Winter 2021

Extension of Restricted Open-Shell Kohn-Sham Methodology to a Density-Functional Tight-Binding Framework

Reuben Szabo
Western Washington University, szabor@wwu.edu

EXTENSION OF RESTRICTED OPEN-SHELL KOHN-SHAM METHODOLOGY TO A DENSITY-FUNCTIONAL TIGHT-BINDING FRAMEWORK

By

REUBEN ALAN SZABO

Accepted in Partial Completion
of the Requirements for the Degree
Master of Science

ADVISORY COMMITTEE

Dr. Tim Kowalczyk, Chair

Dr. Jay McCarty

Dr. Armin Rahmani

WESTERN WASHINGTON UNIVERSITY
Graduate School

Dr. David L. Patrick, Dean

MASTER'S THESIS

In presenting this thesis in partial fulfillment of the requirements for a master's degree at Western Washington University, I grant to Western Washington University the non-exclusive royalty-free right to archive, reproduce, distribute, and display the thesis in any and all forms, including electronic format, via any digital library mechanisms maintained by WWU.

I represent and warrant this is my original work, and does not infringe or violate any rights of others. I warrant that I have obtained written permissions from the owner of any third party copyrighted material included in these files.

I acknowledge that I retain ownership rights to the copyright of this work, including but not limited to the right to use all or part of this work in future works, such as articles or books.

Library users are granted permission for individual, research and non-commercial reproduction of this work for educational purposes only. Any further digital posting of this document requires specific permission from the author.

Any copying or publication of this thesis for commercial purposes, or for financial gain, is not allowed without my written permission.

Reuben Alan Szabo

March 2021

EXTENSION OF RESTRICTED OPEN-SHELL KOHN-SHAM METHODOLOGY TO A DENSITY-FUNCTIONAL TIGHT-BINDING FRAMEWORK

A Thesis
Presented to
The Faculty of
Western Washington University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Science

By
Reuben Alan Szabo
March 2021

ABSTRACT

The restricted open-shell Kohn-Sham (ROKS) approach for singlet excited states provides some advantages over the Δ-self-consistent-field (ΔSCF) method, requiring only a single SCF procedure and avoiding the problem of variational collapse. While ROKS is a powerful tool for DFT, its application to density functional tight-binding (DFTB) could offer significant improvements in time complexity when compared to DFT, enabling excited-state simulations of extended molecular systems on longer timescales than ROKS. In this work we discuss the implementation of an RO-DFTB approach in the DFTB+ package, as well as its suitability for the study of organic dyes and photoactive compounds. For benchmarking, RO-DFTB is evaluated using vertical excitation energies and Stokes shifts, with reference values from ROKS, higher levels of theory, and experiment. Numerical gradients are also computed to evaluate the RO-DFTB analytical geometry optimization. We compare several different versions of the RO-DFTB implementation, and use ΔDFTB to assess whether RO-DFTB can offer improvements over the current excited-state options in DFTB.

ACKNOWLEDGMENT

Many thanks to all those who have contributed their time, knowledge, and support to make this thesis possible. Above all, I would like to express my deep gratitude to Professor Tim Kowalczyk, not only for the financial support and experience that made this project possible, but also for his mentorship and unending patience. I have been very fortunate to have him as my advisor.

I would also like to thank the members of my committee for their feedback and encouragement. I mostly ignored their pragmatic suggestion to simultaneously work on a back-up project in case RO-DFTB didn't come to fruition, but thankfully this did not cause a problem. (Thank you Ryan Lingg, Brian Barragan-Cruz, and Professor Brian Hutchinson for your contributions to the back-up project.)

I would like to thank Professor Benjamin Hourahine for sharing his deep knowledge of DFTB+, even though I was too new to the program at the time to understand most of it. A special thanks also to Megan Deshaye and Gunnar Carlson for their camaraderie in DFTB+ programming.

Thank you to all those who have shared the computational lab space over the years, especially Haley Doran and Megan Deshaye. It was wonderful spending time together indoors with you, and you always made the lab feel like home.

I also owe a great deal of thanks to my parents, for their lifelong support and unwavering faith in my ability.

Finally, I would like to thank my two cats, who slept through my Zoom defense.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter One

INTRODUCTION

1.1 Motivating RO-DFTB

The study of the electronic excited state is vital to our efforts toward meeting our energy needs with solar power.[1,2] The photochemical processes at play in photosynthetic compounds, photovoltaics, and photoisomers demand advanced electronic structure techniques capable of providing high-quality potential energy surfaces (PESs).[3] At the same time, the need for modeling of realistic systems requires approximate methods which balance accuracy and feasibility.[3,4] Excited-state methods which push the envelope in either direction, by improving accuracy or reducing computational cost, are highly coveted. Additionally, computational simulations can provide a cost-effective, time-saving, and less environmentally damaging alternative to experimental synthesis and characterization.

One particular application of excited-state electronic structure methods is in the development of solar thermal fuels (STFs).[5,6] These photoisomers offer a novel way of both absorbing and storing solar energy, by undergoing internal conversion to a metastable isomer after photoexcitation.[7–12] Part of the energy of the absorbed photon is stored chemically, and is then released as heat when some heat or a catalyst is applied to overcome the isomerization energy barrier. The challenge for advancement of these materials is that the four properties shown in Figure 1.1 must all fall within a suitable range in order for the fuel to be viable.[13]

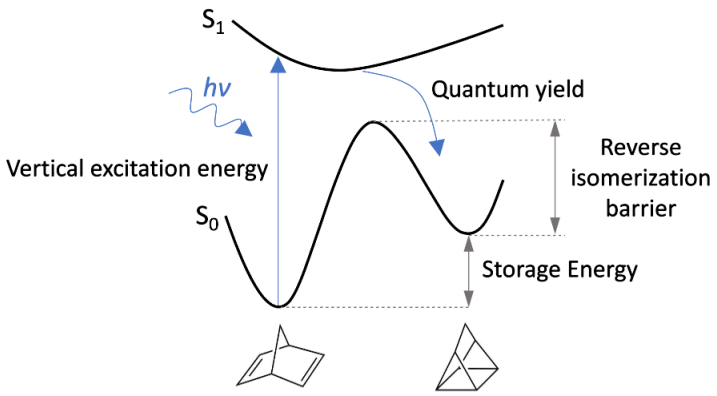

Figure 1.1 Potential energy surface of norbornadiene and its isomer quadricyclane. The four properties central to STF viability and their origin on the ground and excited-state PESs are shown.

The vertical excitation energy is one key STF property which ought to lie within the visible light range of the solar spectrum for effective photon absorption.[14] Unfortunately, the four properties are highly interdependent, so the current challenge is to find compounds for which the tradeoffs are minimal, so that all four properties can be optimized.[11,15–17]

Excited-state simulation is a key component in the discovery of new STFs as part of a screening procedure: rather than choosing STF candidates one at a time, whole libraries of thousands of compounds can be assessed computationally.[8] The ideal electronic method for this study would be able to capture ground and excited-state energies, and would be as resource-efficient as possible without sacrificing the ability to differentiate between good and bad candidates.

This screening procedure can be taken a step further with the inclusion of a machine learning (ML) algorithm.[18,19] By training the neural network with a featurization of the molecules and the results of the computational study, the ML program could learn to predict STF candidacy without the need for electronic structure methods beyond the training and testing stages. In this way the number of molecules in the screening library could be greatly increased.

The general framework of such a ML method is shown in Figure 1.2. There are many

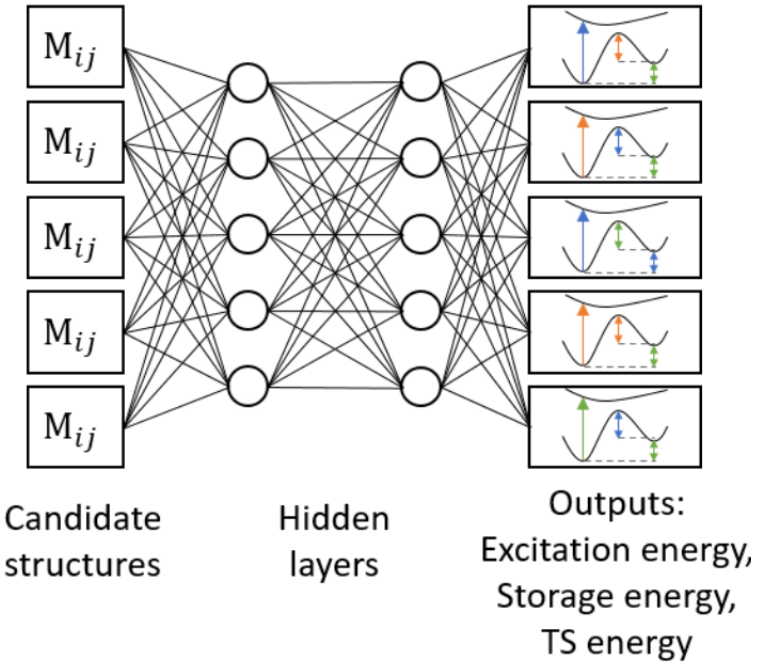

Figure 1.2 Visualization of the molecular inputs, hidden nodes in the neural network, and STF-related outputs for the machine learning solar thermal fuel project.

options available for the featurization of the candidate structures, with the Coulomb matrix (represented as M_{ij} being only one of them.[20] The purpose of the featurization is to capture as many properties of the molecular system as possible, so that the neural network has the best chance of finding patterns between the input and output layers. If the algorithm is successful enough and the library of compounds is diverse, it is possible that the resulting trained neural network could even predict new STF candidate structures not in the original library.

The significance of ML projects like this one is ever-increasing, as molecular simulations of all types are becoming more successful by making use of large data sets with machine learning.[21] But equally as important as the neural network is the electronic structure method used to train the program. In this project, the computational program must offer both accuracy and speed; the most accurate methods will not be able to generate the amount of data required in a reasonable time, but the data that is generated must do a reasonable job of evaluating the STF properties correctly.

The aim of this thesis is to first provide an overview of the theory behind an electronic excited-state implementation which combines two already-established computational tools to provide new functionality. Second, we aim to show that, of the multitude of electronic structure methods available—and the much smaller number of approximate excited-state methods—the method developed for this work offers unique advantages that make it a useful and robust tool for the many applications where highly resource-intensive methods are not an option.

This method is of course restricted open-shell density-functional tight-binding (RO-DFTB), which is specifically designed for singlet excited states and thus well-suited to study photoactive compounds. To put RO-DFTB in context with the other similar methods available, and even explain what its name means, we will start by briefly covering the foundations of molecular orbital theory.

1.2 Molecular orbital theory

From a chemical standpoint, we can identify five contributions to the energy of a molecular system: two kinetic energies T_e and T_n for the electrons and nuclei, respectively; the attractive Coulomb energy E_{ne} between nuclei and electrons; and the two repulsive energies between pairs of electrons, E_{ee}, and pairs of nuclei, E_{nn}.

$$E = -T_e - T_n - E_{ne} + E_{ee} + E_{nn} \tag{1.1}$$

To calculate this energy we need both the wavefunction and the Hamiltonian energy operator, according to the time-independent Schrödinger equation,

$$H\Psi = E\Psi \tag{1.2}$$

Two of the energy contributions can be disregarded with the Born-Oppenheimer approximation, which assumes that the nuclei move much slower than the electrons. The T_n and

E_{nn} energy terms can be calculated classically and added separately, since they depend only on the nuclei. We then define the Hamiltonian operator so that its expectation value yields the remaining three energy terms for a system of N electrons and M atoms. In atomic coordinates, the Hamiltonian is

$$H = -\sum_{i=1}^{N} \frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}} \tag{1.3}$$

The first term corresponds to the electron kinetic energy, the second to the Coulomb interaction of each electron with each nuclei, and the final term accounts for the interaction of every electron with every other electron.

The wavefunction Ψ describes the position and spin of all of the electrons in the molecular system. A single-electron wavefunction consists of a spatial orbital ψ, which depends only on position, and a spin function. These components are separable, as shown in Eqn. 1.4.

$$\begin{aligned} \chi^{\alpha}(\mathbf{x}) &= \psi(\mathbf{r})\alpha(\omega) \\ \chi^{\beta}(\mathbf{x}) &= \psi(\mathbf{r})\beta(\omega) \end{aligned} \tag{1.4}$$

The spin functions $\alpha(\omega)$ and $\beta(\omega)$ represent spin-up ($\uparrow$) and spin-down ($\downarrow$) electrons, respectively. The single-electron wavefunction is the product χ of the spatial orbital and spin function is known as a spin orbital. (Note that the variable $\mathbf{x}$ combines spatial coordinates $\mathbf{r}$ and the spin variable ω.) Because there are two possible spins, each spatial orbital corresponds to two spin orbitals. The spatial orbitals of a molecular system are the molecular orbitals, and each one may be associated with two electrons of opposite spin as we know from the Pauli exclusion principle. But while wavefunctions depend on spin, the Hamiltonian does not.

To describe a multi-electron wavefunction, we need to combine single-electron orbitals. The Hartree product is one option, which is a simple product of all one-electron orbitals. This product is an eigenfunction of the Schrödinger equation, with a Hamiltonian constructed from a sum of the single-electron Hamiltonians and an energy eigenvalue that is the sum of

the single-electron energies. For a two-electron case, with spin-up electron 1 and spin-down electron 2 in molecular orbital χ (i.e. the ground state helium atom), the Hartree product is

$$\Psi^{\mathrm{HP}} = \chi^{\alpha}(\mathbf{x}_1)\chi^{\beta}(\mathbf{x}_2) \tag{1.5}$$

which corresponds to the Hamiltonian $H = h(1) + h(2)$ and total energy $E = \varepsilon_1 + \varepsilon_2$. Of course, a single-electron Hamiltonian $h(1)$ would not include the electron-electron repulsion term in 1.3, but might be able to include an approximation.

The Hartree product, though, does not obey the antisymmetry principle. To fix this, we need to ensure that the wavefunction remains identical when any two electron spins or coordinates are exchanged—except for a change in sign. Hartree products can be efficiently antisymmstrized by using Slater determinants of spin orbitals to define the wavefunction. For example, the wavefunction of helium from above is represented by the Slater determinant in Eqn. 1.6.

$$\begin{aligned}\Psi &= \begin{vmatrix} \chi^{\alpha}(\mathbf{x}_1) & \chi^{\beta}(\mathbf{x}_2) \\ \chi^{\beta}(\mathbf{x}_2) & \chi^{\alpha}(\mathbf{x}_1) \end{vmatrix} \\ &= \frac{1}{\sqrt{2}}\left[\chi^{\alpha}(\mathbf{x}_1)\chi^{\beta}(\mathbf{x}_2) - \chi^{\beta}(\mathbf{x}_1)\chi^{\alpha}(\mathbf{x}_2)\right]\end{aligned} \tag{1.6}$$

The Slater determinant automatically enforces the Pauli exclusion principle, since a wavefunction with two electrons of the same spin in a shared spatial orbital will vanish. In this way it accounts for electron exchange by correlating the motion of electrons with the same spin. However, electrons of opposite spin remain uncorrelated. But since fully accounting for electron correlation is typically done by using multiple determinants, the use of a single Slater determinant is an excellent place to start. This is known as the Hartree-Fock approximation, which we will cover in more depth in section 2.1.

As we will show in the derivation of the HF equations, the energy expression obtained by applying a Slater determinant wavefunction to the Hamiltonian in Eqn. 1.3 accounts for electron-electron repulsion by using two-electron integrals. This means that instead

of instantaneous interactions between each pair of electrons, an electron essentially feels a "field" of all the other electrons. Consequentially, calculating the average repulsion requires knowing every molecular orbital, so the wavefunction must already be known in order to calculate the Hamiltonian. But our only means of calculating the wavefunction is by solving the Schrödinger equation (or some approximation of it), which requires the Hamiltonian first! This non-linearity can only be addressed with an iterative process, which we call the self-consistent field (SCF) procedure.

Central to the idea of self-consistency is the variational principle, which states that the exact wavefunction solution to our Hamiltonian will correspond to the lowest energy eigenvalue; all approximate guesses at the wavefunction will be higher in energy. This means we can build the Hamiltonian with an initial guess wavefunction, find the energy associated with it, and then vary the wavefunction and repeat the process until the energy stops changing.

To derive a means of consistently varying a wavefunction to achieve an energy minimum, we make use of the linear variation method. This employs the Lagrange method of undetermined multipliers and expands each molecular orbital as a linear combination of K basis functions ϕ.

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i}\phi_\mu \tag{1.7}$$

In practice, this introduces another approximation: the exact solution could be obtained with an infinite basis set, but we are constrained to using only finite sets ($K < \infty$). Therefore the size of the basis set has significant implications for the accuracy of the solution. On the smallest end of the spectrum, a minimal basis set includes only as many functions as there are atomic orbitals, so the molecular orbitals are called a Linear Combination of Atomic Orbitals (LCAO). For more accuracy, the basis set can get as large as available resources allow, but since the number of two-electron integrals required increases by (at worst) K^4, large basis sets can quickly become impractical. Because the basis set is chosen beforehand and does not change throughout the calculation, it allows the eigenvalue problem to be solved, and

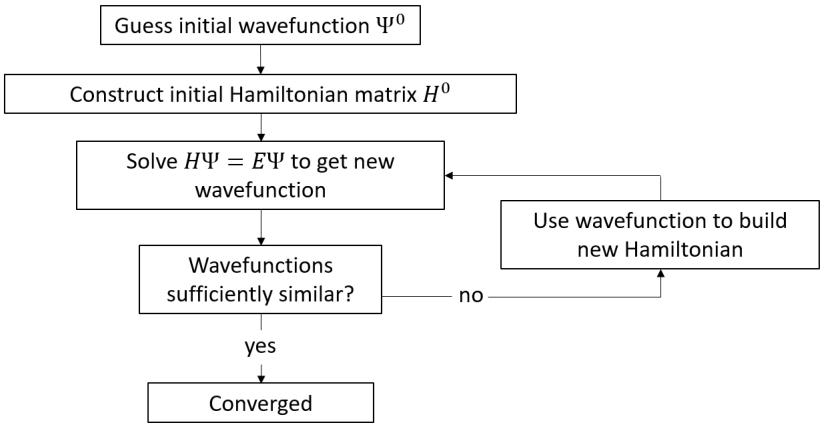

Figure 1.3 A simple self-consistent field procedure.

the energy minimized, with matrices. These are a convenient way to keep track of terms and perform operators during the SCF procedure, which is outlined in Figure 1.3.

