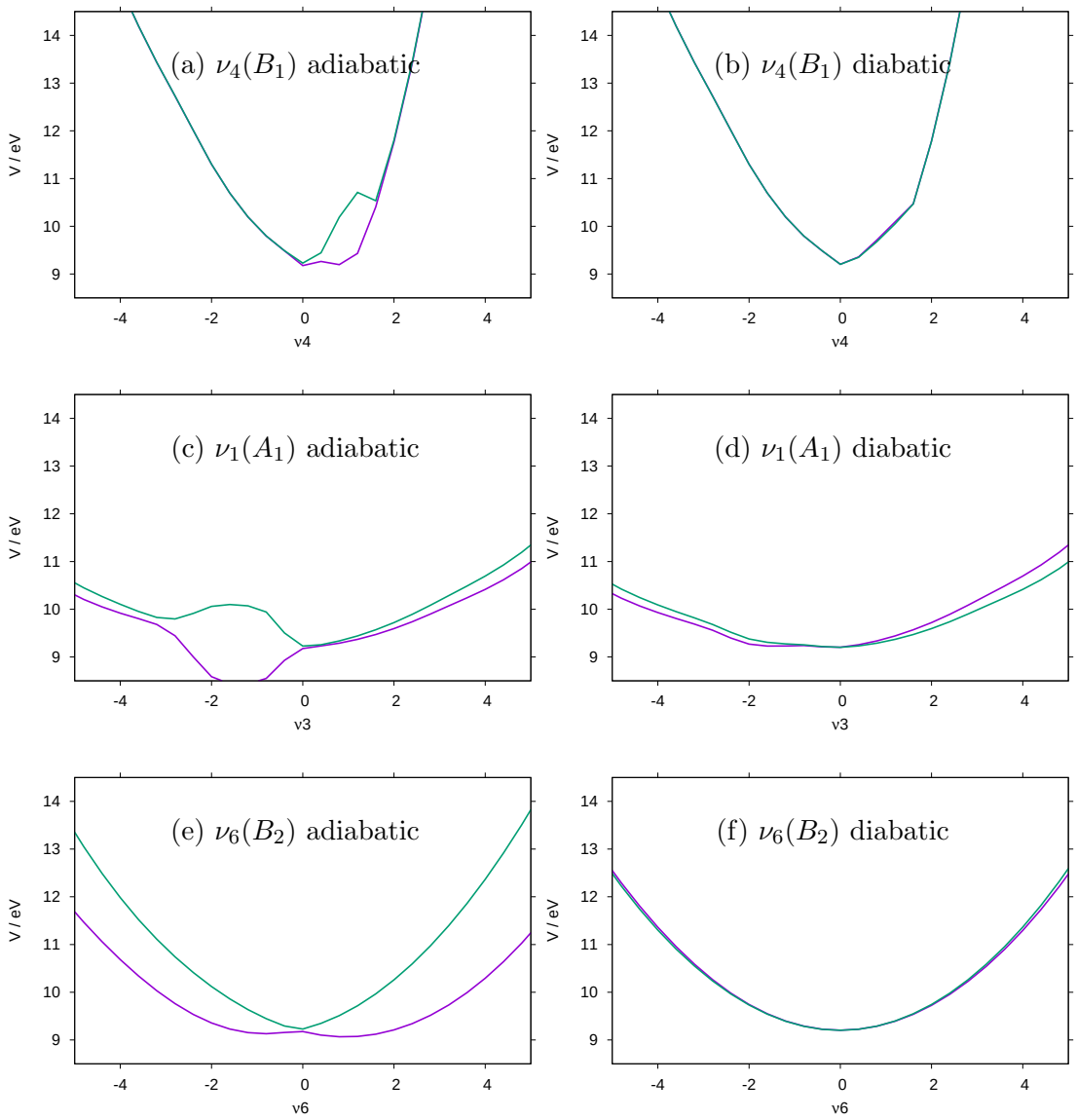


Fig. 6.8: Potential energy surface from the SA1-CAS(4,3)/6-31G* potential energy surfaces of allene from DD-vMCG simulations. Cuts along the key non-degenerate vibrations with all other coordinates kept at the equilibrium geometry.

vertical excitation are shown in Figure 6.9. Transfer takes place at a faster rate when starting at 45° rather than at 75° due to the increased coupling, and an equilibrium is reached, with a 50:50 population in each state, by 60 fs in both cases. An important point to note is that, unlike in the vibronic coupling model these simulations are not related to charge migration. As shown above, the diabatic coupling in the direct dynamics is not along the torsional mode. This is because

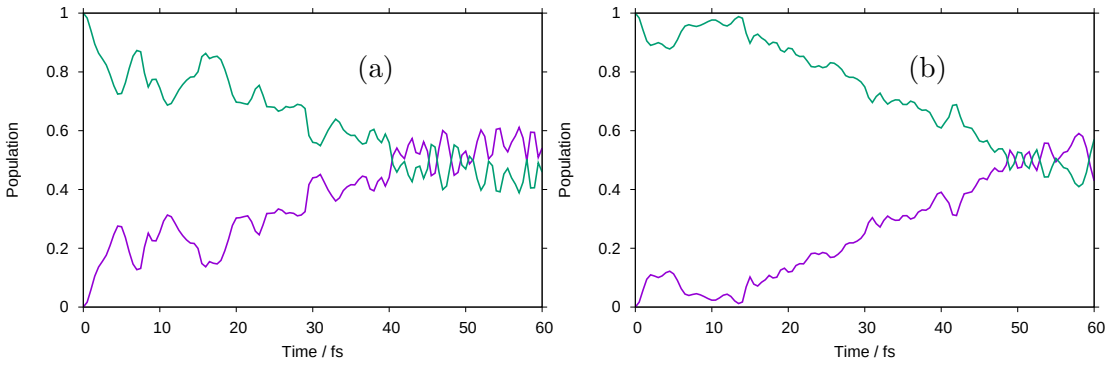


Fig. 6.9: State populations from DD-vMCG simulations of the allene radical cation with SA2-CAS(3,4)/6-31G* potential surfaces. All the population was on the first excited state with the geometry displaced from the Franck-Condon to have a torsion angle of (a) 45° and (b) 75°.

the Quantum chemistry calculations produces CAS orbitals at the equilibrium geometry, which are the diabatic basis, that are not the localised π-orbitals on the different ends of the molecule, but the delocalised orbitals along the molecule.

In order to initiate a calculation localised at one-end, a wavepacket must be created by taking a superposition of eigenstates, i.e.

$$\Psi(t=0) = \frac{1}{\sqrt{2}}(\psi_1 + \psi_2) \tag{6.46}$$

The population of the states at either end of the molecule are then given by the projector of this initial wavefunction onto the evolving wavepacket. Populations for the localised states using this initial superposition are shown in Figure 6.10.

Three calculations were made for each initial angle. In the first, full DD-vMCG dynamics were run and the localised populations are shown in Figure 6.10(a) and (b). The second set of calculations used a single, time-independent GWP in order to represent the electron dynamics only and hence model the charge migration process (Figure 6.10 (c), (d)). A fast oscillation is seen between the orbitals at the two ends of the molecule in the 45° case, while a much slower transfer with the smaller torsion angle of 75°. The results are similar to the model calculations on the $\tilde{A}$ manifold, with the charge migration at 45° fast enough to leave a signal in