Once the SCF routine has converged and a solution of the eigenvalue equation is reached, the molecular orbital energies are directly available as the energy eigenvalues, and the molecular orbital coefficients can be used with the basis functions to visualize the molecular orbital geometries. The total energy can be calculated from the orbital energies, partial atomic charges can be estimated, and analytical gradients can even be derived for geometry optimization. To obtain an optimized molecular geometry, the gradients are used to update the atomic coordinates, the molecular orbitals are converged all over again, and the process continues until the analytical gradients are near zero.

Even though we define the excitation energy as the energy required to promote one electron from an occupied to an unoccupied orbital, it cannot be well approximated by the energy difference between those two orbitals. Firstly, the total energy that MO theory minimizes is summed over only the occupied orbitals, so virtual (unoccupied) orbitals do not get optimized. Secondly, as an electron excites from one orbital to another, all other orbitals shift in response—and on the same timescale.

Instead, to calculate excitation energies the energy of the excited state must be calculated

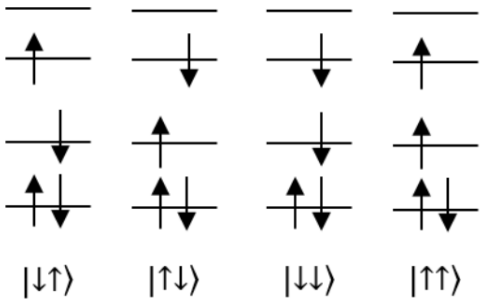

Figure 1.4 Four different spin configurations of a helium atom with a singly excited electron.

independently from the ground state configuration. Oftentimes, ground states are closed-shell systems where all electrons are spin-paired, and the excitation of a single electron from the HOMO to the LUMO creates two partially occupied orbitals with a single electron each. For helium, molecular orbital diagrams of the four possible spin configurations with one electron excited from HOMO to LUMO are shown in Fig. 1.4.

Each of these configurations can be represented with a single determinant, but none of them represent the singlet excited state, which is the state of interest in studying photoexcitation. In fact, the two excited configurations with mixed spins (those on the left) are not even eigenstates of the spin operator; instead, their symmetric combination is a triplet state with the same energy as the parallel-spin states, and their antisymmetric combination is the singlet excited state.

$$\begin{aligned} \text{Triplet states}: &\quad |\uparrow\uparrow\rangle\,, \quad |\downarrow\downarrow\rangle\,, \quad \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ \text{Singlet state}: &\quad \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \end{aligned} \tag{1.8}$$

So not only does the analysis of the singlet excited state require partially filled orbitals, which are much rarer for ground state calculations, but it also cannot be done with a single determinant! Section 1.3 will provide an overview of computational methods available which can handle excited-state calculations.

1.3 Computational methods for excited states

So far the only computational method we have discussed is Hartree-Fock theory, which is the foundation of a family of methods known as *ab initio* methods, since they rely only on the Schrödinger equation and quantum mechanical principles. There are several post-Hartree-Fock *ab initio* methods which improve on Hartree-Fock theory by accounting for electron correlation in different ways. Three of the most common are configuration interaction (CI),[22] various orders of Møller-Plesset perturbation theory (MP2, MP3, MP4),[23] and coupled-cluster (CC) methods.[24] These methods are some of the most accurate electronic structure methods available, and have some excited-state capabilities, but their computational cost is far above that of DFTB; they can generally only be used for small molecules, and are infeasible or intractable for most of the applications where DFTB is used.

Opposite to the post-Hartree-Fock methods on the scale of electronic structure accuracy are the semi-empirical methods, which supplement quantum chemical theory and mathematical approximation with parameterization from *ab initio* methods or experimental data. There is a wide range of these methods available, some of which pertain only to very limited sets of compounds, but some of them are capable of calculating excited states, such as the ZINDO method, which is developed from the earlier INDO method.[25] (INDO stands for 'intermediate neglect of differential overlap'). In general these methods are not successful at both geometry optimizations and excited states, and cannot be relied on for excited-state PESs.

In-between these levels of theory is density functional theory, or DFT.[26,27] Kohn-Sham DFT has earned its place as the most popular electronic structure method today, due to its high accuracy and low computational cost, compared to the other methods available. While DFT is an exact method in theory, in practice it relies on approximate functionals to correct for the effects of electron exchange and correlation and to correct for its non-interacting electron approximation. Although these functionals provide powerful time savings while

rivaling the accuracy of some *ab initio* methods, their overwhelming number and the difficulty of finding a functional which offers consistent accuracy across applications are two of the main drawbacks of the method.

DFT is a very powerful tool for studying medium-sized molecules. But there are still many applications—proteins, liquids, molecular dynamics simulations, large screening studies—for which DFT is too slow. For these applications there is a semi-empirical method, density-functional tight-binding, or DFTB, which borrows from the advantages (and disadvantages) of DFT and makes them available for even larger systems.[28–31] The tight-binding formalism from which it also borrows is a band structure model in solid-state physics which uses a linear combination of atomic orbitals (LCAO) to describe the wavefunction.[32] Combining the LCAO with DFT allows DFTB to be around 1,000 times faster than DFT, while being less dependent on finely tuned parameter sets than other semiempirical models; this gives DFTB excellent transferability.[33]

There are a wide range of excited-state DFT-based methods available.[34] Perhaps most well-known is linear-response time-dependent density functional theory, or TDDFT.[35] This method has the advantage of being able to compute any excited state, and offers an excellent combination of accuracy and efficiency like DFT itself. But there are drawbacks to using the time-dependent Schrödinger equation.[36] For comparison between ground and excited states, time-independent methods are more desirable due to their even-handed treatment of both states.

The ΔSCF method offers an alternative to TDDFT which is both time-independent and comparable in accuracy.[37] As its name suggests, the ΔSCF energy is the difference between two SCF-converged energies. ΔSCF obtains the singlet excited-state energy by optimizing the non-Aufbau triplet and mixed-spin determinants separately. As shown by Ziegler *et al.*,[38,39] a linear combination of these two energies produces the correct singlet excited-state energy: $E_s = 2E_m - E_t$. (For a complete derivation, see section A.1.) The main drawback of the ΔSCF method is that excited orbitals are prone to what is known as variational

collapse. That is, from one SCF iteration to the next, a single occupied excited state can collapse down to the ground state, destroying the excited-state calculation. Modifications have been implemented to correct for this problem, including the maximum overlap method (MOM),[40] but ROKS offers a simpler solution to the calculation of the first singlet excited state that is not prone to variational collapse.[41,42]

ROKS stands for 'restricted open-shell Kohn-Sham.' Like ΔSCF this method only calculates singlet excited states (although it can also be applied to open-shell ground states depending on the implementation). Rather than perform two full SCF routines like ΔSCF, ROKS combines the mixed and triplet determinants into a single Hamiltonian matrix, which affords it some improvement in time complexity as well as avoiding variational collapse.

TDDFT and ΔSCF are both implemented in the DFTB software package DFTB+ (there called TDDFTB and ΔDFTB, respectively).[43] And since ROKS has been shown to be a robust method in DFT capable of offering the advantages mentioned above,[41] an extension of the restricted open-shell formalism to the DFTB framework makes sense. If successful, RO-DFTB could help enhance photochemical computation in applications involving large systems near the upper limit of current electronic excited-state capabilities.

1.4 Overview

Having outlined the context of the present work, in chapter 2 we will detail the mathematical foundation of DFTB and ROKS. We will start with Hartree-Fock theory and restricted open-shell Hartree-Fock (ROHF), and then proceed to Kohn-Sham DFT and DFTB. Chapter 3 will introduce the current implementation of RO-DFTB in the DFTB+ software package,[44] and in chapter 4 we will provide benchmarking statistics along with an analysis of the different options available in RO-DFTB. Finally, chapter 5 will offer a perspective on the outlook of the RO-DFTB method.

Chapter Two

THEORY

2.1 The Hartree-Fock approximation

In this section, the discussion will be constrained to restricted Hartree-Fock (RHF), where all occupied orbitals contain two spin-paired electrons. This eliminates the need to account for spin, so molecular orbitals will be represented with only their spatial component ψ. In addition, the number of occupied orbitals will always be half the number of electrons, or $N/2$.

As we have seen, the Hartree-Fock approximation solves the Schrödinger equation for a wavefunction given by a single Slater determinant. Unlike Hartree products, the Slater determinant is not an eigenfunction of a Hamiltonian made up of one-electron Hamiltonians. However, we can derive the one-electron Fock operator by first using a Slater determinant wavefunction to calculate the energy from the Hamiltonian. With the ground state Slater determinant written in Dirac notation as $|\Psi_0\rangle$, the energy is the expectation value of the Hamiltonian operator: $E = \langle\Psi_0|H|\Psi_0\rangle$. (For a review of Dirac notation, see the mathematical review of Szabo and Ostlund,[45] whose work this section follows closely.)

Inserting the expression for the Hamiltonian from Equation 1.3, we have

$$E = \langle\Psi_0| - \sum_i^N \frac{1}{2}\nabla_i^2 - \sum_i^N \sum_A^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}|\Psi_0\rangle \tag{2.1}$$

The wavefunction $|\Psi_0\rangle$ can also be expressed as $|...\psi_a\psi_b...\rangle$, where orbitals ψ_a and ψ_b are two

of the orbitals in the antisymmetrized product. So if we define the core Hamiltonian, h, as

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_A^M \frac{Z_A}{r_{iA}} \tag{2.2}$$

we can separate it from the two-electron operator and write the energy expression in terms of the $N/2$ orthogonal molecular orbitals:

$$E = 2\sum_a^{N/2} \langle\psi_a|h|\psi_a\rangle + \sum_a^{N/2}\sum_b^{N/2} 2\langle\psi_a\psi_b|\psi_a\psi_b\rangle - \langle\psi_a\psi_b|\psi_b\psi_a\rangle \tag{2.3}$$

The final two terms are the Coulomb and exchange integrals, which arise from applying the two-electron interelectron distance operator r_{ij}^{-1} to the Slater determinant. These integrals have been written using the notation

$$\langle\psi_a\psi_b|\psi_c\psi_d\rangle = \langle ab|cd\rangle = \int \psi_a^*(1)\psi_b(1)\frac{1}{r_{12}}\psi_c^*(2)\psi_d(2)\mathrm{d}\mathbf{r}_1\mathrm{d}\mathbf{r}_2 \tag{2.4}$$

The Coulomb integral is $J_{ab} = \langle\psi_a\psi_b|\psi_a\psi_b\rangle = \langle ab|ab\rangle$, which describes electron-electron repulsion from Coulomb interactions. The exchange operator, $K_{ab} = \langle\psi_a\psi_b|\psi_b\psi_a\rangle = \langle ab|ba\rangle$, arises from the antisymmetry in the Slater determinant, and accounts for the decreased probability of finding two electrons of the same spin near each other. The exchange integral is the only term not multiplied by a factor of two (since there are two paired electrons per orbital) because only electrons of the same spin are affected by the exchange operator.

Now that we have an energy expression, the Hartree-Fock equations can be obtained by minimizing the energy, subject to the constraint that the molecular orbitals remain orthonormal. To do this, we use the Lagrange method of undetermined multipliers and the calculus of variations. This process is briefly laid out here because it will come in handy for understanding the restricted open-shell method later.

The Lagrangian $\mathscr{L}$ subtracts from the energy in Equation 2.3 a constraint (in this case $\langle a|b\rangle - \delta_{ab} = 0$, which ensures orthonormality) multiplied by the Lagrange multiplier ε_{ba}:

$$\mathscr{L}[\{\psi_a\}] = E_0[\{\psi_a\}] - 2\sum_a^{N/2}\sum_b^{N/2} \varepsilon_{ba}(\langle a|b\rangle - \delta_{ab}) \tag{2.5}$$

(The factor of two is arbitrary and its convenience will become apparent momentarily.) To minimize the $\mathscr{L}$, and therefore the energy, we take the first variation of the Lagrangian, $\delta\mathscr{L}$, and set it equal to zero. The first variation is obtained by varying the molecular orbitals by an arbitrary, infinitesimal amount such that $\psi_a \to \psi_a + \delta\psi_a$. After some manipulation, we can write the first variation of the Lagrangian as

$$\begin{aligned}\delta\mathscr{L} =& 2\sum_a^{N/2} \langle\delta\psi_a|h|\psi_a\rangle + 2\sum_a^{N/2}\sum_b^{N/2} 2\,\langle\delta\psi_a\psi_a|\psi_b\psi_b\rangle - \langle\delta\psi_a\psi_b|\psi_b\psi_a\rangle \\ &- 2\sum_a^{N/2}\sum_b^{N/2} \varepsilon_{ba}\,\langle\delta\psi_a|\psi_b\rangle + \text{complex conjugate} = 0\end{aligned} \tag{2.6}$$

In particular, since $\mathscr{L}$ is real and $\langle\psi_a|\psi_b\rangle = \langle\psi_b|\psi_a\rangle^*$, the Lagrange multipliers must make a Hermitian matrix where $\varepsilon_{ab} = \varepsilon_{ba}^*$. Note that the Coulomb and exchange integrals have acquired a factor of two here due to their two-electron nature and the multiplicity of applying the first variation.

Now, introducing the Coulomb and exchange operators (which have as their expectation values the Coulomb and exchange integrals),

$$J_a(1)\psi_b(1) = \left[\int \psi_a^*(2) r_{12}^{-1} \psi_a(2) \mathrm{d}\mathbf{r}_2\right]\psi_b(1) \tag{2.7}$$

$$K_a(1)\psi_b(1) = \left[\int \psi_a^*(2) r_{12}^{-1} \psi_b(2) \mathrm{d}\mathbf{r}_2\right]\psi_a(1) \tag{2.8}$$

we can rearrange to write

$$\begin{aligned}\delta\mathscr{L} =& 2\sum_a^{N/2} \int \mathrm{d}\mathbf{r}_1 \delta\psi_a^*(1)\left[h(1)\psi_a(1) + \sum_b^{N/2}(2J_b(1) - K_b(1))\psi_a(1) - \sum_b^{N/2}\varepsilon_{ba}\psi_b(1)\right] \\ &+ \text{complex conjugate} = 0\end{aligned} \tag{2.9}$$

We can't say whether or not ψ_a^* is zero, since this equation holds for all a, so the term in square brackets must be zero. We can then write

$$h(1)\psi_a(1) + \sum_b^{N/2}(2J_b(1) - K_b(1))\psi_a(1) = \sum_b^{N/2}\varepsilon_{ba}\psi_b(1) \tag{2.10}$$

where we now have an expression for the one-electron Fock operator:

$$f(1) = h(1) + \sum_{a}^{N/2} 2J_a(1) - K_a(1) \tag{2.11}$$

so we can rewrite Equation 2.10 as

$$f(1)\psi_a = \sum_{b}^{N/2} \varepsilon_{ba}\psi_b(1) \tag{2.12}$$

But these equations, the Euler-Lagrange equations, are not quite the eigenvalue equations we want. After a unitary transformation to diagonalize the matrix of Lagrange multipliers ε_{ba}, we obtain a set of eigenvalue equations that are the Hartree-Fock equations.

$$f(1)\psi_i(1) = \varepsilon_i\psi_i(1) \tag{2.13}$$

It is now clear that the Lagrange multipliers, as the eigenvalues of the one-electron Fock operator, are the molecular orbital energies. Notably, the unitary transformation does change the Lagrange multipliers and molecular orbitals, but the transformation can be chosen to localize the orbitals in line with our chemical understanding of bonding orbitals.

The next step is to design a way to solve the set of Hartree-Fock equations, and to do this we make use of the set of basis functions we have chosen. Recall that molecular orbital ψ_i can be expanded with basis functions ϕ_μ to give

$$\psi_i = \sum_{\mu=1}^{K} c_{\mu i}\phi_\mu \tag{2.14}$$

This expansion allows us to write the Fock operator in matrix form:

$$\begin{aligned} F_{\mu\nu} &= \int \phi_\mu^*(1) f(1) \phi_\nu(1) \mathrm{d}\mathbf{r}_1 \\ &= \langle \phi_\mu | h(1) + \sum_{a}^{N/2} 2J_a(1) - K_a(1) | \phi_\nu \rangle \end{aligned} \tag{2.15}$$

But because the basis functions that get used in quantum chemistry are not orthogonal (unlike the molecular orbitals themselves), we also need to introduce the overlap matrix, $\mathbf{S}$.

$$\begin{aligned} S_{\mu\nu} &= \int \phi_\mu^*(1) \phi_\nu(1) \mathrm{d}\mathbf{r}_1 \\ &= \langle \phi_\mu | \phi_\nu \rangle \end{aligned} \tag{2.16}$$

Using these components, the Hartree-Fock equations give rise to the Roothaan equations, which can be written succinctly in matrix form:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\varepsilon} \tag{2.17}$$

Here, the $\mathbf{C}$ matrix is made up of the expansion coefficients $c_{\mu i}$ from Equation 2.14, and $\boldsymbol{\varepsilon}$ is a matrix with the MO energy eigenvalues ε_i on the diagonal. All of the matrices are $K \times K$ square matrices, where K is the number of basis functions. This means that we now have as many molecular orbitals, represented by the columns of the $\mathbf{C}$ matrix, as we have basis functions.