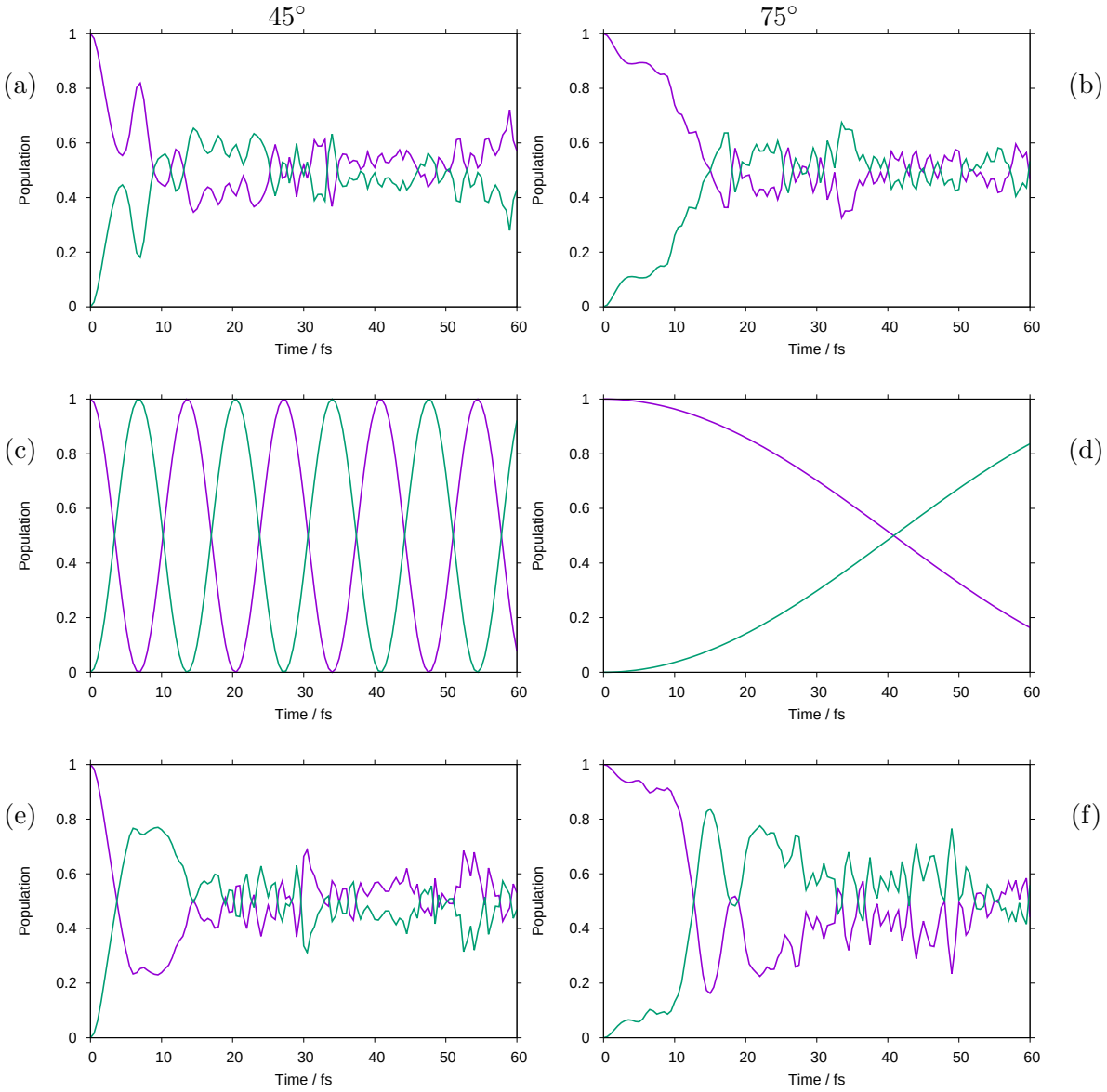


Fig. 6.10: Localised orbital populations from DD-vMCG simulations of the allene radical cation with SA2-CAS(3,4)/6-31G* potential surfaces. The initial wavepacket was localised at one end of the molecule with the geometry displaced from the Franck-Condon to have a torsion angle of either (a),(c),(e) 45° or (b),(d),(e) 75°. (a) and (b) are results from full DD-vMCG calculations. (c) and (d) use a single time-independent wavepacket to simulate charge migration. (e), (f) are DD-vMCG simulations with classical GWPs.

the populations that is not seen at 75°. The 45° population transfer is, however, slower than in the $\tilde{A}$ state pointing to a smaller coupling in the $\tilde{X}$ state. The final calculations were DD-vMCG calculations, but instead using GWPs that follow classical trajectories (Figure 6.10 (e), (f)). A qualitatively similar behaviour is seen to the full DD-vMCG results, but the damping occurs at a slower rate likely due to the fact that a basis of GWPs following classical trajectories cannot follow the full quantum result.

6.3.4 Direct Ehrenfest Dynamics Calculations

In recent years the Robb group at Imperial College London have implemented a classical Ehrenfest propagation method into the Gaussian quantum chemistry program. This has been to great effect in the study of how nuclear motions damp the natural charge migration in a range of molecules [200–202, 214]. However, as has been shown above, classical trajectories may not be suitable for capturing accurately the full quantum motion of the coupled nuclear-electron motion. For this reason, an aim of this project was to couple the DD-vMCG algorithm in Quantics with the Ehrenfest trajectory algorithm in Gaussian to provide a quantum description of nuclei moving over a time-dependent Ehrenfest potential.

The first trial was to precompute a set of Ehrenfest trajectories using Gaussian and then propagate clMCG GWPs along these trajectories to simulate the evolving wavepacket. The performance of this method is shown in the model studies above, where this procedure was used in the eMCG calculations of Section 6.3.2. The quality of these results was, however, not good enough to warrant proceeding with this approach.

The second trial was to use the Ehrenfest surfaces directly directly from Gaussian. This approach, however, required the implementation of second derivatives in the Ehrenfest method. This has recently been accomplished by M. Vacher and A. Jenkins. [205] Figure 6.11 shows the first DD-vMCG results using the second-order

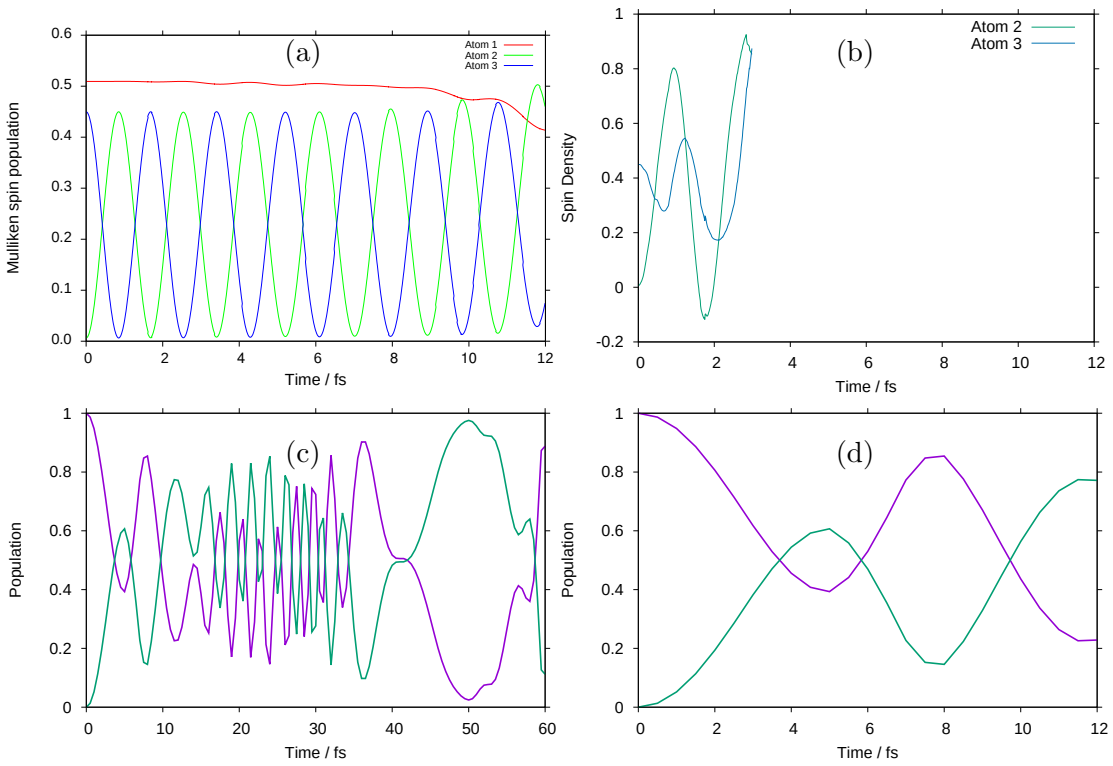


Fig. 6.11: Charge transfer using DD-vMCG with Ehrenfest surfaces at a CAS (4,3)/6-31G* level of theory and an initial electronic wavepacket equally populating the electronic states. The initial geometry is with a 45° torsion angle. (a) Spin densities on the atoms from a simulation with 1 GWP. Atoms 2,3 are the end carbons. Atom 1 is the central carbon. (b) Spin densities on the atoms from a simulation with 12 GWP. (c) State populations from a DD-vMCG calculation with an explicit 2-state representation and 1 GWP. (d) As (c) but showing only the first 12 fs.

Ehrenfest surfaces.

Figure 6.11(a) shows the results from a coupled DD-vMCG-Ehrenfest calculation with a single nuclear wavepacket, starting as above with an electronic wavepacket equally populating the two electronic states. Rather than state populations, the spin density on the atoms in the allene chain are shown. Atoms 2 and 3 are the terminal carbon atoms, with atom 1 being the carbon in the centre. A fast charge migration between the 2 ends is seen, with little damping by the nuclear motion. A calculation with 12 GWPs is shown in Figure 6.11(b). The spin densities from the trajectories at the centres of the GWPs are weighted by the Gross Gaussian Populations. Unfortunately the integrator failed after only 3

fs, but there is definite damping of the migration so that the charge has localised on one end.

DD-vMCG state populations for a single wavepacket using the coupled 2-state representation as before are shown for comparison in Figure 6.11(c). This GWP follows a classical Ehrenfest trajectory. The dynamics is clearly different from that shown in the 24 GWP calculation of Figure 6.10(a), showing the lack of damping in a classical Ehrenfest simulation. Figure 6.11(d) repeats this plot, showing only the first 12 fs. On comparison to the result from the DD-vMCG-Ehrenfest simulation in Figure 6.11(a), the time period for the oscillation is very different. The reason for this, which may be due to the time steps used in this latter simulation missing the fast oscillations, or due to errors in the Ehrenfest calculation giving a spuriously large coupling, are yet to be investigated.

6.4 Summary and Conclusions

In this Chapter a new method for the study of electronic wavepacket motion coupled to nuclear dynamics was presented and tested. The Ehrenfest approach provides a alternative conceptual approach for Direct Dynamics requiring only a single, time-dependent potential energy surface. The key concepts in the use of this Ehrenfest method are that the calculations are intended to complement attosecond timescale experiments in which the dynamics of the electrons of a system are observed and hence the competition between charge migration (electronic motion over a static framework) and charge transfer (electronic motion as a result of nuclear motion). Existing Ehrenfest methods utilise classical nuclei exist whereas the new Ehrenfest method utilises quantum nuclei.

Test calculations on a model system, the allene cation, were used in comparison of methods available within the Quantics software package in order to benchmark the Ehrenfest method. The allene radical cation presented an ideal system for this

initial testing as precalculated surfaces representing the $\tilde{A}/\tilde{B}$ (where $\tilde{B}$ couples the doubly degenerate $\tilde{A}$ states) manifold exist in the MCTDH code and it is known that movement along the torsion angle of the terminal carbons essentially acts as a switch allowing flow of charge to occur between the ends of the molecule. Angles of 75^o and 45^o were used to represent weak and strong coupling, respectively.

An initial study using the full MCTDH method was carried out in order to show the signal of charge migration on static nuclei, the full MCTDH quantum result (quantum dynamic nuclei), the vMCG method (quantum dynamic nuclei) and using a single GWP (classical dynamic nuclei), where the state populations show the movement of charge from one end of the molecule to the other. As expected, the result for the static nuclei showed that the period of oscillation for the charge across the system at 75^o took approximately 16 fs while at 45^o the period of oscillation was approximately 6 fs. Upon inspection of the full MCTDH result, the signal of charge migration can clearly be seen in the first 10 fs of the propagation starting at 45^o with minor features at later times. At 75^o, however, after only 2 fs the charge migration signal is significantly damped. The vMCG method, using only 50 GWPs, replicates the result of the full MCTDH very closely, demonstrating the validity of the vMCG. It is clear that when a single GWP is used the signal of charge migration is damped to a significantly lesser degree and does not replicate the full result well. A comparison between the existing Ehrenfest method (classical trajectories) to the full quantum results shows that while the expectation values along the symmetric stretching mode (Q_3) and H-C-C-H torsional mode (Q_4) are well replicated, due to a bifurcation along the H-C-H out-of-phase bending mode (Q_7) a classical trajectory does not replicate the full result and hence presents a challenge to the new Ehrenfest method.

A hierarchy of methods, representing layers of approximations based on the vMCG method, were then presented detailing the performance of each in com-

parison to the full MCTDH result. The allene test system at the 45^o torsion angle was used and only the Q_3, Q_4 and Q_7 modes were active. As expected, the vMCG method performed well using only 30 GWPs. The first set of approximations made was in the use of classical GWPs (clMCG). These results showed that using 150 standard ($1/\sqrt{2}$) width GWPs replicated the full result well, while the use of 200 narrow (0.4) width GWPs resulted in an over damping of population transfer after the first 6 fs (more significantly after 30 fs) though the signal of charge migration is more significant. The second set of approximations made was in the use of independent (uncoupled) classical GWPs (iMCG). These results showed that using 200 narrow (0.4) width GWPs followed the full result well for the first 6 fs and displayed less damping than the narrow width clMCG after 30 fs. No overall improvement or deterioration with the use of 200 narrow width, equal weighted GWPs was displayed compared to the result where a fitted distribution was used. The final approximation made was in the use of the results of Ehrenfest trajectories (eMCG). Although the results replicated the period of charge migration oscillations in the first 10 fs well, across the entire propagation the damping of the oscillations was poor. It should be noted, however, that the eMCG result may have failed to replicate the clMCG result due to limitations in the current implementation.

The results of a DD-vMCG investigation of the allene cation were then presented. It was found that while the states included in the original allene study were unsuitable for use, the ground ($\tilde{X}(^2\mathrm{E})$) state was a suitable alternative. A 6-31G* CAS(4,3) level of theory for the quantum chemistry calculations was used and, following the Direct Dynamics protocol outlined in Chapter 4.3.1, calculations were propagated using 25 GWPs over 60 fs. The resultant diabatic couplings and potential energy surfaces along selected modes are shown which demonstrate that coupling occurs in particularly in the ν_6 (B_2) mode and that the degenerate-

diabatic/split-adibatic potential surfaces are correct also for the ν_6 mode. Additionally, in the diabatic representation (which is used for propagation in the Dynamics), the surfaces are smooth. However, spurious peaks appear in the diabatic couplings, and the potential energy surfaces representing the ν_3 and ν_4 modes fail to show the correct degenerate behaviour.

As the state population results from these DD-vMCG calculations do not represent the charge migration in the same manner as the earlier study (propagations begin with equilibrium geometry), additional calculations were propagated with the 45^o and 75^o torsion angle starting geometry. In this case it is shown that when a single, time-independent wavepacket is used (analogous to the static nuclei of the full MCTDH calculation) that the period of oscillation between the ends of the molecule occurs over approximately 14 fs at 45^o and takes longer than 60 fs at 75^o, indicating weaker coupling in the $\tilde{X}$ state than in the $\tilde{A}$ state. The full DD-vMCG result shows a similar signal to the earlier results in the first 10 fs of the propagation at both angles after which the results diverge significantly from the full MCTDH result. An additional DD-vMCG calculation was performed using classical GWPs, the performance of which was qualitatively similar to the full DD-vMCG result. This demonstrated a limitation at present in the DD-vMCG method for the representation of these electronic wavepacket calculations.

As a result of the poor performance of the initial Ehrenfest calculations with classical GWPs, a new implementation of Ehrenfest in the Gaussian software package (including second derivatives) was used, the surfaces taken directly from these calculations and the Direct Dynamics performed. The spin densities of coupled DD-vMCG-Ehrenfest calculations using a single nuclear wavepacket and 12 GWPs are shown. The first (single nuclear wavepacket) calculation resulted in the expected oscillation from one end of allene to the other (spin density on the central carbon also shown). The second (12 GWP) calculation failed after 3 fs (due to an

integrator failure) though damping of the charge migration is seen. A third calculation, using 1 GWP following the classical Ehrenfest trajectory, was performed, the resultant state population results of which are shown. No observable damping of the oscillations was achieved and the period of oscillations differed from the single nuclear wavepacket result. It was not possible to definitely state the reason for this, though too large time steps or errors in the Ehrenfest calculation may be the cause and further research is required to further explore this failure.

With the advent of attosecond laser experiments, as well as the identification of systems in which a large number of excited states are required for a complete dynamical representation, the Ehrenfest method presents a potentially useful conceptual alternative to the standard dynamical methods. As shown, the existing Ehrenfest, with classical nuclei, methods under perform even in a relatively simple test system. While the current implementation of the new Ehrenfest method does not yet result in a significant improvement, the failures of the calculation appear to be technical in nature and some potentially promising results are shown (particularly the spin density of the 12 GWP, DD-vMCG-Ehrenfest). The development of this method is continuing.

Chapter 7

Conclusions

In this thesis, new methods for simulating the dynamics of photo-excited molecules including all quantum effects have been tested by applying them to challenging, but realistic systems. The newly developed Direct Dynamics variational Multi-configurational Gaussian method, implemented in the Quantics software package has, thus far, been tested using relatively few active degrees of freedom with a limited number of excited states. The overall aim of this study was essentially to test the limits of this new method using systems of reasonable size, with a larger number of excited states. The DD-vMCG method shows great promise. As a direct dynamics method, it calculates the potential surfaces on-the-fly and so saves the effort of pre-calculating these functions, as is needed in traditional grid-based quantum dynamics methods. This allows complete flexibility for the molecules to go where they want, rather than being restricted by a model potential that can only describe what is in the model. Unlike other direct dynamics methods, however, it includes all the nuclear quantum effects implicitly. These are key for describing situations where the electronic and nuclear dynamics are coupled, such as in non-adiabatic phenomena, and charge migration after the formation of an electronic wavepacket.

Direct dynamics simulations require quantum chemistry calculations to evaluate the potential surfaces and couplings. In order to ensure stability and efficiency

in the propagation a thorough investigation was carried out resulting in a novel method being developed for the systematic characterisation of the active space required in CAS calculations in conjunction with the characterisation of the excited states. This enabled the best selection of active space size, character and number of excited states prior to the Direct Dynamics simulation. Key for future development, the propagation diabatisation scheme, which allows DD-vMCG to be run in the diabatic picture while the surfaces are calculated in the adiabatic picture, is seen to be performing well. The diabatic surfaces are smooth and cross each other, unlike the adiabatic surfaces that form conical intersections instead of crossings. The method is also able to treat a number of states in a general way: here up to eight excited states have been included.

The systems studied, however, show that the calculations take a significant amount of time. The simulation of a 12 atom molecule (2-pyridone) with only 41 basis functions took a number of months to propagate a few femtoseconds. In addition to the time for quantum chemistry calculations, the bulk of this time is taken simply in reading the database of potential surface points. This is something that must be made more efficient for the method to become competitive. The integration of the basis functions also needs improvements. Simulations take small time steps and occasionally fail for no obvious reason. The real limitation of the method, however, is seen to be the accuracy of the electronic structure calculations. Not only are these time-consuming to set up and check, but they may fail in regions of configuration space away from the equilibrium geometry. This is seen strongly in the applications where, for example, the CAS space chosen fails to describe the breaking of the N–H bond.

The Ehrenfest method has been shown to be successful in modelling charge movement after creating an electronic wavepacket when the nuclei are treated classically, but the coupling to the quantum nuclei of the vMCG method still needs

to be completed. First attempts, however, do show that the method describes damping of the charge migration due to the nuclear motion.