To solve this matrix equation, we need to eliminate the overlap matrix by orthogonalizing the basis set. This yields the transformed Roothaan equations, $\mathbf{F}'\mathbf{C}' = \mathbf{C}'\boldsymbol{\varepsilon}$, where $\mathbf{F}'$ and $\mathbf{C}'$ are just the Fock and coefficient matrices in the orthogonal basis. In this basis, the Fock matrix is diagonalized to solve for $\mathbf{C}'$, and then they can both be transformed back to the original basis. Ultimately, all we need is $\mathbf{F}$ to solve for $\mathbf{C}$ and $\boldsymbol{\varepsilon}$.

Before we can go about this, though, we need a more explicit form of the Fock matrix in order to build it. To this end we introduce the charge density, which like other probability densities is the square of the wavefunction. The total charge density is summed over only the $N/2$ occupied orbitals,

$$\rho(\mathbf{r}) = 2\sum_{a}^{N/2} |\psi_a(\mathbf{r})|^2 \tag{2.18}$$

and it can be rewritten by expanding the molecular orbitals according to Equation 2.14:

$$\rho(\mathbf{r}) = \sum_{\mu}^{K}\sum_{\nu}^{K} P_{\mu\nu}\phi_\mu(\mathbf{r})\phi_\nu^*(\mathbf{r}) \tag{2.19}$$

Here we have defined elements of the density matrix $\mathbf{P}$ as

$$P_{\mu\nu} = 2\sum_{a}^{N/2} c_{\mu a}c_{\nu a}^* \tag{2.20}$$

Since the coefficients that make up the density matrix come from the coefficient matrix $\mathbf{C}$, the density matrix is another way of expressing the solution of the Roothaan equations.

Now, we can rewrite the terms of the Fock matrix from Equation 2.15 to separate the core Hamiltonian $H_{\mu\nu}^{\text{core}}$ from the two-electron part

$$\begin{aligned} F_{\mu\nu} &= \int \phi_\mu^*(1) h(1) \phi_\nu(1) \mathrm{d}\mathbf{r}_1 + \sum_a^{N/2} \int \phi_\mu^*(1) [2J_a(1) - K_a(1)] \phi_\nu(1) \mathrm{d}\mathbf{r}_1 \\ &= H_{\mu\nu}^{\text{core}} + \sum_a^{N/2} 2 \langle \mu\nu | aa \rangle - \langle \mu a | a\nu \rangle \end{aligned} \tag{2.21}$$

Finally, we can write the Fock matrix with the molecular orbital expansion from Equation 2.14, which shows its dependence on the density matrix.

$$F_{\mu\nu} = H_{\mu\nu}^{\text{core}} + \sum_\lambda^K \sum_\sigma^K P_{\lambda\sigma} \left[\langle \mu\nu | \sigma\lambda \rangle - \frac{1}{2} \langle \mu\lambda | \sigma\nu \rangle \right] \tag{2.22}$$

Now we are truly faced with the non-linearity of our matrix eigenvalue problem, although the Fock operator has always depended on the wavefunction in the form of the Coulomb and exchange operators. In order to build the Fock matrix, we need to use a guess density matrix, which requires a guess coefficient matrix, which is equivalent to a guess at the molecular orbitals. The only way to solve for the molecular orbitals, then, is with an iterative process like the self-consistent field (SCF) procedure, shown in Figure 2.1.

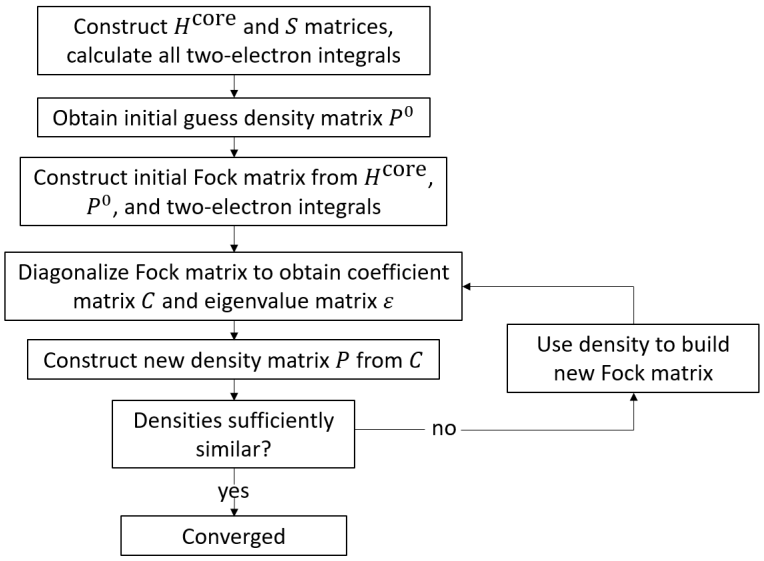

Figure 2.1 The Hartree-Fock self-consistent field (SCF) procedure.

In theory, convergence is straightforward, although it requires a somewhat decent initial guess. In practice, replacing the density matrix entirely on each iteration is quite inefficient, and convergence can be very slow. A variety of convergence acceleration algorithms exist which modify the next input density by combining it with one or more old densities; these are widely used in many different SCF calculations.

Once the SCF procedure has converged, we probably want to calculate the energy of the system. This is *not* the sum of the orbitals energies of the occupied orbitals, since this summation would double-count Coulomb and exchange interactions. Rather, it is the same expectation value of the Fock operator as in Equation 2.3. With subscripts representing expectation values (i.e., $\langle\psi_a|h|\psi_a\rangle = h_{aa}$), the energy equation is

$$E_0 = 2\sum_{a}^{N/2} h_{aa} + \sum_{a}^{N/2}\sum_{b}^{N/2} 2J_{ab} - K_{ab} \tag{2.23}$$

As the expectation values of the one-electron Fock operator, we can write the orbital energies ε_a as

$$\varepsilon_a = f_{aa} = h_{aa} + \sum_{b}^{N/2} 2J_{ab} - K_{ab} \tag{2.24}$$

so the energy becomes

$$E_0 = 2\sum_{a}^{N/2}(h_{aa} + \varepsilon_a) \tag{2.25}$$

By expansion with the basis functions, we can write the energy using the matrices that we have on hand during an SCF calculation: the density matrix, core Hamiltonian matrix, and Fock matrix.

$$E_0 = \frac{1}{2}\sum_{\mu}\sum_{\nu} P_{\nu\mu}(H^{\text{core}}_{\mu\nu} + F_{\mu\nu}) \tag{2.26}$$

For the total energy, we add the internuclear repulsion:

$$E_{\text{total}} = E_0 + \sum_{A}\sum_{B>A}\frac{Z_A Z_B}{R_{AB}} \tag{2.27}$$

There are many more properties that can be calculated, including dipole moments, but one that will be relevant to later discussion of the DFTB method are the charges associated

with each orbital and atom. In DFTB, this is obtained with a Mulliken population analysis, and it can be in Hartree-Fock as well. The expression for the total number of electrons, N, is the sum of integrated charge densities for each occupied orbital:

$$N = 2\sum_{a}^{N/2} \int |\psi_a(\mathbf{r})|^2 \mathrm{d}\mathbf{r} \tag{2.28}$$

Substituting in the basis function expansion, we have

$$\begin{aligned} N &= \sum_{\mu}^{K}\sum_{\nu}^{K} P_{\mu\nu}S_{\nu\mu} \\ &= \sum_{\mu}^{K} (\mathbf{PS})_{\mu\mu} \end{aligned} \tag{2.29}$$

Here, $(\mathbf{PS})_{\mu\mu}$ is the number of electrons in basis function ϕ_μ. If our basis functions are centered on atoms, then the partial charge on atom A is given by

$$q_A = Z_A - \sum_{\mu \in A} (\mathbf{PS})_{\mu\mu} \tag{2.30}$$

where $\mu \in A$ includes only the basis functions that are part of atom A. In DFTB, it is the charges, rather than the density matrix, which are carried from the end of one iteration to the next.

This concludes our discussion of restricted Hartree-Fock. But before we move on to density functional theory and its tight-binding counterpart DFTB, we will take the time to introduce the restricted open-shell method in the framework of the Hartree-Fock approximation. The theory outlined in the next section will be built on later for DFT and DFTB.

2.2 Restricted open-shell Hartree-Fock

The restricted Hartree-Fock theory we have covered has been limited to closed-shell systems where α and β spins occupy the same spatial orbitals. The restricted open-shell method in Hartree-Fock theory (ROHF) shares the advantages of restricted Hartree-Fock while also

allowing for partially occupied orbitals.[46–48] This makes it well-suited for studying the singlet excited state, but it can be applied to many open-shell systems. However, because the Fock operator depends on the symmetry of the system—how many orbitals are partially filled and what their degeneracies are—our discussion of ROHF (and restricted open-shell methods in general) will be limited to the singlet excited state.

The challenge of accounting for open shells in a restricted formalism is two-fold. First, accounting for different spin states of the open orbitals requires working with a wavefunction which is a linear combination of Slater determinants. As we have seen, the singlet excited state requires both the triplet determinant and a mixed determinant. Second, the partially occupied orbitals require a different Fock operator than the doubly occupied orbitals. This creates so-called coupling operators, the operators of one shell acting on another shell.[49]

In the singlet excited state there are three types of orbitals: those in the closed shell, c, the two open shells, a and b, and the orbitals in the unoccupied virtual shell, v. On the diagonal are the block matrices of operators between two shells of the same type. The coupling operators are the off-diagonal blocks between two different shells. The virtual Fock operator doesn't really matter as it doesn't contribute to the total energy (since these orbitals are unoccupied), but it is useful to include to obtain virtual orbital energies.[50] Of course, the ground state Fock matrix includes virtual orbitals as well, and treats them with the same operator as the closed shells. (The virtual orbitals of the ground state are simply all the orbitals between the number of occupied orbitals, $N/2$, and the number of basis functions K.) We will use the closed-shell Fock matrix for the ROHF virtual shell as well.

Energy minimization with the constraint of MO orthonormality is a bit more complicated for ROHF. Recall the result of setting the first variation of the Lagrangian equal to zero for the closed-shell ground state case: we obtained the Euler-Lagrange equations, Eq. 2.12, with the Lagrange multipliers making up a Hermitian matrix where $\varepsilon_{ij} = \varepsilon_{ji}^*$. In ROHF, we follow

a similar procedure, starting from the ROHF singlet energy:

$$E_s = 2\sum_{k}^{N^c} h_{kk} + h_{aa} + h_{bb} + \sum_{kl}^{N^c}(2J_{kl} - K_{kl}) + \sum_{k}^{N^c}(2J_{ka} - K_{ka}) + \\ + \sum_{k}^{N^c}(2J_{kb} - K_{kb}) + J_{ab} + K_{ab} \tag{2.31}$$

This energy is derived using unrestricted Hartree-Fock in the appendix. Indices k and l refer to the N^c closed-shell orbitals, a and b refer to the first and second open shell orbitals, respectively, and i and j refer to any orbital.

Next we set the first variation of the Lagrangian, $\mathscr{L} = E_s - \sum_{ij} \varepsilon_{ij}(\langle\psi_i|\psi_j\rangle - \delta_{ij})$, equal to zero. Separating the terms by shell and introducing the closed and open-shell Fock operators, we have

$$\delta\mathscr{L} = \sum_{k}\left[\langle\delta\psi_k|F_c|\psi_k\rangle - \sum_{j}\varepsilon_{kj}(\langle\delta\psi_k|\psi_j\rangle)\right] + \langle\delta\psi_a|F_a|\psi_a\rangle - \sum_{j}\varepsilon_{aj}(\langle\delta\psi_a|\psi_j\rangle) \\ + \langle\delta\psi_b|F_b|\psi_b\rangle - \sum_{j}\varepsilon_{bj}(\langle\delta\psi_b|\psi_j\rangle) + \text{complex conjugate} = 0 \tag{2.32}$$

where

$$F_c = 2h + 2\sum_{k}(2J_k - K_k) + (2J_a - K_a) + (2J_b - K_b) \tag{2.33}$$

$$F_a = h + \sum_{k}(2J_k - K_k) + J_b + K_b \tag{2.34}$$

$$F_b = h + \sum_{k}(2J_k - K_k) + J_a + K_a \tag{2.35}$$

These can be re-written in terms of the α and β mixed and triplet Fock operators:[41]

$$F_c = 2F_m^{\alpha} + 2F_m^{\beta} - F_t^{\alpha} - F_t^{\beta} \tag{2.36}$$

$$F_a = 2F_m^{\beta} - F_t^{\beta} \tag{2.37}$$

$$F_b = 2F_m^{\alpha} - F_t^{\beta} \tag{2.38}$$

Here the spin designations have been dropped from the definitions in the appendix (since

the α and β spin orbitals share spatial orbitals), so

$$F_m^\alpha = h + \sum_k^{N^c}(2J_k - K_k) + J_a + J_b - K_b \tag{2.39}$$

$$F_m^\beta = h + \sum_k^{N^c}(2J_k - K_k) + J_a + J_b - K_a \tag{2.40}$$

$$F_t^\alpha = h + \sum_k^{N^c}(2J_k - K_k) + J_a + J_b \tag{2.41}$$

$$F_t^\beta = h + \sum_k^{N^c}(2J_k - K_k) + J_a + J_b - K_a - K_b \tag{2.42}$$

These Fock operators correspond to the $|{\downarrow}{\uparrow}\rangle$ mixed determinant and $|{\downarrow}{\downarrow}\rangle$ triplet determinant, though either spin configuration of the determinants would have worked.

The variational conditions are the same as in the closed-shell ground state:

$$F_i \left|\psi_i\right\rangle = \sum_j \left|\psi_j\right\rangle \varepsilon_{ji} \tag{2.43}$$

$$\varepsilon_{ji} = \varepsilon_{ij}^* \tag{2.44}$$

But now that orbitals of different shells have different Fock operators, these can be rewritten:[51]

$$F_i \left|\psi_i\right\rangle = \sum_j \left|\psi_j\right\rangle \left\langle\psi_j|F_i|\psi_i\right\rangle \tag{2.45}$$

$$\langle\psi_j|F_i - F_j|\psi_i\rangle = 0 \tag{2.46}$$

To create a unified Fock operator[49] which satisfies the variational conditions, we follow the general SCF operator technique of Hirao and Nakatsuji,[50] introducing the operator

$$G_{ji} = \lambda_{ji}F_j + (1 - \lambda_{ji})F_i \tag{2.47}$$

where λ_{ji} is an arbitrary non-zero real number. The variational conditions can now be written

$$\left[F_i - \sum_j \left|\psi_j\right\rangle \left\langle\psi_j\right| G_{ji}\right] \left|\psi_i\right\rangle = 0 \tag{2.48}$$

After adding $|\psi_i\rangle \langle\psi_i|F_i|\psi_i\rangle$ to both sides and multiplying the operator in the square brackets by the projection operator $|\psi_i\rangle \langle\psi_i|$ (since ψ_i is the only orbital it acts on), we have

$$\left[F_i |\psi_i\rangle \langle\psi_i| - \sum_{j\neq i} |\psi_j\rangle \langle\psi_j| G_{ji} |\psi_i\rangle \langle\psi_i|\right] |\psi_i\rangle = \varepsilon_{ii} |\psi_i\rangle \tag{2.49}$$

For the virtual orbitals, we introduce the operator

$$V = \sum_{vw} |\psi_v\rangle \langle\psi_v| F_c |\psi_w\rangle \langle\psi_w| \tag{2.50}$$

where v and w index over the virtual orbitals. Summing the operator from Equation 2.49 over all orbitals and adding V, our unified operator is

$$F' = \sum_i \left[F_i |\psi_i\rangle \langle\psi_i| - \sum_{j\neq i} |\psi_j\rangle \langle\psi_j| G_{ji} |\psi_i\rangle \langle\psi_i|\right] + V \tag{2.51}$$

We then set $|\lambda_{ij}| = 1$ and $\lambda_{ji} = -\lambda_{ij}$, symmetrize the virtual orbitals, and obtain in matrix form the operator

$$\mathbf{F} = \begin{pmatrix} \mathbf{F}_c & \mathbf{F}_c - \mathbf{F}_a & \mathbf{F}_c - \mathbf{F}_b & \mathbf{F}_c \\ \mathbf{F}_c - \mathbf{F}_a & \mathbf{F}_a & \mathbf{F}_a - \mathbf{F}_b & \mathbf{F}_a \\ \mathbf{F}_c - \mathbf{F}_b & \mathbf{F}_a - \mathbf{F}_b & \mathbf{F}_b & \mathbf{F}_b \\ \mathbf{F}_c & \mathbf{F}_a & \mathbf{F}_b & \mathbf{F}_c \end{pmatrix} \tag{2.52}$$

This unified coupling operator replaces the Fock operator in the SCF routine, and is in theory all that is necessary to define the ROHF method. However, there are some additional considerations for implementing a restricted open-shell method. Firstly, the matrix representation in Eq. 2.52 is in the molecular orbital (MO) basis, so the block Fock matrices used to build it must be transformed from the atomic orbital (AO) basis to the MO basis. This can be accomplished with density matrices of the corresponding shell, which serve as projection operators:

$$\mathbf{P}^{\sigma}_{\mu\nu} = \sum_{i=1}^{N^{\sigma}} c_{\mu i} c^{*}_{\nu i}, \qquad \sigma \in (c, a, b, v) \tag{2.53}$$

So in the AO basis, the ROHF Fock matrix is the sum of the 16 blocks:

$$\mathbf{F} = \sum_{\sigma} \mathbf{S}\mathbf{P}^{\sigma}\mathbf{F}_{\sigma}\mathbf{P}^{\sigma}\mathbf{S} + \sum_{\sigma}\sum_{\gamma\neq\sigma} \mathbf{S}\mathbf{P}^{\sigma}(\mathbf{F}_{\sigma} - \mathbf{F}_{\gamma})\mathbf{P}^{\gamma}\mathbf{S}, \qquad \sigma, \gamma \in (c, a, b, v) \tag{2.54}$$

where the overlap matrix $\mathbf{S}$ is necessary if the basis set is not orthogonal.

A second consideration is that the off-diagonal blocks of the ROHF Fock matrix in the MO basis are also known as the error matrices, since they will be zero at convergence. This is in addition to the normal SCF convergence criteria that the density matrix is no longer changing. In principle, both conditions should be met at convergence.

Third, an issue can arise in ROHF calculations if the orbital energies between different shells are too similar.[52,53] In this case the higher-energy shells can be artificially raised to separate the orbitals; at convergence this does alter the orbital energies but not the total energy, since the diagonal block Fock operators are technically arbitrary. This level-shifting procedure adds an energy λ_σ to the diagonal of the block Fock matrix for shell $\sigma \in (c, a, b, v)$, i.e. $\mathbf{F}_\sigma \rightarrow \mathbf{F}_\sigma + \lambda_\sigma \mathbf{I}_\sigma$.

Finally, the calculation of analytical gradients for geometry optimization requires a special energy-weighted density matrix $\mathbf{W}$. For single-determinant unrestricted methods, this matrix is

$$\mathbf{W} = \mathbf{P}^\alpha \mathbf{F}^\alpha \mathbf{P}^\alpha + \mathbf{P}^\beta \mathbf{F}^\beta \mathbf{P}^\beta \tag{2.55}$$

while to accommodate the two separate density matrices, for ROHF the energy-weighted density matrix is

$$\mathbf{W}_{\mathrm{ROHF}} = \mathbf{P}^\alpha_t (2\mathbf{F}^\alpha_m \mathbf{P}^\alpha_m + 2\mathbf{F}^\beta_m \mathbf{P}^\beta_m + \mathbf{F}^\alpha_t \mathbf{P}^\alpha_t - \mathbf{F}^\beta_t \mathbf{P}^\beta_t) \tag{2.56}$$

With the foundations of restricted open-shell theory covered and these practical considerations noted, we turn briefly to Kohn-Sham density functional theory before delving into density-functional tight-binding.

2.3 Density functional theory

Kohn-Sham density functional theory is entirely analogous to Hartree-Fock theory, but with an enlightened simplification: instead of dealing with electron interactions in the Schrödinger

equation, only the non-interacting kinetic energy is solved for.[54] The effects of all particle interactions are then accounted for with a potential which is a function of the electron density (and the density is the same as in a fully interacting system). To see how this works out, we will start with the Kohn-Sham DFT energy expression, which is a sum of the non-interacting kinetic energy, the Coulomb repulsions as a function of electron density, an exchange-correlation energy, and the electron-nuclei interaction energy:

$$E_{KS}[\rho] = 2\sum_{i}^{N/2} \langle\psi_i| - \frac{1}{2}\nabla^2|\psi_i\rangle + J[\rho] + E_{\mathrm{XC}}[\rho] + E_{\mathrm{ext}}[\rho] \tag{2.57}$$

Here $J[\rho]$ is the same as the Coulomb interaction term we are familiar with, but written in terms of the electron density

$$J[\rho] = \frac{1}{2}\int \frac{\rho(\mathbf{r})\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|} \tag{2.58}$$

and the external potential is the the same electron-nuclei interaction we have used in Hartree-Fock theory, but again written with the density instead of the wavefunction:

$$E_{\mathrm{ext}} = \sum_{A}^{M} Z_A \int \frac{\rho(\mathbf{r})}{|\mathbf{r}-\mathbf{R}_A|}\mathrm{d}\mathbf{r} \tag{2.59}$$

According to the Hohenberg and Kohn theorems, the electron density ρ uniquely determines the external potential.[55] Also according to the Hohenberg and Kohn theorems, this energy functional obeys the variational principle, so we can minimize the energy in much the same way as in Hartree-Fock theory. The Lagrangian is

$$\mathscr{L} = E_{\mathrm{KS}} - 2\sum_{i}^{N/2}\sum_{j}^{N/2} \varepsilon_{ij}\left(\langle\psi_i|\psi_j\rangle - \delta_{ij}\right) \tag{2.60}$$

We take the first derivative rather than the first variation, recalling that the density is

$$\rho(\mathbf{r}) = 2\sum_{i}^{N/2} |\psi_i(\mathbf{r})|^2 \tag{2.61}$$

which means that the derivative of the Kohn-Sham energy, a functional of the density, can be written

$$\frac{\delta E_{\mathrm{KS}}[\rho]}{\delta\psi_j^*(\mathbf{r})} = \frac{\delta E_{\mathrm{KS}}[\rho]}{\delta\rho(\mathbf{r})}\frac{\partial\rho(\mathbf{r})}{\partial\psi_j^*(\mathbf{r})} = \frac{\delta E_{\mathrm{KS}}[\rho]}{\delta\rho(\mathbf{r})}2\psi_j(\mathbf{r}) \tag{2.62}$$

So, setting the derivative of the Lagrangian equal to zero, we have

$$\frac{\delta \mathscr{L}[\rho]}{\delta \psi_j^*(\mathbf{r})} = 2\left(-\frac{1}{2}\nabla^2\psi_j(\mathbf{r})\right) + 2\left(\frac{\delta J[\rho]}{\delta\rho} + \frac{\delta E_{\mathrm{XC}}[\rho]}{\delta\rho} + \frac{\delta E_{\mathrm{ext}}[\rho]}{\delta\rho}\right)\psi_j(\mathbf{r}) - 2\sum_i^{N/2}\varepsilon_{ij}\psi_j(\mathbf{r}) = 0 \tag{2.63}$$

With

$$\begin{aligned} V_{\mathrm{KS}} =& \frac{\delta J[\rho]}{\delta\rho} + \frac{\delta E_{\mathrm{XC}}[\rho]}{\delta\rho} + \frac{\delta E_{\mathrm{ext}}[\rho]}{\delta\rho} \\ =& \int \frac{\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r'} + V_{\mathrm{ext}}(\mathbf{r}) + V_{\mathrm{XC}}(\mathbf{r}) \end{aligned} \tag{2.64}$$

the energy minimization yields the Kohn-Sham equations:

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \varepsilon_i\psi_i(\mathbf{r}) \tag{2.65}$$

We now have a way to obtain the electron density, and we can use almost exactly the same self-consistent field procedure as in Hartree-Fock to iteratively reach the energy minimum. Introducing a finite basis to expand the orbitals ψ_i, as was done to convert the Hartree-Fock equations to the Roothaan equations, we can write the Kohn-Sham equations in matrix form,[27]

$$\sum_\nu^N (F_{\mu\nu} - \varepsilon_i S_{\mu\nu})c_{\nu i} = 0 \tag{2.66}$$

The overlap ($\mathbf{S}$), coefficient ($\mathbf{C}$), and eigenvalue ($\boldsymbol{\varepsilon}$) matrices are defined in the same way as in Hartree-Fock. However the Fock matrix is now different: the core Hamiltonian contains only the kinetic energy operator, the Coulomb matrix is defined separately, there are no terms from the exchange operator, and the exchange-correlation matrix is added, which varies depending on the exchange-correlation functional chosen.

$$F_{\mu\nu} = H_{\mu\nu}^{\mathrm{core}} + J_{\mu\nu} - F_{\mu\nu}^{\mathrm{XC}} \tag{2.67}$$

Just as in Hartree-Fock, the DFT self-consistent field procedure (shown in Fig. 2.2) requires an initial guess of the density. This is used to create the Kohn-Sham potential,

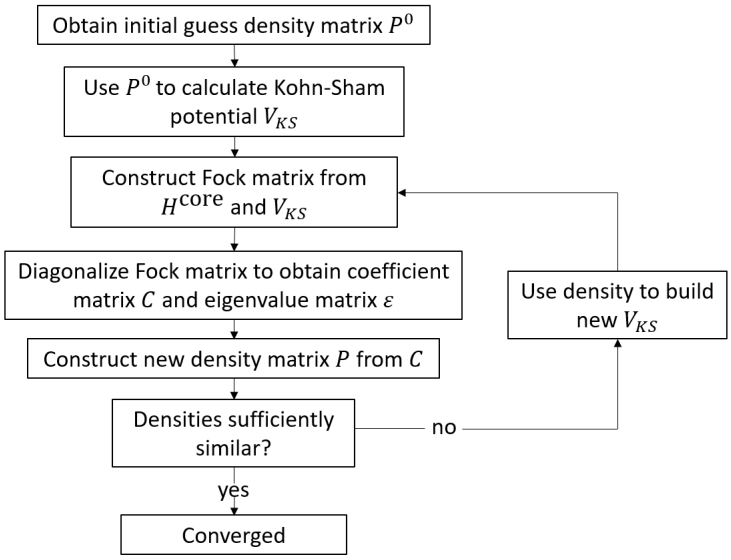

Figure 2.2 The Kohn-Sham DFT self-consistent field procedure.

which is used to build the Fock matrix; diagonalizing the matrix and solving for the molecular orbitals is the same as in HF, and the SCF procedure continues until convergence is reached.

With this understanding of the Kohn-Sham DFT Fock operator, we see that the α and β mixed and triplet determinant Fock operators for the ROKS Fock matrix will be nearly identical to the ROHF Fock operators, but without the exchange operator and plus the exchange-correlation potential.[56] That is,

$$F_\sigma^\omega = h + 2\sum_k^{N^c} J_k + J_a + J_b + V_{\text{XC}}^{\omega,\sigma} \tag{2.68}$$

where σ is either the mixed or triplet determinant and ω is either the α or β spin. The closed and open-shell Fock operators are made of the same linear combinations of these four Fock operators, so the rest of the ROKS Fock matrix formulation is identical to ROHF.[57]

Now we are ready to advance to the approximations of the DFT energy that give rise to the density-functional tight-binding method.

2.4 Density-functional tight-binding

To understand the fundamentals of DFTB, we'll start with the Kohn-Sham DFT energy expression written in the form

$$E[\rho] = \sum_i^{N/2} f_i \left\langle \psi_i \right| -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \frac{1}{2}\int \frac{\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r'}\left|\psi_i\right\rangle + E_{\text{XC}}[\rho] + E_{nn} \tag{2.69}$$

where f_i is the electron occupancy of the orbital (between 0 and 2 for a spin-restricted calculation), and E_{nn} is the internuclear repulsion term. Now we want to decompose the density $\rho(\mathbf{r})$ into an approximate reference $\rho_0(\mathbf{r})$ plus a small fluctuation $\delta\rho(\mathbf{r})$ which makes up the difference between the true density $\rho(\mathbf{r})$ and the approximate density. We can exchange $\rho(\mathbf{r})$ for $\rho_0(\mathbf{r}) + \delta\rho(\mathbf{r})$ to rewrite the energy as

$$\begin{aligned} E[\rho_0+\delta\rho] = &\sum_i^{N/2} f_i \left\langle \psi_i \right| -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r})\left|\psi_i\right\rangle + \frac{1}{2}\int\int \frac{(\rho_0+\delta\rho)(\rho_0'+\delta\rho')}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r'}\mathrm{d}\mathbf{r} \\ &+ E_{\text{XC}}[\rho_0+\delta\rho] + E_{nn} \end{aligned} \tag{2.70}$$

Note that, for instance, $\sum_i f_i \left\langle \psi_i \right| V_{\text{XC}}[\rho_0] \left|\psi_i\right\rangle = \int V_{\text{XC}}[\rho_0](\rho_0+\delta\rho)\mathrm{d}\mathbf{r}$, so terms can be moved between the Dirac notation and the integral notation. Additionally, the operator inside the sum will turn out to be the reference Hamiltonian for DFTB, H^0, so we will rearrange to add in the (double-counted) Coulomb term and the exchange-correlation potential for density ρ_0, and then we will correct for these terms with subtractions outside the sum:

$$\begin{aligned} E[\rho_0+\delta\rho] = &\sum_i^{N/2} f_i \left\langle \psi_i \right| -\frac{1}{2}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r'} + V_{\text{XC}}[\rho_0]\left|\psi_i\right\rangle \\ &- \frac{1}{2}\int\int \frac{\rho_0'(\rho_0+\delta\rho)}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} - \int V_{\text{XC}}[\rho_0](\rho_0+\delta\rho)\mathrm{d}\mathbf{r} \\ &+ \frac{1}{2}\int\int \frac{\delta\rho'(\rho_0+\delta\rho)}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} + E_{\text{XC}}[\rho+\delta\rho] + E_{nn} \end{aligned} \tag{2.71}$$

Rewriting E_{XC} for the corrected density with a third-order Taylor expansion,

$$\begin{aligned} E_{\text{XC}}[\rho_0+\delta\rho] = &E_{\text{XC}}[\rho_0] + \int \left.\frac{\delta E_{\text{XC}}}{\delta\rho}\right|_{\rho_0}\delta\rho\mathrm{d}\mathbf{r} + \frac{1}{2}\int\int \left.\frac{\delta^2 E_{\text{XC}}}{\delta\rho\delta\rho'}\right|_{\rho_0}\delta\rho\delta\rho'\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} \\ &+ \frac{1}{6}\int\int\int \left.\frac{\delta^3 E_{\text{XC}}}{\delta\rho\delta\rho'\delta\rho''}\right|_{\rho_0}\delta\rho\delta\rho'\delta\rho''\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'}\mathrm{d}\mathbf{r''} \end{aligned} \tag{2.72}$$

and using $(\delta E_{\mathrm{XC}}/\delta\rho)_{\rho_0} = V_{\mathrm{XC}}[\rho_0]$, we can substitute this into the energy expression and rearrange to cancel the first-order fluctuations:

$$\begin{aligned} E =& \sum_i^{N/2} f_i \left\langle \psi_i \right| -\frac{1}{2}\nabla^2 + V_{\mathrm{ext}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r'})}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r'} + V_{\mathrm{XC}}[\rho_0] \left| \psi_i \right\rangle \\ &- \frac{1}{2}\int\int \frac{\rho_0'\rho_0}{|\mathbf{r}-\mathbf{r'}|}\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} + E_{\mathrm{XC}}[\rho_0] - \int V_{\mathrm{XC}}[\rho_0]\rho_0 \mathrm{d}\mathbf{r} + E_{nn} \\ &+ \frac{1}{2}\int\int \left(\frac{\delta\rho\delta\rho'}{|\mathbf{r}-\mathbf{r'}|} + \left.\frac{\delta^2 E_{\mathrm{XC}}}{\delta\rho\delta\rho'}\right|_{\rho_0} \right) \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} \\ &+ \frac{1}{6}\int\int\int \left.\frac{\delta^3 E_{\mathrm{XC}}}{\delta\rho\delta\rho'\delta\rho''}\right|_{\rho_0} \delta\rho\delta\rho'\delta\rho''\mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'}\mathrm{d}\mathbf{r''} \end{aligned} \tag{2.73}$$

The first line of Eq. 2.73 is called the band structure energy, and is a sum of the orbital energies using only the reference Hamiltonian. The second line is the repulsive energy term, which contains the internuclear repulsion in addition to Coulomb interactions and exchange-correlation effects. The third term is the second-order energy correction, which accounts for charge fluctuations, and the fourth is the third-order correction, which describes changes in chemical hardness. We can represent the total third-order DFTB energy equivalently as

$$\begin{aligned} E_{\mathrm{DFTB3}} =& \sum_i^{N/2} f_i \left\langle \psi_i \right| H^0 \left| \psi_i \right\rangle + E_{\mathrm{rep}}[\rho_0] + E_{\mathrm{2nd}}[\rho_0, \delta\rho] + E_{\mathrm{3rd}}[\rho_0, \delta\rho, \delta^2\rho] \\ =& E_{\mathrm{band}}[\rho_0] + E_{\mathrm{rep}}[\rho_0] + E_{\mathrm{2nd}}[\rho_0, \delta\rho] + E_{\mathrm{3rd}}[\rho_0, \delta\rho, \delta^2\rho] \end{aligned} \tag{2.74}$$

DFTB1 neglects both second- and third-order terms, while DFTB2 omits only the third-order energy. Since the band energy and repulsive energy rely solely on the reference density, the iterative SCC procedure is only needed in DFTB2 and DFTB3.

With the familiar basis set expansion we can write

$$\psi_i = \sum_\mu^M c_{\mu i}\phi_\mu \tag{2.75}$$

where only a minimal basis set is used, so that molecular orbitals are formed of linear combinations of atomic orbitals (LCAO). The basis functions give rise to the secular equations

$$\sum_\nu^M \left(H^0_{\mu\nu} - \varepsilon_i S_{\mu\nu} \right) c_{\nu i} = 0 \tag{2.76}$$

where the overlap, coefficient, and eigenvalue matrices are analogous to Kohn-Sham and Hartree-Fock, and the Fock matrix is now referred to as the Hamiltonian.