The simulations of formamide, formimidic acid, 2-pyridone and 2-hydroxypyridine show some interesting properties. After photo-excitation to high lying states, all relax quickly to lower states. Formimidic acid and formamide both quickly lose a hydrogen atom to aid the loss of energy. All seem fairly photostable to this high energy radiation. In all cases, the chemistry takes place in the S_1 state. Formamide predominantly undergoes loss of N–H but also breaking of the backbone. In contrast, formimidic acid does not fully dissociate, but mostly loses the N–H atom. Interestingly, in formimidic acid, the higher states exclusively lose the O–H atom. Proton transfer is clearly not an important channel for these molecules. Once a hydrogen atom is lost, however, the molecules will be in a position to accept a hydrogen atom from another source to isomerise.

2-pyridone and 2-hydroxypyridine are more photostable than the related formamide and formimidic acid. After exciting to high lying states, fast relaxation takes place with no loss of protons. For both molecules, the chemistry again takes place on S_1 and is predominantly ring breaking. In the case of 2-hydroxypyridine, there is also O–H dissociation which takes place after initial out-of-plane twisting. This is in contrast to formimidic acid in which the analogous O–H dissociation takes place in plane.

In conclusion, the DD-vMCG shows promise of being a general method able to follow excited state dynamics of medium sized molecules accurately. An important feature of the method is that the full potential surfaces are contained in the database, which will provide information on the energetics of regions of configuration space accessible. The next steps with respect to the formamide and 2-pyridone studies is to lift the timing restriction and use a larger, more accurate, basis set. The next steps with respect to the development of the DD-vMCG method will be

to improve the algorithm to allow the calculations to be more efficient. New tools are also required to help analyse the large amount of data obtained. For example, an automatic way of analysing product formation is required as visual inspection is not practical for large basis set sizes and complex molecules.

Bibliography

[1] M. Ben-Nun and T. J. Martínez. Ab initio molecular dynamics study of cistrans photoisomerization in ethylene. *Chem. Phys. Lett.*, 298:57–65, 1998.

[2] N. Luehr, A. G. B. Jin, and T. J. Martínez. Ab Initio Interactive Molecular Dynamics on Graphical Processing Units (GPUs). *J. Chem. Theory Comput.*, 11:4536–4544, 2015.

[3] D. A. Case, D. S. Cerutti, T. E. Cheatham III, T. A. Darden, R. E. Duke, T. J. Giese, H. Gohlke, A. W. Goetz, D. Greene, N. Homeyer, S. Izadi, A. Kovalenko, T. S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. M. Merz, G. Monard, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. R. Roe, A. Roitberg, C. Sagui, C. L. Simmerling, W. M. Botello-Smith, J. Swails, R. C. Walker, J. Wang, R. M. Wolf, X. Wu, L. Xiao, D. M. York, and P. A. Kollman. Amber 2017, 2017.

[4] B. R. Brooks, C. L. Brooks III, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and Karplus M. Charmm: The biomolecular simulation program. *J. Comp. Chem.*, 30:1545–1615, 2009.

[5] S. Jang and G. A. Voth. A derivation of centroid molecular dynamics and other approximate time evolution methods for path integral centroid variables. *J. Chem. Phys.*, 111:2371–2384, 1999.

[6] S Habershon, T E Markland, and D E Manolopoulos. Competing quantum effects in the dynamics of a flexible water model. *J. Chem. Phys.*, 131:24501–24511, 2009.

[7] U. W. Schmitt and G. A. Voth. The computer simulation of proton transport in water. *J. Chem. Phys.*, 111:9361–9381, 1999.

[8] Wikipedia contributors. Moose test – wikipedia, the free encyclopedia, 2017. https://en.wikipedia.org/w/index.php?title=Moose_test&oldid=828171522.

[9] K. K. G. Smith, J. A. Poulsen, G. Nyman, and P. J. Rossky. A new class of ensemble conserving algorithms for approximate quantum dynamics: Theoretical formulation and model problems. *J. Chem. Phys.*, 142:244112–14, 2015.

[10] K. K. G. Smith, J. A. Poulsen, G. Nyman, A. Cunsolo, and P. J. Rossky. Application of a new ensemble conserving quantum dynamics simulation algorithm to liquid para-hydrogen and ortho-deuterium. *J. Chem. Phys.*, 142:244113–10, 2015.

[11] S. Pronk, S. Pll, R. Schulz, P. Larsson, P. Bjelkmar, R. Apostolov, M. R. Shirts, J. C. Smith, P. M. Kasson, D. van der Spoel, B. Hess, and E. Lindahl. Gromacs 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, 29(7):845–854, 2013.

[12] A. de Ruiter and C. Oostenbrink. Extended thermodynamic integration: Efficient prediction of lambda derivatives at nonsimulated points. *J. Chem. Theory Comput.*, 12(9):4476–4486, 2016.

[13] W. Plazinski, A. Lonardi, and P. H. Hnenberger. Revision of the gromos 56a6carbo force field: Improving the description of ring-conformational equilibria in hexopyranose-based carbohydrates chains. *J. Comp. Chem.*, 37(3):354–365, 2016.

[14] M. Barbatti, G. Granucci, M. Ruckenbauer, F. Plasser, J. Pittner, M. Persico, and H. Lischka. Newton-x: a package for newtonian dynamics close to the crossing seam. *www.newtonx.org*, 2012.

[15] M. Barbatti, M. Ruckenbauer, F. Plasser, J. Pittner, G. Granucci, M. Persico, and H. Lischka. Newton-x: a surface-hopping program for nonadiabatic molecular dynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science*, 4(1):26–33, 2014.

[16] C. R. Woese. *The Genetic Code: The Molecular Basis for Genetic Expression*. Modern perspectives in biology. Harper & Row, 1967.

[17] L. E. Orgel. Evolution of Genetic Apparatus. *J. Mol. Biol.*, 38:381–393, 1968.

[18] Crick F. H. C. The origin of the genetic code. *J. Mol. Biol.*, 38:367–379, 1968.

[19] K. Kruger, P. J. Grabowski, A. J. Zaug, J. Sands, D. E. Gottschling, and T. R. Cech. Self-Splicing RNA - Auto-Excision and Auto-Cyclization of the Ribosomal-RNA Intervening Sequence of Tetrahymena. *Cell*, 31:147–157, 1982.

[20] C. Guerriertakada, K. Gardiner, T. Marsh, N. Pace, and S. Altman. The RNA Moiety of Ribonuclease-P is the Catalytic Subunit of the Enzyme. *Cell*, 35:849–857, 1983.

[21] P. Nissen, J. Hansen, N. Ban, P. B. Moore, and T. A. Steitz. The structural basis of ribosome activity in peptide bond synthesis. *Science*, 289(5481):920–930, AUG 11 2000.

[22] K. Lang, M. Erlacher, D. N. Wilson, R. Micura, and N. Polacek. The role of 23S ribosomal RNA residue A2451 in peptide bond synthesis revealed by atomic mutagenesis. *Chem. Biol.*, 15(5):485 – 492, 2008.

[23] I. R. Craig and D. E. Manolopoulos. Quantum statistics and classical mechanics: real time correlation functions from ring polymer molecular dynamics. *J. Chem. Phys.*, 121:3368–3373, 2004.

[24] S. Habershon, D. E. Manolopoulos, T. E. Markland, and T. F. Miller. Ring-Polymer Molecular Dynamics: Quantum Effects in Chemical Dynamics from Classical Trajectories in an Extended Phase Space. *Ann. Rev. Phys. Chem.*, 64:387–413, 2013.

[25] J. O. Richardson and S. C. Althorpe. Ring-polymer molecular dynamics rate-theory in the deep-tunneling regime: Connection with semiclassical instanton theory. *J. Chem. Phys.*, 131:214106–12, 2009.

[26] S. Mai, M. Richter, M. Heindl, M. F. S. J. Menger, A. Atkins, M. Ruckenbauer, F. Plasser, M. Oppel, P. Marquetand, and L. González. Sharc2.0: Surface hopping including arbitrary couplings program package for non-adiabatic dynamics. sharc-md.org, 2018.

[27] M. Richter, P. Marquetand, J. González-Vázquez, I. Sola, and González L. SHARC: ab initio molecular dynamics with surface hopping in the adiabatic

representation including arbitrary couplings. *J. Chem. Theory Comput.*, 7(5):1253–1258, 2011.

[28] S. Mai, P. Marquetand, and L. González. A general method to describe intersystem crossing dynamics in trajectory surface hopping. *Int. J. Quantum Chem.*, 115:1215–1231, 2015.

[29] M. Messineo, J. S. Clark, D. F. Figer, R-P. Kudritzki, F. Najarro, R. M. Rich, K. M. Menten, V. D. Ivanov, E. Valenti, C. Trombley, C-H. R. Chen, and B. Davies. Massive Stars in the W33 Giant Molecular Complex. *Astrophys. J.*, 805:110, 2015.

[30] A. H. Zewail. Femtochemistry. Past, present, and future. *Pure Appl. Chem.*, 72:2219–2231, 2000.

[31] A. H. Zewail. Laser Femtochemistry. *Science.*, 242:1645–1653, 1988.