Before solving the DFTB Kohn-Sham equations, we make a few more approximations. First, the basis functions are restricted to the valence shell, which means that core orbitals should be orthogonalized between atoms. This orthogonalization essentially creates a pseudopotential which reduces the Kohn-Sham potential in the reference Hamiltonian:

$$H^0_{\mu\nu} = \langle\phi_\mu| \left[-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}} - \sum_{\kappa_\alpha} |\phi^\alpha_\kappa\rangle \, \varepsilon_{\kappa_\alpha} \, \langle\phi^\alpha_\kappa| \right] |\phi_\nu\rangle \tag{2.77}$$

(Recall that where V_{KS} is another way to write the Coulomb interactions, external potential, and exchange correlation potential in the first line of Eq. 2.73.) However, this pseudopotential only applies to terms where μ and ν are not on atom α, so it mostly appears on three-center terms. In DFTB we make the two-center approximation, and ignore the pseudopotential effect when μ and ν are on the same atom. We additionally write the Kohn-Sham potential as a sum of single-atom potentials, so the Hamiltonian matrix elements are

$$H^0_{\mu\nu} = \langle\phi_\mu| -\frac{1}{2}\nabla^2 + V^\alpha_{\mathrm{KS}} + V^\beta_{\mathrm{KS}}|\phi_\nu\rangle \,, \quad \mu \in \alpha, \; \nu \in \beta \tag{2.78}$$

While this approximation is the same as the non-orthogonal tight-binding methods and extended Hückel theory,[33] DFTB has the advantage of basis sets and Hamiltonian elements derived from DFT calculations. In DFT the basis sets are obtained by solving the Kohn-Sham equations for each atom, with the modification of an additional confinement potential to compress the radius of the orbital and account for the fact that the free atomic orbitals of a free atom will not represent bound electrons well.

$$\left[-\frac{1}{2}\nabla^2 + V_{\mathrm{KS}} + \left(\frac{r}{r^0}\right)^2 \right] \phi_\mu = \varepsilon_\mu \phi_\mu \tag{2.79}$$

Here r^0 is the confinement radius, which is typically around twice the covalent radius of the atom. Within the Kohn-Sham potential in the reference Hamiltonian are also the neutral atom potentials which are necessary for the Coulomb interaction energy; these are obtained from DFT in the same way as the basis set.

With the (pseudo)atomic orbitals and densities parameterized from DFT, the matrix elements in DFTB are obtained using

$$H^0_{\mu\nu} = \begin{cases} \varepsilon^{\text{free}}_{\mu}, & \mu = \nu \\ \langle \phi_\mu | -\frac{1}{2}\nabla^2 + V^{\alpha}_{\text{KS}} + V^{\beta}_{\text{KS}} | \phi_\nu \rangle \,, & \mu \in \alpha,\ \nu \in \beta,\ \alpha \neq \beta \\ 0, & \text{otherwise} \end{cases} \tag{2.80}$$

In DFTB1, all that is needed to calculate the total energy is the band energy from the Hamiltonian and the repulsive energy. E_{rep} depends only on the reference density ρ_0, which is a superposition of the atomic densities:

$$\rho_0(\mathbf{r}) = \sum_{\alpha}^{M} \rho_0^{\alpha}(\mathbf{r}_\alpha) \tag{2.81}$$

where $\mathbf{r}_\alpha$ is the distance between $\mathbf{r}$ and the atom α nucleus. Ignoring interactions between 3 or more atoms (these are small and the densities are neutral, so are easily cancelled by nuclei), the repulsive energy can be written

$$E_{\text{rep}} = \sum_{\alpha}^{M} E_{\text{rep}}[\rho_0^{\alpha}] + \frac{1}{2}\sum_{\alpha}^{M}\sum_{\beta\neq\alpha}^{M} \left(E_{\text{rep}}[\rho_0^{\alpha} + \rho_0^{\beta}] - E_{\text{rep}}[\rho_0^{\alpha}] - E_{\text{rep}}[\rho_0^{\beta}] \right) \tag{2.82}$$

But rather than calculating E_{rep} values for known atomic densities, the entire functional is just fitted to the difference between the DFT energy and the band energy as a function of interatomic distance. This leaves DFTB1 independent of the need for a self-consistent procedure, but also incapable of accounting for charge fluctuations between atoms. To account for these, we need the second-order energy term.

By rewriting the density fluctuation $\delta\rho$ as a sum of atomic density fluctuations (the same way as the overall density), we can expand the atomic density fluctuations with a multipole expansion. Then, by keeping only the monopole term, we have

$$\delta_\alpha = \Delta q_\alpha F^{\alpha}_{00} Y_{00} \tag{2.83}$$

The Mulliken charge Δq_α is the difference between the electron population on atom α and the number of valence electrons on the neutral atom, i.e. $\Delta q_\alpha = q_\alpha - Z_\alpha$. The Mulliken

population q_α is determined in the same way as in Hartree-Fock, so

$$q_\alpha = \frac{1}{2}\sum_i^N f_i \sum_{\mu\in\alpha}^M \sum_\nu^M (C_{i\mu}C_{i\nu}S_{\mu\nu} + C_{i\nu}C_{i\mu}S_{\nu\mu}) \tag{2.84}$$

The functions F_{00}^α and Y_{00} create a spherical radial dependence for the charge fluctuation, so the second-order energy becomes

$$\begin{aligned} E_{\text{2nd}} =& \frac{1}{2}\sum_{\alpha,\beta}^M \Delta q_\alpha \Delta q_\beta \int\int \left(\frac{1}{|\mathbf{r}-\mathbf{r'}|} + \left.\frac{\delta^2 E_{\text{XC}}}{\delta\rho\delta\rho'}\right|_{\rho_0} \right) F_{00}^\alpha F_{00}^\beta Y_{00}^2 \mathrm{d}\mathbf{r}\mathrm{d}\mathbf{r'} \\ =& \frac{1}{2}\sum_{\alpha,\beta}^M \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} \end{aligned} \tag{2.85}$$

The function $\gamma_{\alpha\beta}$ is then parameterized in DFT such that when $\alpha = \beta$ it describes the electron-electron interaction within atom α, and so $\gamma_{\alpha\alpha}$ becomes the chemical hardness, or Hubbard parameter U_α. It is a function of Hubbard parameters for both atoms and of the interatomic distance.

Now the DFTB2 energy is

$$E_{\text{DFTB2}} = \sum_i^N \langle\psi_i|H^0|\psi_i\rangle + \frac{1}{2}\sum_{\alpha\beta}^M \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta}(U_\alpha, U_\beta, R_{\alpha\beta}) + E_{\text{rep}} \tag{2.86}$$

and the Hamiltonian relies on the Mulliken charges, according to

$$H_{\mu\nu} = \langle\phi_\mu|H^0|\phi_\nu\rangle + \frac{1}{2}S_{\mu\nu}\sum_\eta^M(\gamma_{\alpha\eta} + \gamma_{\beta\eta})\Delta q_\eta, \quad \mu\in\alpha,\ \nu\in\beta \tag{2.87}$$

Since the Mulliken charges are obtained from the coefficient matrix, and equivalently the wavefunctions ψ_i, DFTB2 must now follow a self-consistent procedure. Since convergence is based on the charges, this is called the self-consistent charge (SCC) procedure, which is outlined in Figure 2.3.

The third-order energy correction for DFTB3 is parameterized from DFT in much the same way, introducing the derivative of the Hubbard parameter and resulting in the function $\Gamma_{\alpha\beta}$, the derivative of the γ function with respect to charge.

$$E_{\text{DFTB3}} = \sum_i^N \langle\psi_i|H^0|\psi_i\rangle + \frac{1}{2}\sum_{\alpha\beta}^M \Delta q_\alpha \Delta q_\beta \gamma_{\alpha\beta} + \frac{1}{3}\sum_{\alpha\beta} \Delta q_\alpha^2 \Delta q_\beta \Gamma_{\alpha\beta} + E_{\text{rep}} \tag{2.88}$$

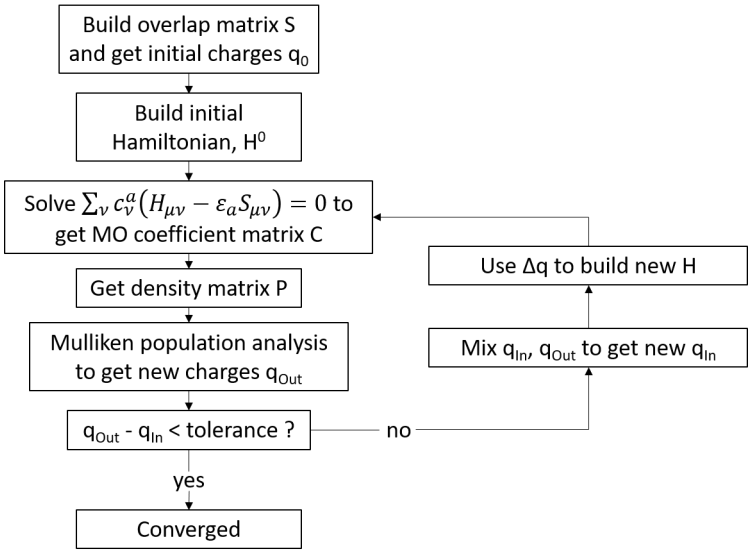

Figure 2.3 The DFTB2 self-consistent charge procedure

For a more detailed discussion of DFTB3 see, for example, Refs. [58] and [59]. This work focuses on DFTB2 only; third-order corrections should be compatible with the RO-DFTB implementation, future testing will need to confirm this.

Before examining the results of RO-DFTB testing at the DFTB2 level, we will outline the major points of the current implementation.

Chapter Three

METHODS

3.1 The RO-DFTB routine

This development of RO-DFTB takes place in the open-source DFTB+ software package, version 19.1.[44] Some aspects of ΔDFTB were employed in building the non-Aufbau mixed and triplet determinants, particularly the routine for adjusting the electron filling of the orbitals.

In addition to building the RO Hamiltonian, the successful RO-DFTB routine requires adjusting the SCC procedure to accommodate steps where both determinants are treated side-by-side, such as building the separate mixed and triplet Hamiltonians and handling separate mixed and triplet charge vectors. Furthermore, the updated energy-weighted density matrix must be incorporated for calculation of the gradients for geometry optimization. The current RO-DFTB routine is depicted in Figure 3.1, which shows both the SCC and geometry optimization iterative procedures.

To focus first on the SCC routine, one of the main features of the RO-DFTB procedure is the separation of the first SCC iteration from subsequent iterations. This iteration serves as an initial guess at the ground state of the system, using the initial charges to build a ground state Hamiltonian. Being derived solely from the electron populations of the neutral atoms plus any overall charge of the system, these initial charges cannot be separated into

the mixed and triplet determinant charges without an orbital structure or electron densities. The easiest way to obtain mixed and triplet determinant charges is by diagonalizing the ground state Hamiltonian and then performing the mixed and triplet non-Aufbau electron fillings, followed by the Mulliken population analysis on each electron density. So, the first SCC iteration serves to build the mixed and triplet charges.

Subsequent SCC iterations build mixed and triplet unrestricted Hamiltonians, use them to create the RO Hamiltonian, and obtain new mixed and triplet densities and charges by diagonalizing the Hamiltonian and applying the two non-Aufbau fillings. In this way the only truly restricted part of the RO-DFTB routine is the Hamiltonian matrix diagonalization; in all other areas unrestricted mixed and triplet duplicates of most matrices exist. The duplicate charges in particular create an ambiguity as to how best to to mix the two charge vectors for convergence acceleration. In the typical DFTB SCC routine, charges are mixed at the end of an iteration and used to start the next iteration. But mixing the two RO-DFTB charge vectors separately was unsatisfactory given their interdependence via the RO Hamiltonian. Nor could they be summed into one vector, because separate charges are needed at the start of the SCC step for generating the Hamiltonians. Instead, mixed and triplet output charges are appended into a single vector and mixed together; then they are separated for the beginning of the next SCC iteration.

However, even this procedure for updating charges has a potential issue: at the beginning of SCC iteration n, for example, the input charges are a result of mixing the charges of iteration $n-1$ with some number of charges from earlier iterations. These input charges are used to create the iteration n mixed and triplet Hamiltonians, but in building the RO Hamiltonian the density matrices are also required to project the coupling operators to the MO basis. However, the density matrices have not been mixed, and correspond to iteration $n-1$, which means that they do not correspond to the charges of the current iteration or to the Hamiltonians derived from those charges. As shown in chapter 4, the performance of this routine was poor enough and the severity of this discrepancy seemed substantial enough

to pursue a different method of convergence acceleration.

The alternative to charge mixing that was developed uses the same charge mixer algorithms to instead update the RO Hamiltonian. Since DFTB+ uses sparse matrix storage, the Hamiltonian is readily obtained as a vector, and can be mixed with previous iterations in exactly the same way as the charges. This does change the organization of the SCC routine, though. The mixed and triplet charges are now passed directly from one iteration to the next, and the RO Hamiltonian is combined with its previous iteration(s) in the middle of the SCC procedure, just before diagonalization.

Typically in DFTB+, convergence is determined before mixing the charges: if new and old charges are similar within the set tolerance, mixing is not required and the SCC routine has converged. In RO-DFTB, convergence can be determined in the same way with either charge mixing or Hamiltonian mixing, and this is the method used most extensively in development and testing. Determining convergence based on changes in the Hamiltonian instead offers no advantage, since the Hamiltonian is built from the charges. However, the RO Hamiltonian matrix does give another option for determining convergence, since the off-diagonal error matrices should go to zero. While building the RO Hamiltonian, the maximum value in the error matrices is obtained which can be used in place of the maximum charge difference; thus Hamiltonian convergence is compatible with either charge or Hamiltonian mixing.

While the Hamiltonian mixing method seems to be free of any inconsistencies between charges and densities, care does have to be taken when calculations require the mixed and triplet determinant Hamiltonians after the RO Hamiltonian mixing procedure. In this case, unrestricted Hamiltonians from the beginning of the iteration should not be used. Instead, new mixed and triplet Hamiltonians should be built from the charges obtained after the updated RO Hamiltonian is diagonalized. This is only an issue after SCC convergence during a geometry optimization calculation, when the energy-weighted density matrix requires Hamiltonians from the separate determinants.

Geometry optimization in the RO-DFTB routine has its own ambiguities in the treat-

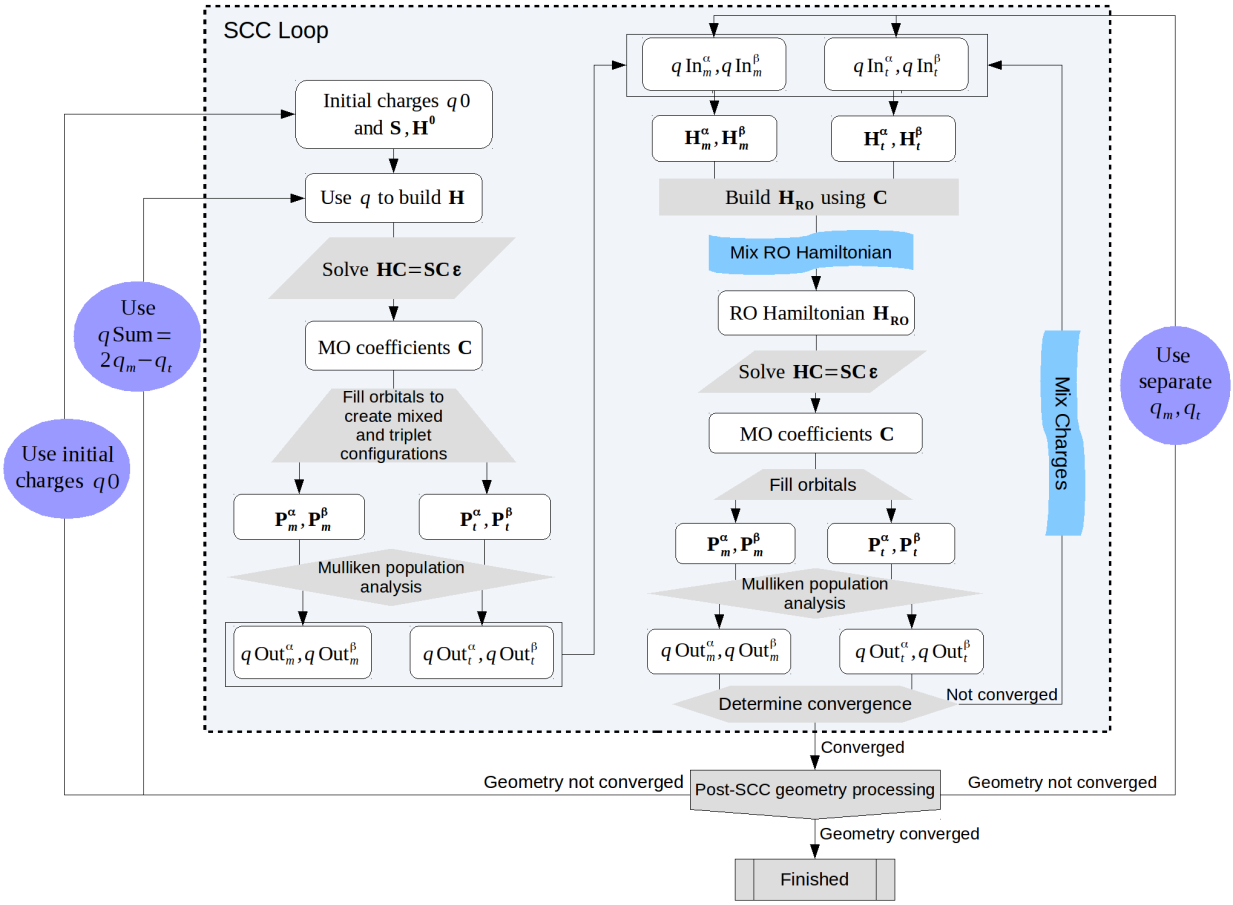

Figure 3.1 Flowchart of an RO-DFTB geometry optimization calculation. The SCC loop contains the first iteration to the left of the main loop, and the options for mixing either the RO Hamiltonian or the charges are shown in the main SCC loop. The geometry loop, outside of the SCC box, shows the three charge input options for beginning the next geometry step.

ment of the mixed and triplet charges: how should these charges be passed to the next geometry step? Ground state calculations and ΔDFTB use the output charges of the previous geometry step to begin new steps, but this means that geometry steps after the first will be fundamentally different from a single point calculation at the same geometry, in that they will not begin with a guess of the ground state Hamiltonian. It's also not clear whether keeping the mixed and triplet charges separate imposes unnecessary restrictions, since they are not being optimized individually. For these reasons, two other charge input methods for geometry optimization are available in RO-DFTB as well, both of which use a single charge vector and build a non-RO Hamiltonian on the first SCC iteration of all geometry steps. The first uses a single vector of the Ziegler sum of the mixed and triplet charges, while the second method resets the charges at each geometry step to the charges derived from neutral atoms which are used for the initial geometry.

Combining these three charge input methods with the two mixing methods yields six different RO-DFTB routines. These are abbreviated with "Ham" or "charge" for the mixing method, and "qMT" for separate (qMixed and qTriplet) geometry input charges, "qSum" for Ziegler summed charges, and "q0" for reset charges.

Level shifting has also been implemented in RO-DFTB, which allows each shell to be raised by any energy amount. While this does enhance the convergence and overall performance of RO-DFTB, it does add an additional layer of complexity for testing the six routines. A full breakdown of RO-DFTB testing is provided in the next section.

3.2 RO-DFTB testing

In addition to the six baseline routines, three degrees of level shifting and three different convergence acceleration mixers are used to evaluate RO-DFTB. The first degree of level shifting, "noLS," does not shift any shells, while the second, "virtLS," shifts the virtual shell by 0.1 hartree (H), and the third, "LS," adds 0.1 H and 0.2 H to the first and second open

shells, respectively, and 0.3 H to the virtual shell. The mixing algorithms we have used are the Broyden mixer,[60] the DIIS mixer,[61] and simple mixer.

Typically, level shifting is only incorporated when a calculation does not converge without it. However, we have repeated successful calculations for each degree of level shifting to better understand its effects and its usefulness in RO-DFTB. Similarly, we have distinguished between the different mixers more than usual to examine the performance of the RO-DFTB routines in more detail. Additionally, we have fixed the maximum number of SCC steps to 500 for both Broyden and DIIS mixers, and 1000 for the simple mixer. While these maximums are generous, it is possible that in some cases convergence could be reached by further increasing this cutoff.

While most testing has used the method of convergence based on charges, some additional testing has been done with Hamiltonian-based convergence. Where not otherwise specified, charge-based convergence is used.

Finally, to evaluate the geometry optimization of RO-DFTB, we have used numerical calculations of the energy gradients to compare against the analytical gradients computed in DFTB+. These numerical gradients require single-point energy calculations for each atom in the x-, y-, and z-coordinates, and we calculate energies on both sides of the true coordinate. That is, we use the symmetric difference quotient:

$$E'(x) = \lim_{\delta x \to 0} \frac{E(x + \delta x) - E(x - \delta x)}{2\delta x} \tag{3.1}$$

Chapter Four

RESULTS AND DISCUSSION

4.1 Evaluating RO-DFTB

We have multiple goals for the evaluation of the RO-DFTB method. First, we aim to either identify a single RO-DFTB routine which outperforms the others, or define a set of guidelines for choosing which RO-DFTB routine will be best suited to a particular purpose. Second, we aim to benchmark RO-DFTB against results from ΔDFTB, higher levels of theory (including ROKS), and experiment to assess its usefulness as an approximate excited-state method.

To this end, we evaluate RO-DFTB in the following sections first on four metrics: convergence success rate, compute time, vertical excitation energy, and Stokes shift. We then examine energy variability and geometry optimizations in more detail, including a comparison of the analytical and numerical gradients.

We have used four molecular test sets for evaluating RO-DFTB performance. Three of these sets offer reference values for vertical excitation energies, which can be used to evaluate RO-DFTB single-point energy calculations, and two offer reference values for Stokes shifts, which can be used as one measure of RO-DFTB energies after geometry optimization.

The first set contains 14 small molecules used extensively in method development: H_2, N_2, O_2^{2+}, O_2^{2-}, CO, NO^+, CN^-, OH^-, CO_2, ozone, water, acetone, methylcyclopropene (MCP), and toluene. O_2 was included initially, but has been left out because its ground state is a

triplet rather than a singlet. Despite their small size, some of these compounds have excited states that are rather difficult to model.

The second set contains 28 organic dyes from the benchmarking set designed by Thiel *et al.* specifically for excited states in Ref. [62]. Third is a set of π-conjugated organic molecules used to benchmark the restricted ensemble-referenced Kohn-Sham (REKS)[63] method in Ref. [64]. The fourth set is shown in Fig. 4.1 and consists of six large organic dyes from a subset of the collection used to benchmark ROKS in Ref. [65]. Three of the original nine dyes are excluded because the mio parameter set for DFTB2 used in these benchmarks does not include fluorine or zinc atoms.[31,66]

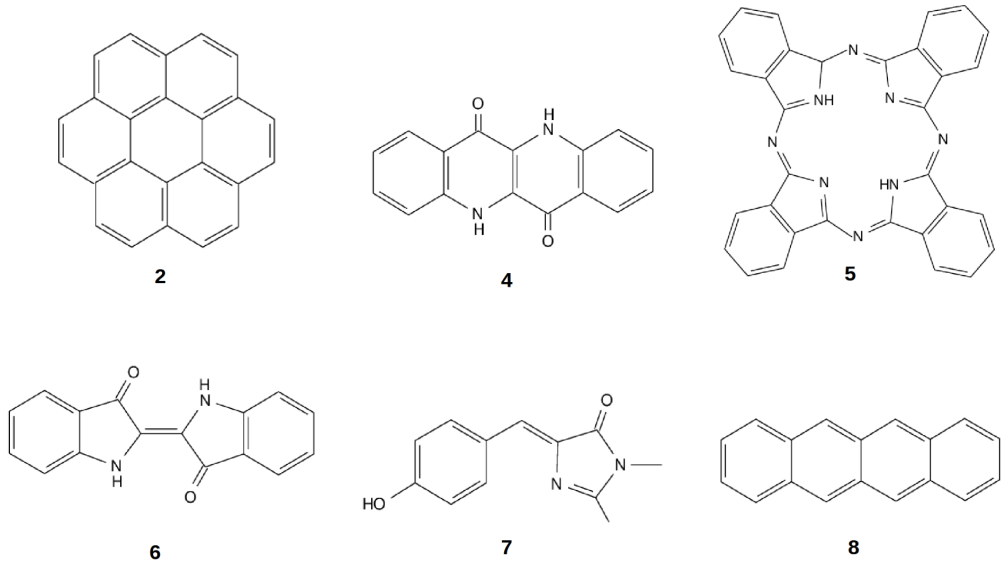

Figure 4.1 Six large organic dyes from a subset of the collection used in Ref. [65]. Compounds **1**, **3**, and **9** included fluorine or zinc and were excluded since these atoms are not part of the mio parameter set used for DFTB2.

Altogether, the four sets comprise 56 molecules, two of which are duplicates that appear in more than one test set. When the sets are evaluated individually these molecules, acetone and butadiene, appear in both sets. For all averaging across test sets, they are counted only once for a total of 54 compounds.

4.2 RO-DFTB convergence

The convergence success of RO-DFTB single-point calculations is shown in Figure 4.2 for 18 RO-DFTB versions: these combine the six mixing and charge input routines with the three different mixers. Of course, the different charge input methods should have no effect before the molecule begins the second geometry step. Indeed, the Hamiltonian mixing routines are independent of charge input; but charge mixing shows slightly different convergence with the qMT charge input. Variations such as these will be discussed in more detail in section 4.6.

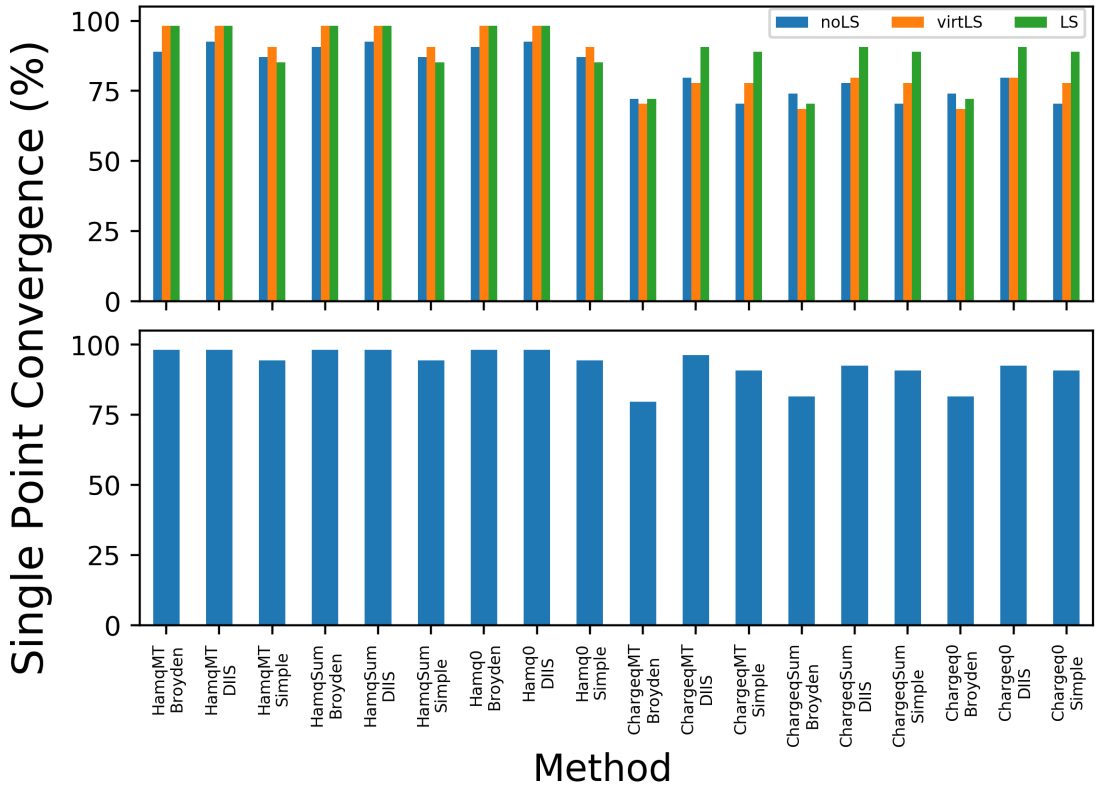

Figure 4.2 RO-DFTB convergence by version for single-point calculations at the optimized ground state geometry. (Top) The different degrees of level shifting are shown separately, and (bottom) level shifting is included when the single-point calculation does not converge.

With the Hamiltonian mixing methods, the Broyden mixer converges for 48 out of 54 molecules, while the DIIS mixer converges for 50. When just virtual level shifting is added,

both mixers converge for all but one molecule (large dye **5**). Convergence is significantly lower for the charge mixing methods, although level shifting increases DIIS convergence from 42 compounds to 49. Interestingly though, level shifting offers no advantage with the Broyden mixer. The Broyden mixer's overall poor performance for charge mixing methods is odd, especially since the Broyden mixer is generally the most robust for ΔDFTB calculations.

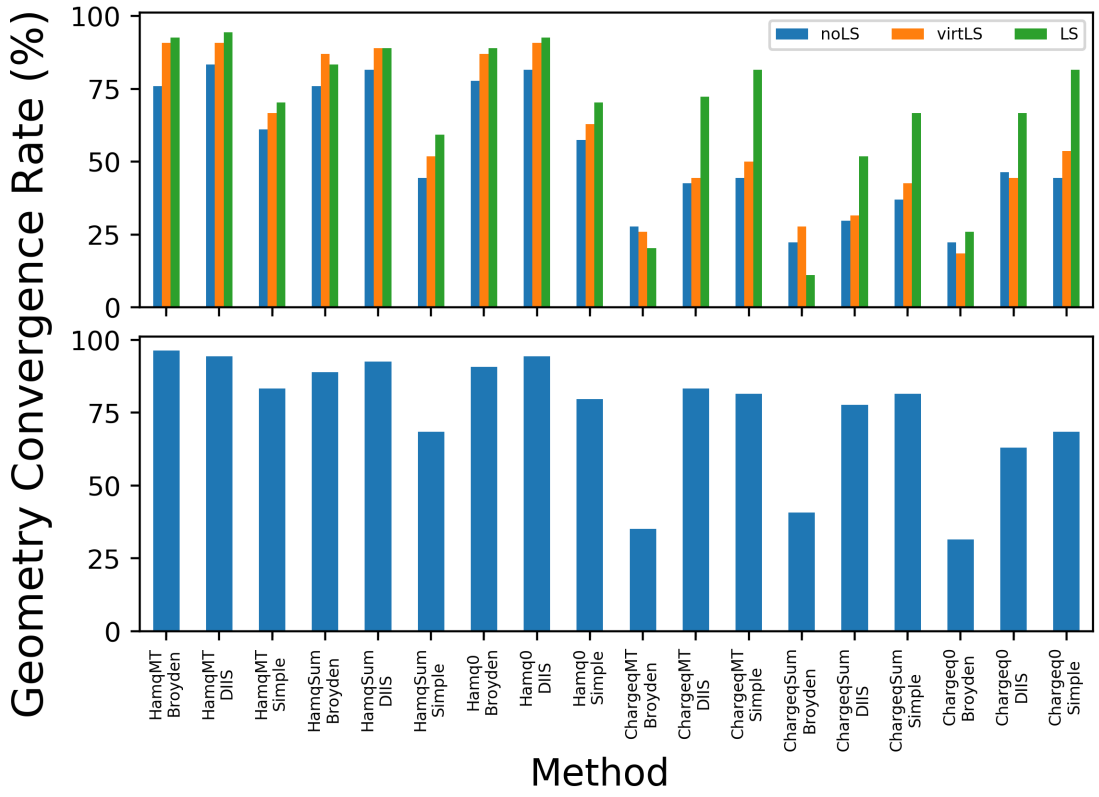

Figure 4.3 Geometry optimization convergence of RO-DFTB by method across all data sets. (Top) with different level shifting options shown explicitly, and (bottom) with level shifting included only when the optimization is not reached without it.

The performance differences between Hamiltonian and charge mixing methods are even more stark after geometry optimization. Figure 4.3 shows that the Broyden mixer is still significantly worse than either DIIS or simple mixers for charge mixing methods, while it is comparable to the DIIS mixer with Hamiltonian mixing. As with single-point calculations, level shifting still improves convergence for all versions except the Broyden mixer with charge

mixing. The advantages of level shifting are especially great for the other charge mixing versions, where convergence is below 50 % without it.

Notably, the HamqMTBroyden version (Hamiltonian mixing, qMT charge inputs, and the Broyden mixer) performs the best: only the oxygen anion and large dye **5** fail to converge, and these compounds fail for all of the RO-DFTB versions. The HamqMTDIIS and Hamq0DIIS versions each fail for one additional molecule: methyliminium and cytosine, respectively. With charge mixing, the best version is chargeqMTDIIS, which fails to converge for 9 of the 54 compounds.

The convergence rate of the Hamiltonian mixing versions is even more impressive when compared to ΔDFTB. The geometry optimization of ΔDFTB was most successful with the Broyden mixer, where it failed to converge for 8 of the 54 compounds. However, it should be noted that ΔDFTB calculations did not attempt to utilize the maximum overlap method (MOM) or similar variations[67] in order to improve convergence.

A more thorough picture of RO-DFTB convergence can be gained by examining method convergence by molecule. In Figure 4.4, the number of mixers converged is shown for each molecule with each of the three charge input methods, before and after level shifting is added. Without level shifting, nearly half of the test molecules fail to converge with any mixer for any charge input method. Those that do converge are mainly the neutral or positively charged diatomics and acetone from the small molecule set, a few of the molecules in the REKS set, and around half of the Thiel set compounds. Interestingly, the nucleotides thymine and uracil both fail to converge with ΔDFTB, but converge across all charge input methods here. However, the charge mixing methods are particularly poor at converging molecules from the large dye set.