[32] Douhal A. and Lahmani F. and Zewail A. H. Proton-transfer reaction dynamics. *Chemical Physics*, 207:477–498, 1996. Laser and molecular beam studies of chemical reaction dynamics.

[33] E. Stoye. Femtochemistry pioneer Ahmed Zewail dies. *Chemistry World*, 2016.

[34] R. M. Bowman and A. H. Zewail. From strongly bound to repulsive surface dynamics. *Chem. Phys. Lett.*, 161:297–302, 1989.

[35] M. Dantos, M. J. Rosker, and A. H. Zewail. Femtosecond real-time probing of reactions. II. The dissociation reaction of ICN. *J. Chem. Phys.*, 89:6128–6140, 1988.

[36] T. S. Rose, M. J. Rosker, and A. H. Zewail. Femtosecond real-time probing of reactions. IV. The reactions of alkali halides. *J. Chem. Phys.*, 91:7415–7436, 1989.

[37] S. Ullrich, T. Schultz, M. Z. Zgierski, and A. Stolow. Electronic relaxation dynamics in DNA and RNA bases studied by time-resolved photoelectron spectroscopy. *Phys. Chem. Chem. Phys.*, 6:2796–2801, 2004.

[38] P. Kukura, D. W. McCamant, S. Yoon, D. B. Wandschneider, and R. A. Mathies. Structural observation of the primary isomerization in vision with femtosecond-stimulated Raman. *Science*, 310:1006–1009, 2005.

[39] M. Kondo, I. A. Heisler, D. Stoner-Ma, P. J. Tonge, and S. R. Meech. Ultrafast proton transfer in the green fluorescent protein: Analysing the instantaneous emission at product state wavelengths. *J. Photochem. Photobiol. A*, 234:21–26, 2012.

[40] J. J. van Thor, G. Y. Georgiev, M. Towrie, and J. T. Sage. Ultrafast and low barrier motions in the photoreactions of the green fluorescent protein. *J. Biol. Chem.*, 280:33652–33659, 2005.

[41] F. Krausz and M. Ivanov. Attosecond physics. *Rev. Mod. Phys.*, 81:163–234, 2009.

[42] M. F. Kling and M. J. J. Vrakking. Attosecond Electron Dynamics. *Ann. Rev. Phys. Chem.*, 59:463–492, 2008.

[43] G. W. Richings, I. Polyak, K. E. Spinlove, G. A. Worth, I. Burghardt, and B. Lasorne. Quantum dynamics simulations using Gaussian wavepackets: the vMCG method. *Int. Rev. Phys. Chem.*, 34:269–308, 2015.

[44] K. E. Spinlove, M. Vacher, M. Bearpark, M. A. Robb, and G. A. Worth. Using quantum dynamics simulations to follow the competition between charge

migration and charge transfer in polyatomic molecules. *Chem. Phys.*, 482:52–63, 2016.

[45] G. A. Worth, K. Giri, G. W. Richings, M. H. Beck, A. Jäckle, and H.-D. Meyer. Quantics package, Version 1.1, 2015.

[46] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, and J. V. Ort. Gaussian 09, Revision A.02, 2009.

[47] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.*, 28:1049–1070, 1926.

[48] Hinchliffe. *Molecular Modelling for Beginners.* John Wiley and sons, New York, 2011.

[49] A. Szabo and N. S. Ostlund. *Modern Quantum Chemistry.* McGraw Hill, New York, 1989.

[50] H Köppel, W Domcke, and L S Cederbaum. Multimode molecular Dynamics beyond the Born-Oppenheimer approximation. *Adv. Chem. Phys.*, 57:59–246, 1984.

[51] G. A. Worth and L. S. Cederbaum. Beyond Born-Oppenheimer: Conical intersections and their impact on molecular dynamics. *Ann. Rev. Phys. Chem.*, 55:127–158, 2004.

[52] M Baer. Topological effects in molecular systems : an attempt towards a complete theory. *Chem. Phys.*, 259:123–147, 2000.

[53] M Baer. Adiabatic and diabatic representations for atom-molecule collisions: Treatment of the collinear arrangement. *Chem. Phys. Lett*, 35:112–118, 1975.

[54] G. N. Lewis. The atom and the molecule. *J. Am. Chem. Soc.*, 1448:762–785, 1916.

[55] W. Heitler and F. London. Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik. *Z. Phys.*, 44:455–472, 1927.

[56] L. Pauling. The Shared-Electron Chemical Bond. *Proc. Natl. Acad. Sci. USA*, 14:359–362, 1928.

[57] L. Pauling. The Nature of the Chemical Bond. Application of Results Obtained from the Quantum Mechanics and from a Theory of Paramagnetic Susceptibility to the Structure of Molecules. *J. Am. Chem. Soc.*, 53:1367–1400, 1931.

[58] S. Shaik and P. C. Hiberty. *A Chemist's Guide to Valence Bond Theory*. John Wiley and Sons, Hoboken, NJ, USA, 2008.

[59] Nobel Foundation. *Nobel Lectures, Chemistry 1963-1970*. Elsevier Publishing Company, Amsterdam, 1972.

[60] J. E. Lennard-Jones. The electronic structure of some diatomic molecules. *Trans. Farad. Soc.*, 25:668–686, 1929.

[61] E. Hückel. Zur Quantentheorie der Doppelbindung. *Z. Phys.*, 60:423–465, 1930.

[62] E. Hückel. Quantentheoretische Beiträge zum Benzolproblem - I. Die Elektronenkonfiguration des Benzols und verwandter Verbindungen. *Z. Phys.*, 70:204–286, 1931.

[63] E. Hückel. Grundzüge der Theorie Ungesättigter und Aromatischer Verbindungen. *Z. Elektrochem.*, 43:752–788, 1937.

[64] E. Hückel. Grundzüge der Theorie Ungesättigter und Aromatischer Verbindungen. *Z. Elektrochem.*, 43:827–849, 1937.

[65] R. S. Mulliken. Electronic Structures of Polyatomic Molecules and Valence VI. On the Method of Molecular Orbitals. *J. Chem. Phys.*, 3:375–378, 1935.

[66] R. S. Mulliken. Electronic Population Analysis on LCAOMO Molecular Wave Functions. I. *J. Chem. Phys.*, 23:1833–1840, 1955.

[67] Nobelprize.org. The Nobel Prize in Chemistry 1966. http://www.nobelprize.org/nobel_prizes/chemistry/laureates/1966/.

[68] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev. B*, 136:864–871, 1964.

[69] L. H. Thomas. The calculation of atomic fields. *Math. Proc. Cam. Phil. Soc.*, 23:542–548, 1927.

[70] E. Fermi. Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente. *Z. Phys.*, 48:73–79, 1928.

[71] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev. A*, 140:1133–1138, 1965.

[72] S. S. Shaik and P. C. Hiberty. *A Chemist's Guide to Valence Bond Theory.* Wiley, 2007.

[73] J. B. Li and R McWeeny. VB2000: Pushing valence bond theory to new limits. *Int. J. Q. Chem.*, 89:208–216, 2002.

[74] S. Shaik and P. C. Hiberty. Valence bond theory, its history, fundamentals, and applications: A primer. In Lipkowitz, K. B. and Larter, R. and Cundari, T. R., editor, *Rev. Comp. Chem.*, volume 20 of *Reviews in Computational Chemistry*, pages 1–100. Wiley-VCH, Inc., 2004.

[75] N. Bohr. Der Bau der Atome und die physikalischen und chemischen Eigenschaften der Elemente. *Z. Phys.*, 9:1–67, 1922.

[76] D. R. Hartree. The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods. *Math. Proc. Cam. Phil. Soc.*, 24:89–110, 1928.

[77] D. R. Hartree. The Wave Mechanics of an Atom with a non-Coulomb Central Field. Part II. Some Results and Discussion. *Math. Proc. Cam. Phil. Soc.*, 24:111–132, 1928.

[78] J.C. Slater. *Quantum Theory of Atomic Structure*. International series in pure and applied physics. McGraw-Hill, 1960.

[79] W. Pauli. On the connection of the arrangement of electron groups in atoms with the complex structure of spectra. *Z. Phys.*, 31:765–783, 1925.

[80] W. Pauli. The Connection Between Spin and Statistics. *Phys. Rev.*, 58:716–722, 1940.

[81] J. C. Slater. Note on Hartree's method. *Phys. Rev.*, 35:0210–0211, 1930.

[82] V. Fock. Approximation method for the solution of the quantum mechanical multibody problems. *Z. Phys.*, 61:126–148, 1930.

[83] D. R. Hartree and W. Hartree. Self-Consistent Field, with Exchange, for Beryllium. *Proc. R. Soc. London A Math. Phys. Sci.*, 150:9–33, 1935.

[84] C. C. J. Roothaan. New Developments in Molecular Orbital Theory. *Rev. Mod. Phys.*, 23:69–89, 1951.

[85] G. G. Hall. The molecular orbital theory of chemical valency viii. a method of calculating ionization potentials. *Proc. R. Soc. A*, 205:541–552, 1951.

[86] F. Jensen. *Introduction to Computational Chemistry.* John Wiley & Sons, 2006.

[87] C. J. Cramer. *Essentials of Computational Chemistry: Theories and Models.* John Wiley & Sons, 2004.

[88] S. F. Boys. Electronic wave functions. I. A general method of calculation for the stationary states of any molecular system. *Proc. R. Soc. A*, 200:542–554, 1950.

[89] F. Jensen. *Introduction to Computational Chemistry. Chapter 5.* John Wiley & Sons, 2006.

[90] E. Schrödinger. Quantisierung als Eigenwertproblem. *Ann. Phys.*, 385:437–490, 1926.

[91] K. Anderson, P.-Å. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. Second-Order Perturbation Theory with a CASSCF Reference Function. *J. Phys. Chem.*, 94:5483–5488, 1990.

[92] O. Sinanoğlu. Many-Electron Theory of Atoms, Molecules and their Interactions. *Adv. Chem. Phys.*, 6:315–412, 1964.

[93] R. P. Feynman. Space-Time Approach to Non-Relativistic Quantum Mechanics. *Rev. Mod. Phys.*, 20:367, 1948.

[94] A. Abedi, N. T. Maitra, and E. K. U. Gross. Exact Factorization of the Time-Dependent Electron-Nuclear Wave Function. *Phys. Rev. Lett.*, 105:123002, 2010.

[95] A. Abedi, N. T. Maitra, and E. K. U. Gross. Correlated electron-nuclear dynamics: Exact factorization of the molecular wavefunction. *J. Chem. Phys.*, 137:22A530, 2012.

[96] S. K. Min, F. Agostini, and E. K.U. Gross. Coupled-Trajectory Quantum-Classical Approach to Electronic Decoherence in Nonadiabatic Processes. *Phys. Rev. Lett.*, 115:073001–5, 2015.

[97] S. K. Min, F. Agostini, I. Tavernelli, and E. K.U. Gross. Ab Initio Nonadiabatic Dynamics with Coupled Trajectories: A Rigorous Approach to Quantum (De)Coherence. *J. Phys. Chem. Lett.*, 8:3048–3055, 2017.

[98] A. D. McLachlan. A variational solution of the time-dependent Schrödinger equation. *Mol. Phys.*, 8:39, 1964.

[99] P. A. M. Dirac. Note on exchange phenomena in the Thomas atom. *Proc. Cambridge Philos. Soc.*, 26:376–385, 1930.

[100] H.-D. Meyer, U. Manthe, and L. S. Cederbaum. The Multi-Configurational Time-Dependent Hartree Approach. *Chem. Phys. Lett.*, 165:73–78, 1990.

[101] J. Frenkel. *Wave Mechanics*. Clarendon Press, Oxford, U.K., 1934.

[102] M. H. Beck, A. Jäckle, G. A. Worth, and H.-D. Meyer. The multiconfiguration time-dependent Hartree method: A highly efficient algorithm for propagating wavepackets. *Phys. Rep.*, 324:1–105, 2000.

[103] J. C. Light, I. P. Hamilton, and J. V. Lill. Generalized discrete variable approximation in quantum mechanics. *J. Chem. Phys.*, 82:1400–1409, 1985.

[104] J. C. Light and T. Carrington Jr. Discrete variable representations and their utilization. *Adv. Chem. Phys.*, 114:263–310, 2000.

[105] H.-D. Meyer and G. A. Worth. Quantum molecular dynamics: propagating wavepackets and density operators using the multiconfiguration time-dependent Hartree method. *Theo. Chem. Acc.*, 109:251–267, 2003.

[106] U. Manthe, H.-D. Meyer, and L. S. Cederbaum. Wave-Packet Dynamics within the Multiconfiguration Hartree Framework: General Aspects and application to NOCl. *J. Chem. Phys*, 97:3199–3213, 1992.

[107] M. H. Beck and H.-D. Meyer. An efficient and robust integration scheme for the equations of motion of the multiconfiguration time-dependent Hartree (MCTDH) method. *Z. Phys. D*, 42:113–129, 1997.

[108] G. A. Worth, H. D. Meyer, and L. S. Cederbaum. Relaxation of a system with a conical intersection coupled to a bath: A benchmark 24-dimensional wave packet study treating the environment explicitly. *J. Chem. Phys.*, 109:3518–3529, 1998.

[109] I. Burghardt, H.-D. Meyer, and L. S. Cederbaum. Approaches to the approximate treatment of complex molecular systems by the multiconfiguration time-dependent Hartree method. *J. Chem. Phys.*, 111:2927–2938, 1999.

[110] I. Burghardt, M. Nest, and G. A. Worth. Multiconfigurational system-bath dynamics using gaussian wavepackets: Energy relaxation and decoherence induced by a finite-dimensional bath. *J. Chem. Phys.*, 119:5364–5378, 2003.

[111] G. A. Worth and I. Burghardt. Full quantum mechanical molecular dynamics using Gaussian wavepackets. *Chem. Phys. Lett.*, 368:502–508, 2003.

[112] I. Burghardt, K. Giri, and G. A. Worth. Multimode quantum dynamics using Gaussian wavepackets: The Gaussian-based multiconfiguration time-dependent Hartree (G-MCTDH) method applied to the absorption spectrum of pyrazine. *J. Chem. Phys.*, 129:174104–174114, 2008.

[113] E. J. Heller. Time-dependent approach to semiclassical dynamics. *J. Chem. Phys.*, 62:1544–1555, 1975.

[114] E. J. Heller. Frozen Gaussians: A very simple semiclassical approximation. *J. Chem. Phys.*, 75:2923–2931, 1981.

[115] S. Sawada, R. Heather, B. Jackson, and H. Methiu. A strategy for time dependent quantum mechanical calculations using a Gaussian wave packet representation of the wave function. *J. Chem. Phys.*, 83:3009–3027, 1985.

[116] G. A. Worth, M. A. Robb, and I. Burghardt. A novel algorithm for non-adiabatic direct dynamics using variational Gaussian wavepackets. *Faraday Discuss.*, 127:307–323, 2004.

[117] B. Lasorne, M. A. Robb, and G. A. Worth. Direct quantum dynamics using variational multi-configuration Gaussian wavepackets. Implementation details and test case. *Phys. Chem. Chem. Phys.*, 9:3210–3227, 2007.

[118] G. A. Worth, M. A. Robb, and B. L. Lasorne. Solving the time-dependent Schrödinger equation for nuclear motion in one step: Direct dynamics of non-adiabatic systems. *Mol. Phys.*, 106:2077–2091, 2008.

[119] B. Lasorne, M. J. Bearpark, M. A. Robb, and G. A. Worth. Direct quantum dynamics using variational multi-configuration Gaussian wavepackets. *PCCP*, 432:3210–3227, 2006.

[120] B. Lasorne, M. J. Bearpark, M. A. Robb, and G. A. Worth. Controlling S_1/S_0 Decay and the Balance between Photochemistry and Photostability in

Benzene: A Direct Quantum Dynamics Study. *J. Phys. Chem. A*, 112:13017–13027, 2008.

[121] M. Araújo, B. Lasorne, A. L. Magalhães, G. A. Worth, M. J. Bearpark, and M. A. Robb. The Molecular Dissociation of Formaldehyde at Medium Photoexcitation Energies: A Quantum Chemistry and Direct Quantum Dynamics Study. *J. Chem. Phys.*, 131:144301–144308, 2009.

[122] C. S. M. Allan, B. Lasorne, G. A. Worth, and M. A. Robb. A straightforward method of analysis for direct quantum dynamics: application to the photochemistry of a model cyanine. *J. Phys. Chem. A*, 114:8713–8729, 2010.

[123] D. Asturiol, B. Lasorne, G. A. Worth, M. A. Robb, and L. Blancafort. Exploring the sloped-to-peaked S_2/S_1 seam of intersection of thymine with electronic structure and direct quantum dynamics calculations. *PCCP*, 12:4949–4958, 2010.

[124] D. Mendive-Tapia, B. Lasorne, G. A. Worth, M. A. Robb, and M. J. Bearpark. Towards converging non-adiabatic direct dynamics calculations using frozen-width variational Gaussian product basis functions. *J. Chem. Phys.*, 548:22A548–10, 2012.

[125] J. Ischtwan and M. A. Collins. Molecular potential energy surfaces by interpolation. *J. Chem. Phys.*, 100:8080–8088, 1994.

[126] M. A. Collins and D. H. Zhang. Application of interpolated potential energy surfaces to quantum reactive scattering. *J. Chem. Phys.*, 111:9924–9931, 1999.

[127] T. J. Frankcombe, M. A. Collins, and G. A. Worth. Converged quantum dynamics with modied Shepard interpolation and Gaussian wave packets. *Chem. Phys. Lett.*, 489:242–247, 2010.

[128] T. J. Frankcombe. Using Hessian update formulae to construct modified Shepard interpolated potential energy surfaces: application to vibrating surface atoms. *J. Chem. Phys.*, 140:114106–114108, 2014.

[129] B. Esry and H. Sadeghpour. Split diabatic representation. *Phys. Rev. A*, 68:42706, 2003.

[130] G. W. Richings and G. A. Worth. A Practical Diabatisation Scheme for Use with the Direct-Dynamics Variational Multi-Configuration Gaussian Method. *J. Phys. Chem. A*, 119:12457–12470, 2015.

[131] A. W. Hofmann. On Formamide. *J. Am. Chem. Soc.*, 16:72–74, 1863.

[132] K. Smeykal and M Köthing. VEB Leuna-Werke. DD 13619, 1954.