Even when level shifting is included, the large dye set remains challenging for the charge mixing methods. The small molecule set also sees limited convergence improvement, mainly for CN^-. But level shifting makes a huge difference for the Thiel set and REKS set compounds which had failed to converge. As for the charge inputs, the q0 and qMT methods

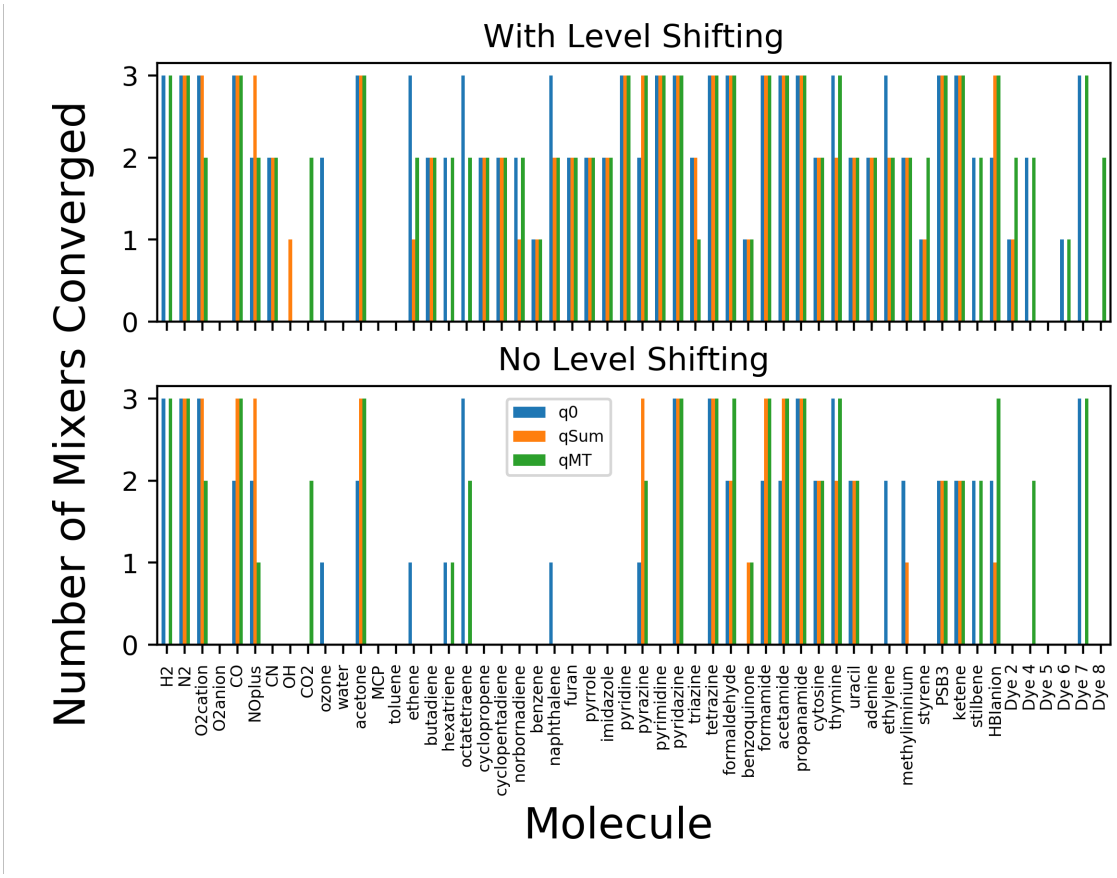

Figure 4.4 Differences in convergence between the charge mixing methods with and without level shifting. Where level shifting is used, it is only when convergence otherwise fails.

have the broadest success, but the qSum method is the only one which converges for OH^-, and also stands out for NO^+ and pyrazine. However, the qSum method never converges for H_2 and only converges with one mixer for dye **2**, while it fails for the other large dyes.

The qSum method is also noticeably different from the qMT and q0 methods for the Hamiltonian mixing routines, shown in Fig. 4.5. In general it performs slightly less consistently, except in the cases of water and methyliminium. It is interesting that there is no overlap between charge and Hamiltonian mixing in the compounds for which the qSum method outperforms qMT and q0 methods. This perhaps speaks to the degree of difference between the Hamiltonian and charge mixing routines. Across all four test sets, Hamiltonian mixing routines converge far more consistently than charge mixing both with and without

level shifting. However, the Hamiltonian routines struggle most with some of the compounds in the small molecule set, which mirrors the case for charge mixing. But Hamiltonian mixing routines converge well for almost all other compounds, including the dyes other than dye **5**.

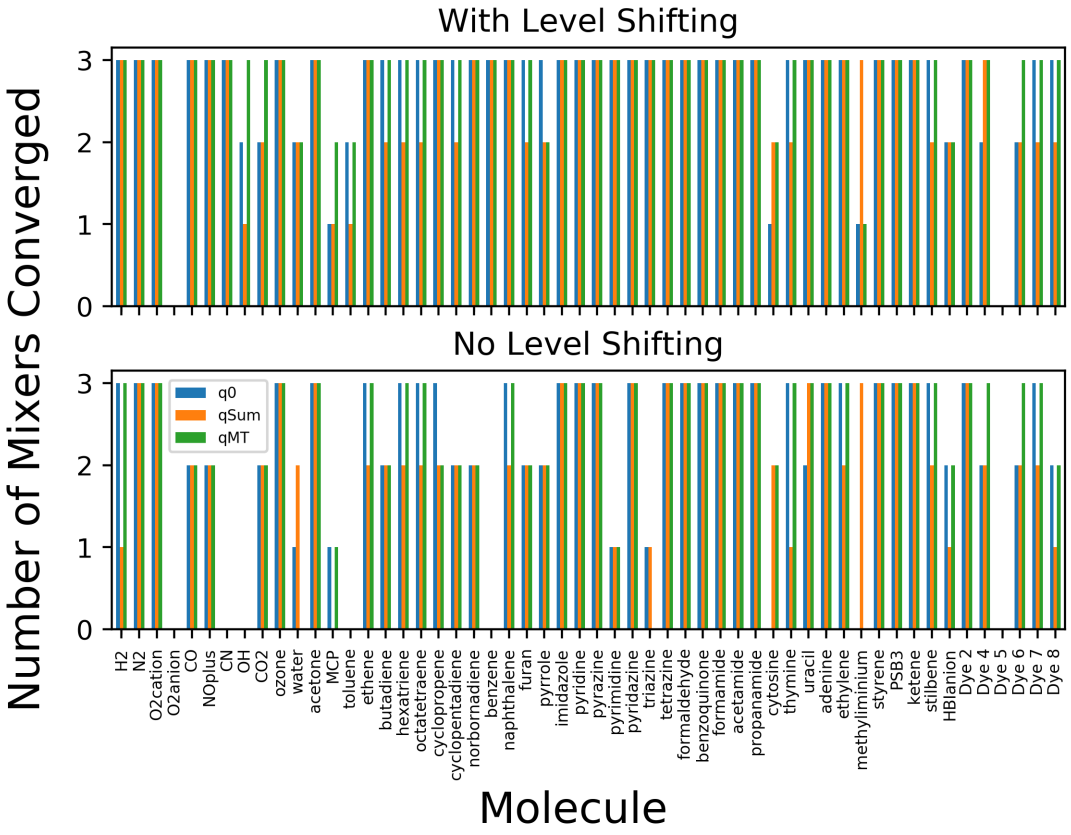

Figure 4.5 Differences in convergence between Hamiltonian mixing methods with and without level shifting. Level shifting is used only when convergence otherwise fails.

There is some similarity between the molecules that Hamiltonian mixing routines struggle to converge without level shifting and those which fail to converge for ΔDFTB, including benzene and triazine. But for other compounds, the Hamiltonian mixing routines perform quite differently from ΔDFTB. Where ΔDFTB fails for dye **2**, RO-DFTB does fine without level shifting. But ΔDFTB easily converges dye **5**, which RO-DFTB cannot even with level shifting. The one molecule which all methods fail to converge is the O_2^{2-} anion, and even ROKS/ωb97x-s/6-311+G* could not obtain a single-point energy (although level shifting

was not added for ROKS).

These differences in convergence between RO-DFTB and ΔDFTB suggest that RO-DFTB may be of some use for cases where ΔDFTB calculations fail. Additionally, the broad success of the Hamiltonian mixing routines indicate that these methods in particular may prove to be a worthy excited-state tool to complement ΔDFTB.

4.3 RO-DFTB compute time

With Hamiltonian mixing appearing superior to charge mixing based on single-point and geometry convergence, it also makes sense to compare these methods on the time it takes them to geometry optimize. While ultimately not as important as accuracy, a key advantage of DFTB is its resource efficiency, which we hope RO-DFTB can contribute to by offering some further time advantage over ΔDFTB.

There are some trends that we might expect between time complexities of the different RO-DFTB versions. First, the simple mixer is sure to be much slower than the Broyden and DIIS mixers, as it is a much less effective algorithm. Second, the q0 and qSum methods may be slightly slower than the qMT charge input method, since they require an additional SCC iteration at each geometry step. Third, we might anticipate that Hamiltonian mixing would take longer than charge mixing, since even in sparse form the Hamiltonian vector is much larger than the charge vectors.

But as shown in Figure 4.6, Hamiltonian mixing seems to offer significant time-savings over charge mixing. (Note that the timescale is quite small because the time has been normalized by the number of electrons in the molecule, to mitigate any bias against versions more capable of converging larger molecules.)

For Hamiltonian mixing, the qMT method is significantly faster than the other charge input methods, and this is consistent across mixers. The Broyden mixer also performs faster on average than the DIIS mixer for all charge input methods. This is in contrast to the charge mixing routines, where the DIIS mixer is faster than the Broyden mixer, especially for the qMT method and to a much lesser extent for the qSum method. While still curious, this does follow the trend of poor compatibility between charge mixing and the Broyden mixer.

While the distinctions between degrees of level shifting in Figure 4.6 aren't directly comparable (since the set of converged molecules that they are averaged over can be substantially

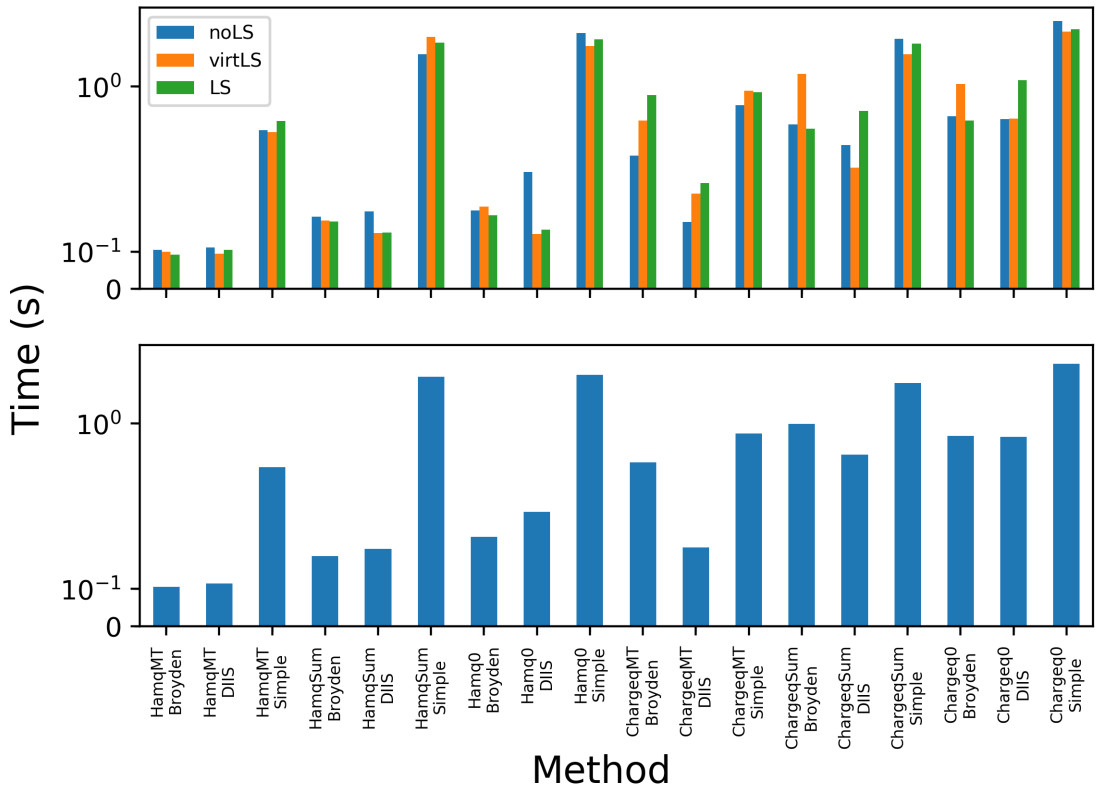

Figure 4.6 Geometry optimization compute time for each RO-DFTB version, normalized by the number of electrons and averaged over all molecules that converge, with a log scale to account for the large runtimes of the simple mixer. (Top) level shifting results are reported separately, and (bottom) level shifting is incorporated for compounds that do not converge without it.

different), the separate times do show that charge mixing is not slower than Hamiltonian mixing due to an increased reliance on level shifting. Even without level shifting the charge mixing methods are much slower. In fact, for RO-DFTB versions where convergence is largely similar with and without level shifting (the Hamiltonian mixing versions), the data suggest that level shifting might not increase the runtime at all, or even decrease it.

Although we would hope that RO-DFTB would take around half the time of a ΔDFTB calculation due to its single SCC procedure, the use of Hamiltonian mixing probably does decrease these savings. To evaluate both the compute times and the number of SCC steps

per geometry step, the HamqMTBroyden version of RO-DFTB (with supplementary level shifting) has been compared to ΔDFTB for each molecule that they both converge. Figure 4.7 shows the average SCC steps per geometry step, as well as both the time per electron and per SCC step—which is measured in only a few milliseconds. It is worth noting that these times represent only a single optimization, and the overhead required has not been accounted for.

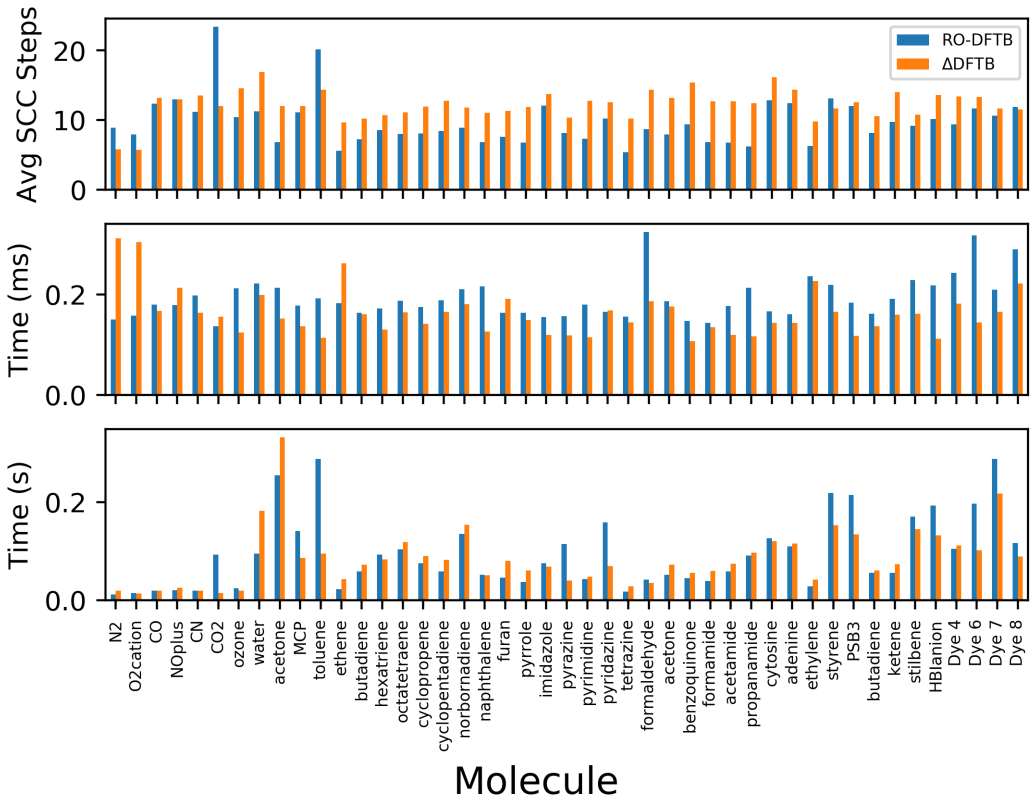

Figure 4.7 Comparison of the HamqMTBroyden RO-DFTB version with as-needed level shifting against ΔDFTB, also with the Broyden mixer. Only molecules where both versions converge are shown: (Top) the average number of SCC steps per geometry step; (middle) the time to geometry convergence normalized by the number of electrons and total SCC steps; and (bottom) the time to convergence normalized by number of electrons.

While RO-DFTB is faster than ΔDFTB in more than half of the test molecules where

they both converge, the time is only reduced modestly—and for several molecules RO-DFTB is significantly slower. These compounds are typically larger, although the trend is not universal. The smallest molecules, though, are the only ones where RO-DFTB takes less time on average per SCC step. This trend is not entirely unexpected due to the resource requirements of Hamiltonian mixing: the Hamiltonian is of comparable size to the charge vectors when the number of electrons is very low, but as the number of atoms increases, the Hamiltonian increases quadratically, while the charge vector increases linearly.

The reduction in the number of SCC steps with RO-DFTB is not quite half, but still does represent a significant decrease for most molecules. RO-DFTB does best for the Thiel set of small dyes, where the ratio of RO-DFTB SCC steps to ΔDFTB SCC steps is 0.67. In contrast, this ratio is 1.11 for the small molecule set, where the RO-DFTB time per SCC step is lower. For the REKS set and large dye set, the ratio of RO-DFTB:ΔDFTB SCC steps is 0.83 and 0.88, respectively.

This comparison of RO-DFTB and ΔDFTB runtimes does not take into account the additional time of running a second or third calculation with level shifting when convergence fails. This does worsen RO-DFTB time requirements, but the majority of compounds should not need level shifting, and indeed most of the test molecules do not with the HamqMT-Broyden version. It would also be quite simple to automate the addition of level shifting to molecules that fail to converge, and this would certainly make RO-DFTB easier to use for a screening study or machine learning application.