[133] G. Peinze. VEB Leuna-Werke. DD 61536, 1967.

[134] E. Germann and H. Fritz. Verfahren zur herstellung von formamid, April 28 1966. DE 1215130B.

[135] H. Bipp and H. Kieczka. Formamides. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

[136] H. Hohenschutz, M. Strohmeyer, M. Herr, and H. Kiefer. Continuous manufacture of formamide, February 19 1980. CA 1072118/DE 2623173.

[137] C. E. A. Winslow, J. Broadhurst, R. E. Buchanan, C. Krumwiede, L. A. Rogers, and G. H. Smith. The Families and Genera of the Bacteria: Final Report of the Committee of the Society of American Bacteriologists on Characterization and Classification of Bacterial Types. *J. Bacteriol.*, 5:191–229, 1920.

[138] A. T. Fuller. The Formamide Method for the Extraction of Polysaccharides from Hæmolytic Streptococci. *Br. J. Exp. Pathol.*, 19:130–139, 1938.

[139] Copeland H. F. The Kingdoms of Organisms. *The Quarterly Review of Biology*, 13:383–420, 1938.

[140] R. H. Whittaker. New Concepts of Kingdoms of Organisms. *Science*, 163:150–160, 1969.

[141] S. Hamada and H. D. Slade. Biology, Immunology and Cariogenicity of Streptococcus mutans. *Microbiol. Mol. Biol. Rev.*, 44:331–384, 1980.

[142] P. Longerich. *Holocaust: The Nazi Persecution and Murder of Jews*. Oxford University Press, 2010.

[143] W. Heerdt. Verfahren zur schaedlingsbekaempfung, December 27 1926. DE 438818.

[144] L. Andrussow. Über die katalytische oxydation von ammoniak-methan-gemischen zu blausure. *Angewandte Chemie*, 48:593–595, 1935.

[145] E. Gail, S. Gos, R. Kulzer, J. Lorsch, A. Rubo, M. Sauer, R. Kellens, J. Reddy, N. Steier, and W. Hasenpusch. Cyano compounds, inorganic. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

[146] L. Andrussow. Über die schnell verlaufenden katalytischen prozesse in strmenden gasen und die ammoniak-oxydation (v). *Berichte der deutschen chemischen Gesellschaft (A and B Series)*, 60:2005–2018, 1927.

[147] L. Andrussow. Production of hydrocyanic acid, November 14 1933. US 1934838.

[148] J. M. Pirie. The Manufacture of Hydrocyanic Acid by the Andrussow Process. *Plat. Met. Rev.*, 2:7–11, 1958.

[149] unknown. Procédé pour retirer des substances volatiles d'enceintes ou règne une pression réduite / werkwijze voor het winnen van vluchtige stoffen uit onder verminderde druk staande ruimten, July 27 1944. FR 906114(A)/NL 63462(C).

[150] J. McKillop. Combat chronology of the USAAF, 2006.

[151] D. L. Miller. *Masters of the Air: America's Bomber Boys Who Fought the Air War Against Nazi Germany*. Simon & Schuster: New York, 2006.

[152] O. W. Holloway and O. B. Isaac. Classified list of ots reports, 1947. Technical Reports Section, Science and Technology Division, Library of Congress.

[153] P. L. Magill. Formamide. *Ind. Eng. Chem.*, 26:611–614, 1934.

[154] A. Höhn and Updated by Staff. Formamide. In *Kirk-Othmer Encyclopedia of Chemical Technology*. John Wiley & Sons, Inc., 2000.

[155] T. Lebleu, H. Kotsuki, J. Maddaluno, and J. Legros. Formylation of amines through catalyst- and solvent-free transamidation reaction. *Tetrahedron Lett.*, 55:362–364, 2014.

[156] M. N. Jachak, A. B. Avhale, V. J. Medhane, and R. B. Toche. A convenient route for the synthesis of pyrazolo[3,4-d]pyrimidine, pyrazolo[3,4-b][1,6]naphthyridine and pyrazolo[3,4-b]quinoline derivatives. *J. Heterocyclic Chem.*, 43:1169–1175, 2006.

[157] I. Nouira, I. K. Kostakis, C. Dubouilh, E. Chosson, M. Iannelli, and T. Besson. Decomposition of formamide assisted by microwaves, a tool for synthesis of nitrogen-containing heterocycles. *Tetrahedron Lett.*, 49:7033–7036, 2008.

[158] F. Cataldo, G. Patanè, and G. Compagnini. Synthesis of HCN Polymer from Thermal Decomposition of Formamide. *J. Macromol. Sci. A*, 46:1039–1048, 2009.

[159] P. M. Kosaka, Y. Kawano, and D. F. S. Petri. Dewetting and surface properties of ultrathin films of cellulose esters. *J. Colloid Interface Sci.*, 316:671–677, 2007.

[160] Z. P. Liu, R. Z. Ma, M. Osada, N. Iyi, Y. Ebina, K. Takada, and T. Sasaki. Synthesis, anion exchange, and delamination of Co-Al layered double hydroxide: Assembly of the exfoliated nanosheet/polyanion composite films and magneto-optical studies. *J. Am. Chem. Soc.*, 128:4872–4880, 2006.

[161] S. G. Fleischman, S. S. Kuduva, J. A. McMahon, B. Moulton, R. D. B. Walsh, N. Rodriguez-Hornedo, and M. J. Zaworotko. Crystal engineering of the composition of pharmaceutical phases: Multiple-component crystalline solids involving carbamazepine. *Crystal Growth & Design*, 3:909–919, 2003.

[162] M. A. Farajzadeh, L. Goushjuii, D. Djozan, and J. K. Mohammadi. Dispersive liquid-liquid microextraction combined with gas chromatography for extraction and determination of class 1 residual solvents in pharmaceuticals. *J. Sep. Sci.*, 35:1027–1035, 2012.

[163] A. G. Diress and C. A. Lucy. Self-assembled coating for modification of the electro-osmotic flow in nonaqueous capillary electrophoresis using formamide. *Can. J. Chem.*, 85:540–546, 2007.

[164] P. M. Solomon. Interstellar Molecules. *Physics Today*, 26:32–40, 1973.

[165] D. Bockelée-Morvan, D. C. Lis, J. E. Wink, D. Despois, J. Crovisier, R. Bachiller, D. J. Benford, N. Biver, P. Colom, J. K. Davies, E. Gérard, B. Germain, M. Houde, D. Mehringer, R. Moreno, G. Paubert, T. G. Phillips,

and H. Rauer. New molecules found in comet C / 1995 O1 (Hale-Bopp). Investigating the link between cometary and interstellar material. *Astron. Astrophys.*, 353:1101–1114, 2000.

[166] D. C. Lis, D. M. Mehringer, D. Benford, M. Gardner, T. G. Phillips, D. Bockelée-Morvan, N. Biver, P. Colom, J. Crovisier, D. Despois, and H. Rauer. New Molecular Species in Comet C/1995 O1 (Hale-Bopp) Observed with the Caltech Ssubmillimeter Observatory. *Earth, Moon, and Planets*, 78:13–20, 1997.

[167] W. A. Schutte, A. C. A. Boogert, A. G. G. M. Tielens, D. C. B. Whittet, P. A. Gerakines, J. E. Chiar, P. Ehrenfreund, J. M. Greenberg, E. F. van Dishoeck, and T. de Graauw. Weak ice absorption features at 7.24 and 7.41 mu m in the spectrum of the obscured young stellar object W 33A. *Astron. Astroph.*, 343:966–976, 1999.

[168] W. H. Flygare, R. C. Benson, H. L. Tigelaar, R. H. Rubin, and G. W. Jr. Swenson. Chemical and rotational stability of interstellar formamide. In *Molecules in the galactic environment*. John Wiley and Sons Inc; New York, 1973.

[169] C. A. Gottlieb, P. Palmer, L. J. Rickard, and B. Zuckerman. Studies of Interstellar Formamide. *Astroph. J.*, 182:699–710, 1973.

[170] R. Saladino, C. Crestini, S. Pino, G. Costanzo, and E. Di Mauro. Formamide and the origin of life. *Phys. Life Rev.*, 9:84–104, 2012.

[171] B. M. Jones, C. J. Bennett, and R. I. Kaiser. Mechanistical Studies on the Production of Formamide (H_2NCHO) within Interstellar Ice Analogs. *Astrophy. J.*, 734:78, 2011.

[172] V. S. Nguyen, T. M. Orlando, J. Leszczynski, and M. T. Nguyen. Theoretical Study of the Decomposition of Formamide in the Presence of Water Molecules. *J. Phys. Chem. A*, 117:2543–2555, 2013.

[173] V. S. Nguyen, H. L. Abbott, M. M. Dawley, T. M. Orlando, J. Leszczynski, and M. T. Nguyen. Theoretical study of formamide decomposition pathways. *J. Phys. Chem. A*, 115:841–851, 2011.

[174] D. Guzmán-Angel, R. Inostroza-Rivera, S. Gutiérrez-Oliva, B. Herrera, and A. Toro-Labbé. Role of water in intramolecular proton transfer reactions of formamide and thioformamide. *Theor. Chem. Acc.*, 135:37–10, 2016.

[175] X. C. Wang, J. Nichols, M. Feyereisen, M. Gutowski, J. Boatz, A. D. J. Haymet, and J. Simons. Ab initio quantum chemistry study of formamide-formamidic acid tautomerization. *J. Phys. Chem.*, 95:10419–10424, 1991.

[176] N. V. Tukachev, V. A. Bataev, A. V. Abramenkov, and I. A. Godunov. Structure and conformational dynamics of formamide molecule in the ground and lowest excited electronic states. *Comp. Theor. Chem.*, 1080:23–32, 2016.

[177] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, Y. Liu, A. W. Lloyd, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, and M. Wang. MOLPRO, version 2015.1, a package of ab initio programs, 2015.

[178] P. J. Knowles and H.-J. Werner. An Efficient Second Order MCSCF Method for Long Configuration Expansions. *Chem. Phys. Lett.*, 115:259–267, 1985.

[179] R. Brause, M. Schmitt, and K. Kleinermanns. Improved determination of structural changes of 2-pyridone-$(H_2O)_1$ upon electronic excitation. *J, Phys. Chem. A*, 111:3287–3293, 2007.

[180] S. Mata, V. Cortijo, W. Caminati, J. L. Alonso, M. E. Sanz, J. C. López, and S. Blanco. Tautomerism and Microsolvation in 2-Hydroxypyridine / 2-Pyridone. *J. Phys. Chem. A*, 114:11393–11398, 2010.

[181] H. Beg, D. Das, S. Ash, and A. Misra. Computation of polarizability, hyperpolarizability and hardness as descriptor for enol-keto tautomerizations of 2-hydroxy pyridines. *Comp. Theor. Chem.*, 1017:200–207, 2013.

[182] A. Samanta, B. K. Paul, S. Kar, and N. Guchhait. Excited state lactim to lactam type tautomerization reaction in 5-(4-fluorophenyl)-2-hydroxypyridine: Spectroscopic study and quantum chemical calculation. *J. Fluoresc.*, 21:95–104, 2011.

[183] M. J. Scanlan and I. H. Hillier. On the mechanism of proton transfer in the 2-hydroxypyridine $\leftrightarrows$ 2-pyridone tautomeric equilibrium. *Chem. Phys. Lett.*, 107:330–332, 1984.

[184] R Nimlos, D F Kelley, and E R Bernstein. Spectroscopy, Structure, and Proton Dynamics of 2-Hydroxypyridine and Its Clusters with Water and Ammonia. *J. Phys. Chem.*, 93:643–651, 1989.

[185] K. Axel. Antiinfectives for Systemic Use, 1. Antibacterials. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

[186] J. Ritz, H. Fuchs, H. Kieczka, and W. C. Moran. Caprolactam. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

[187] H. Berneth. Methine dyes and pigments. In *Ullmann's Encyclopedia of Industrial Chemistry*. Wiley-VCH Verlag GmbH & Co. KGaA, 2000.

[188] Kränzlein G., M. Heyse, and P. Ochwat, 1926.

[189] B. Chmura, M. F. Rode, A. L. Sobolewski, L. Lapinski, and M. J. Nowak. A Computational Study on the Mechanism of Intramolecular Oxo-Hydroxy Phototautomerism Driven by Repulsive $\pi\sigma^*$ State. *J. Phys. Chem. A*, 112:13655–13661, 2008.

[190] C. S. Peng and A. Tokmakoff. Identification of Lactam Lactim Tautomers of Aromatic Heterocycles in Aqueous Solution Using 2D IR Spectroscopy. *J. Phys. Chem. Lett.*, 3:3302–3306, 2012.

[191] J. C. Poully, J. P. Schermann, N. Nieuwjaer, F. Lecomte, G. GréGoire, C. Desfrancois, G. A. Garcia, L. Nahon, D. Nandi, L. Poisson, and M. Hochlaf. Photoionization of 2-pyridone and 2-hydroxypyridine. *PCCP*, 12:3566–3572, 2010.

[192] D. Hammoutène, M. Hochlaf, I. Ciofini, and C. Adamo. Electronic Spectrum of 2-Pyridone$^+$: Ab Initio and Time-Dependent Density Functional Calculations. *Int. J. Quant. Chem.*, 110:498–504, 2010.

[193] J. A. Frey, R. Leist, C. Tanner, H. M. Frey, and S. Leutwyler. 2-pyridone: The role of out-of-plane vibrations on the $S_1 \leftrightarrow S_0$ spectra and S_1 state reactivity. *J. Chem. Phys.*, 125:114308–14, 2006.

[194] Q.-S. Li, W.-H. Fang, and J.-G. Yu. Theoretical studies of proton-transfer reactions of 2-hydroxypyridine–$(H_2O)_n(n = 0 - 2)$ in the ground and excited states. *J. Phys. Chem. A*, 109:3983–3990, 2005.

[195] A. L. Sobolewski. The mechanism of excited-state hydrogen transfer in 2-hydroxypyridine. *Chem. Phys. Lett*, 211:293–299, 1993.

[196] L. Poisson, D. Nandi, B. Soep, M. Hochlaf, M. Boggio-Pasqua, and J.-M. Mestdagh. A roaming wavepacket in the dynamics of electronically excited 2-hydroxypyridine. *PCCP*, 16:581–587, 2014.

[197] S. Kopec and H. Köppel. Theoretical analysis of the $S_2 \leftarrow S_0$ vibronic spectrum of the 2-pyridone dimer. *J. Chem. Phys.*, 144, 2016.

[198] A. L. Kuleff, S. Lünnemann, and L. S. Cederbaum. Electron-correlation-driven charge migration in oligopeptides. *Chem. Phys.*, 414:100–105, 2013.

[199] A. L. Kuleff and L. S. Cederbaum. Ultrafast correlation-driven electron dynamics. *J. Phys. B*, 47:124002, 2014.

[200] D. Mendive-Tapia, M. Vacher, M. J. Bearpark, and M. A. Robb. Coupled electron-nuclear dynamics: charge migration and charge transfer initiated near a conical intersection. *J. Chem. Phys.*, 139:44110, 2013.

[201] M. Vacher, M. J. Bearpark, and M. A. Robb. Communication: Oscillating charge migration between lone pairs persists without significant interaction with nuclear motion in the glycine and Gly-Gly-NH-CH_3 radical cations. *J. Chem. Phys.*, 140:201102, 2014.

[202] M. Vacher, J. Meisner, D. Mendive-Tapia, M. J. Bearpark, and M. A. Robb. Electronic Control of Initial Nuclear Dynamics Adjacent to a Conical Intersection. *J. Phys. Chem. A*, 119:5165–5172, 2014.

[203] D. V. Shalashilin. Quantum mechanics with the basis set guided by Ehrenfest trajectories: theory and application to spin-boson model. *J. Chem. Phys.*, 130:244101–244111, 2009.

[204] D. V. Shalashilin. Multiconfigurational Ehrenfest approach to quantum coherent dynamics in large molecular systems. *Farad. Discuss.*, 153:105–116, 2011.

[205] M. Vacher, D. Mendive-Tapia, M. J. Bearpark, and M. A. Robb. The second-order Ehrenfest method. *Theo. Chem. Acc.*, 133:1505–1512, 2014.

[206] M. Amarouche, F. X. Gadea, and J. Durup. A proposal for the theoretical treatment of multi-electronic-state molecular dynamics with the whole DIM basis (HWD). A test on the evolution of excited Ar_3^+ cluster ions. *Chem. Phys.*, 130:145–157, 1989.

[207] M. F. Hermann. Dynamics by semi-classical methods. *Ann. Rev. Phys. Chem.*, 45:83–111, 1994.

[208] A. W. Jasper, S. Nangia, C. Zhu, and D. G. Truhlar. Non-Born-Oppenheimer molecular dynamics. *Acc. Chem. Res.*, 39:101–108, 2006.

[209] D. V. Shalashilin. Nonadiabatic dynamics with the help of multiconfigurational Ehrenfest method: Improved theory and fully quantum 24D simulation of pyrazine. *J. Chem. Phys.*, 132:244111, 2010.

[210] T. Yonehara, K. Hanasaki, and K. Takatsuka. Fundamental approaches to nonadiabaticity: Toward a chemical theory beyond the Born-Oppenheimer paradigm. *Chem. Rev.*, 112:499–542, 2012.

[211] K. Saita and D. V. Shalashilin. On-the-fly ab initio molecular dynamics with multiconfigurational Ehrenfest method. *J. Chem. Phys.*, 137:22A506–8, 2012.

[212] M. Vacher, M. J. Bearpark, and M. A. Robb. Direct methods for non-adiabatic dynamics: connecting the single-set variational multi-configuration Gaussian (vMCG) and Ehrenfest perspectives. *Theor. Chem. Acc.*, 135:187–11, 2016.

[213] S. Mahapatra, G. A. Worth, H.-D. Meyer, L. S. Cederbaum, and H. Köppel. The $\tilde{A}^2E/\tilde{B}^2B_2$ photoelectron bands of allene beyond the linear coupling scheme: An *ab initio* dynamical study including all fifteen vibrational modes. *J. Phys. Chem. A*, 105:5567–5576, 2001.

[214] M. Vacher, F. E. A. Albertani, A. J. Jenkins, I. Polyak, M. Bearpark, and M. A. Robb. Electron and nuclear dynamics following ionisation of modified bismethylene-adamantane. *Farad. Discuss.*, 194:95–115, 2016.