Some of the runtime differences between RO-DFTB and ΔDFTB may also be due to these two methods converging to significantly different geometries. After all, excited-state geometry optimization is not as simple as a ground state optimization. For all RO-DFTB and ΔDFTB calculations, the input geometry corresponds to the optimized ground state, which in some cases is significantly different than the singlet excited-state geometry. With this in mind, it means little if RO-DFTB is faster than ΔDFTB yet converges to a geometry that is much less accurate. Of course, the same is true for single-point energy calculations, so

we now turn to evaluating the accuracy of RO-DFTB. In the next section, we use reference data[62,64] and ROKS calculations for vertical excitation energies to evaluate the single-point accuracy of RO-DFTB at the ground state geometry.

4.4 Vertical excitation energies

The reference values for benchmarking RO-DFTB vertical excitation energies come from the small molecule set, Thiel set, and REKS set. Data for the small molecule set come from ROKS and ground state DFT calculations with the ωB97X-D exchange-correlation functional and the 6-311+G* basis set. (The ROKS calculations converged for all compounds in the set except for O_2^{2-}.) Reference values for the Thiel set are best estimates from Thiel *et al.*, Ref. [62]. Reference values for the REKS set are from state-interaction state-averaged REKS (SSR) with the LRC-ωPBEh functional and the 6-31G* basis set, as described in Ref. [64].

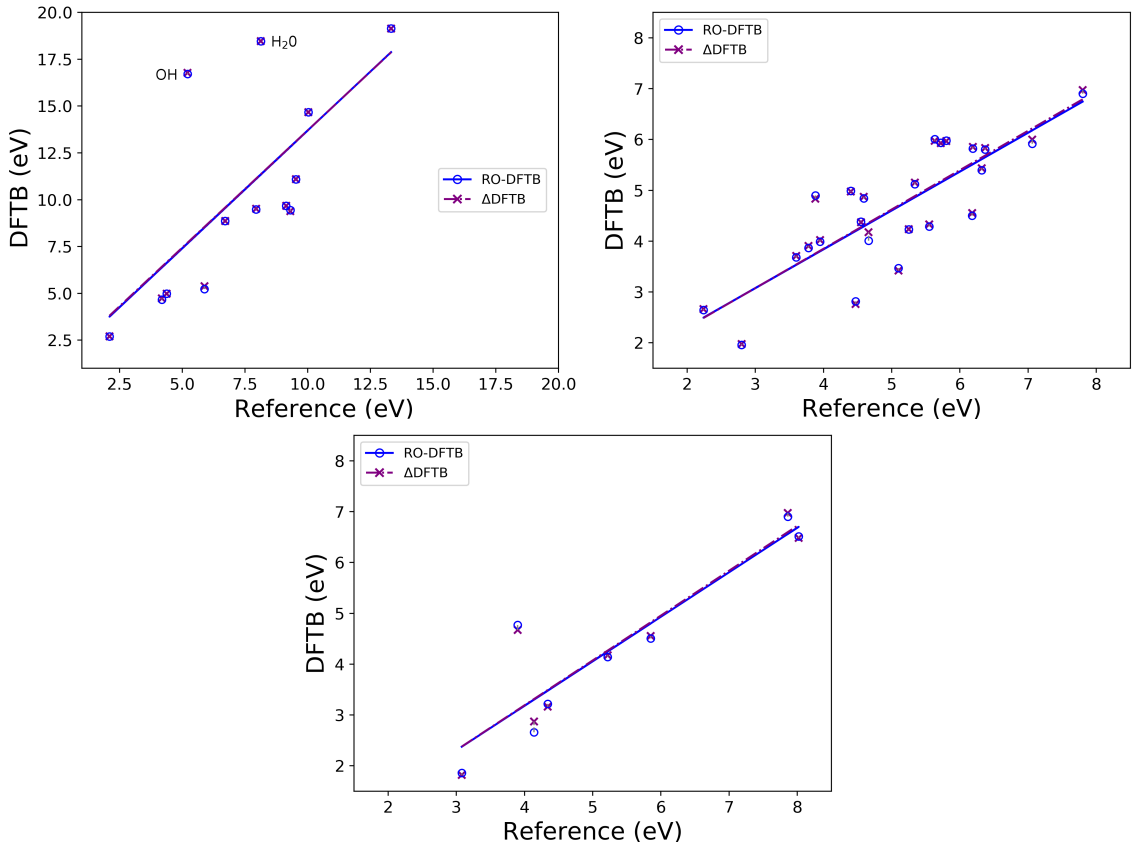

Figure 4.8 Vertical excitation energies of HamqMTBroyden RO-DFTB and ΔDFTB against reference values for the small molecule set (calculated with ROKS ωB97X-D/6-311+G*), Thiel set,[62] and REKS set.[64] Regression statistics are shown in Table 4.1.

Scatter plots for the HamqMTBroyden version of RO-DFTB with supplementary level

shifting are shown in Figure 4.8, where they are compared with ΔDFTB. The trends for the three separate test sets are distinct, but ΔDFTB and RO-DFTB are in extremely close agreement across them. Both tend to overestimate excitation energies for the small molecules, especially OH^- and water, which lie significantly apart from the linear regression. In contrast, nearly all of the REKS compounds are significantly underestimated by both RO-DFTB and ΔDFTB, with the exception of ketene (ethenone), which is overestimated by about the same amount. For the Thiel set, there is a general trend toward underestimation, although formaldehyde is overestimated by as much as 1 eV.

	Small Molecules			Thiel Set			REKS Set		
method	m	b	R^2	m	b	R^2	m	b	R^2
RO-DFTB	1.46	-1.19	0.72	0.76	0.78	0.68	0.88	-0.32	0.82
ΔDFTB	1.45	-1.13	0.72	0.77	0.75	0.69	0.88	-0.34	0.84

Table 4.1 Slope m and y-intercept b (in eV) and coefficient of determination for RO-DFTB and ΔDFTB vertical excitation energies.

The largest discrepancies between RO-DFTB and ΔDFTB are with toluene in the small molecules set (0.16 eV), cytosine in the Thiel set (0.17 eV) and PSB3 in the REKS set (0.21 eV); for all other compounds the difference is less than 0.1 eV. In each of these three cases RO-DFTB arrives at a lower energy than ΔDFTB, and in fact RO-DFTB undershoots the ΔDFTB energy 70 % of the time; this is opposite the case for ROKS, which typically gives a larger excitation energy than ΔSCF.

Despite the small differences over a few molecules, the regressions for RO-DFTB and ΔDFTB for each test set are practically identical, though RO-DFTB has slightly less favorable correlations with the reference excitation energies for the Thiel and REKS sets, as shown in Table 4.1. The errors between RO-DFTB and ΔDFTB are certainly much smaller than the errors of either method against accepted reference values, but these approximate methods can still be very useful when systematic errors are accounted for. Table 4.2 shows

that the larger RMSE values are accompanied by signed and unsigned errors which are very close, which shows the systematic over- and underestimation in the small molecule set and REKS set, respectively. For the Thiel set, where the error is not as systematic, the lower RMSE (which is even lower for RO-DFTB than ΔDFTB) still supports the usefulness of these versions.

These error statistics are averaged over only the molecules which converge with all three mixers and ΔDFTB, which shows a slight variation between HamqMT RO-DFTB mixers, although in theory these mixers should not change the energy at convergence. For the most part these inconsistencies are small (even negligible for the small molecule set), but they become more pronounced for sets including larger molecules.

	Small Molecules			Thiel Set			REKS Set		
method	MSE	MAE	RMSE	MSE	MAE	RMSE	MSE	MAE	RMSE
HamqMTBroyden	3.32	3.32	4.83	-0.40	0.67	0.84	-0.98	1.17	1.19
HamqMTDIIS	3.32	3.32	4.84	-0.41	0.68	0.85	-0.98	1.17	1.19
HamqMTSimple	3.32	3.32	4.84	-0.43	0.70	0.87	-1.04	1.23	1.25
ΔDFTB	3.33	3.33	4.85	-0.38	0.66	0.82	-0.96	1.13	1.15

Table 4.2 Vertical excitation energy errors (mean signed error, MSE; mean absolute error, MAE; and root mean squared error, RMSE) by test set for three RO-DFTB versions and ΔDFTB, reported in eV.

This variation between mixers suggests that we examine the error statistics for all the RO-DFTB versions to compare against HamqMT. We include only those molecules for which all RO-DFTB versions converge, for a total of 38 molecules when level shifting is incorporated. As we would expect, the charge input methods have no effect on the energy without a geometry optimization. For Hamiltonian mixing with the Broyden and simple mixers, and charge mixing with the Broyden and DIIS mixers, the mean absolute errors are nearly identical, at 1.55 eV. But charge mixing with the simple mixer has a higher MAE of 1.58 eV, and Hamiltonian mixing with the DIIS mixer has the lowest MAE, at 1.52 eV. These

differences mainly come from just three molecules: ethylene, where charge mixing with the simple mixer gives a vertical excitation error more 1.3 eV larger than the other methods; and thymine and uracil, where the DIIS mixer with Hamiltonian mixing results in vertical excitation errors 0.56 eV and 0.57 eV lower than the Broyden mixer, respectively.

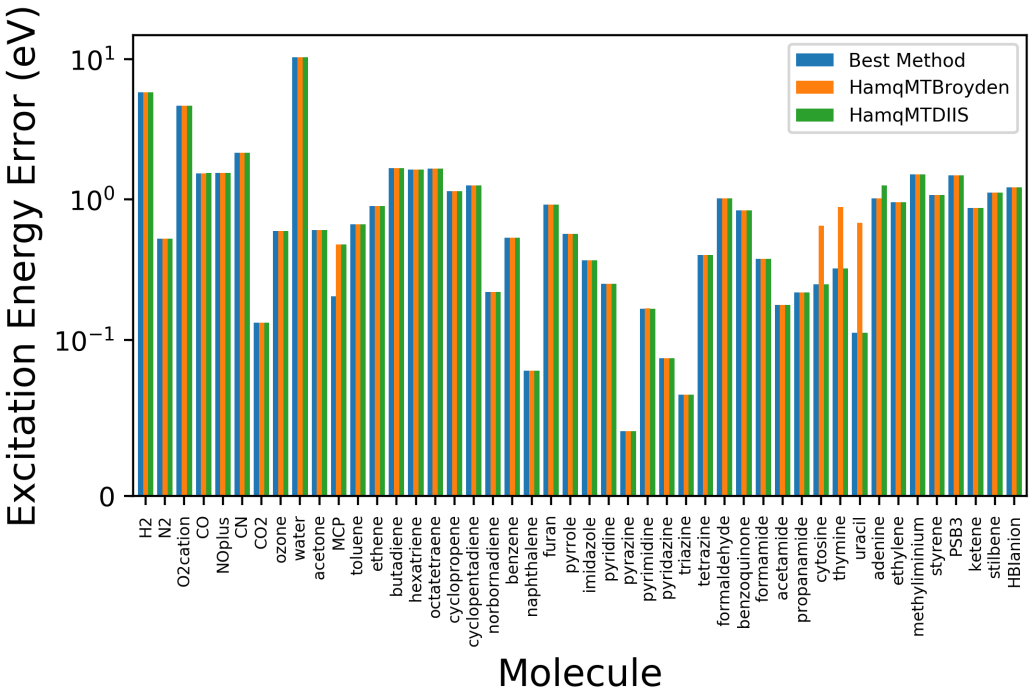

Figure 4.9 Vertical excitation energy errors of each molecule are shown for HamqMTBroyden and HamqMTDIIS RO-DFTB and for the RO-DFTB version which has the lowest error.

To look at the errors for molecules where not all RO-DFTB versions converge, we identify the most accurate RO-DFTB version for each molecule and compare it with the HamqMT-Broyden and HamqMTDIIS versions. As shown in Figure 4.9, there is one molecule where a different RO-DFTB version converged to a more accurate single-point energy: methylcyclo-propene. The charge mixing versions with the DIIS mixer yield a vertical excitation energy error of 0.2 eV, less than half of all Hamiltonian mixing errors, which are 0.48 eV. However, this only occurs with full level shifting: otherwise, all charge mixing versions have an error of -4.19 eV. This is quite a large difference, and dependent entirely on the mixer and degree

of level shifting, neither of which should contribute in theory.

Interestingly, all of the nucleotides have particularly large variations between mixers as well; cytosine is also significantly more accurate with the DIIS mixer, in addition to thymine and uracil. However, the the Broyden mixer achieves a lower error for the fourth nucleotide, adenine. These variations in the RO-DFTB versions which will be explored in more detail in section 4.6. For now, we will move on to evaluating RO-DFTB accuracy in calculating Stokes shifts.

4.5 Stokes shifts

To evaluate the accuracy of the excited-state geometry optimizations performed in RO-DFTB, we can use reference data for Stokes shifts, which we have for two of the test sets. The Stokes shift is the difference between the vertical excitation energy and the emission energy, where the emission energy is taken as the difference between the geometry optimized excited-state energy and the ground state energy at that geometry. The small molecule reference data comes from ROKS with the ωB97X-D functional and 6-311+G* basis set, and the reference values for the dye set are experimental values compiled in Ref. [65]. One molecule has been excluded from the small molecule set, since the ROKS Stokes shift for OH^- was an unrealistically high 25 eV, well above its ionization energy.

Scatter plots of the HamqMTBroyden version of RO-DFTB (with level shifting as needed) and ΔDFTB are shown in Figure 4.10 for each test set. Although ΔDFTB and RO-DFTB differ more significantly than with vertical excitation energies, their agreement is still quite strong. In the small molecule set, the ΔDFTB data include one Stokes shift that RO-DFTB does not: even though RO-DFTB does converge the geometry optimization of water, the ground state calculation at this geometry fails. This is likely because one of the hydrogens has dissociated, but it has only moved about 2 Å further than with ΔDFTB, which predicts an O-H 'bond' length of 3.15 Å. In fact, the first excited state of water does not involve dissociation, but has a linear structure with bond lengths of around 1.23 Å, which is the structure achieved with ROKS. In any case, the water optimization is a failure for both RO-DFTB and ΔDFTB, but the RO-DFTB regression has the advantage of missing this outlier.

The other discrepancy between ΔDFTB and RO-DFTB comes from CO_2. In this case, RO-DFTB managed to converge to a geometry very close to the one predicted by ROKS (this time the excited state is bent rather than linear), while ΔDFTB maintained the linear geometry of the ground state and its Stokes shift suffered for it. The Stokes shift which both

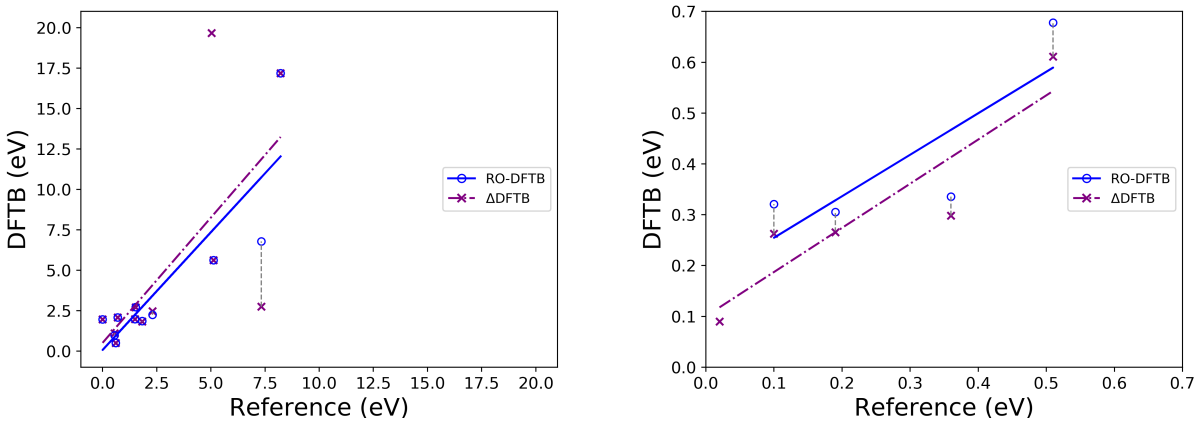

Figure 4.10 Stokes shifts of HamqMTBroyden RO-DFTB and ΔDFTB versus reference values for the small molecule set and large dye set, reported in eV. Small molecules reference data obtained from ROKS ωB97X-D/6-311+G*, while dye set reference values are from Ref. [65]. Regression statistics are shown in Table 4.3.

	Small Molecules			Dye Set		
method	m	b	R^2	m	b	R^2
RO-DFTB	1.46	0.06	0.77	0.81	0.17	0.69
ΔDFTB	1.55	0.50	0.46	0.87	0.10	0.83

Table 4.3 Slope m and y-intercept b (in eV) and coefficient of determination for RO-DFTB and ΔDFTB Stokes shifts.

RO-DFTB and ΔDFTB overestimate to the greatest extent is H_2, which is somewhat of a special case anyway as it dissociates in the excited state, having a bond order of zero.

For the dye set, although the discrepancies appear larger on the plot, they are all significantly less than 0.1 eV and represent a deviation of around 10 % of the Stokes shift, which is about the magnitude of the difference between the RO-DFTB and ΔDFTB Stokes shifts for methylcyclopentene (MCP) and toluene in the small molecule set. It is noticeable though that RO-DFTB arrives at consistently higher Stokes shifts than ΔDFTB for the dye molecules, which is the case for about 68 % of all Stokes shifts calculated (including those without reference values). This is opposite the trend toward underestimation of single-point energies, although systematic disagreements between ΔDFTB and RO-DFTB are less

important to understand than the actual errors.

While neither RO-DFTB or ΔDFTB consistently over- or under-estimate Stokes shifts for the small molecules, they both overestimate the Stokes shifts for the dye molecules except for dye **8**, with RO-DFTB having a more significant systematic error. The regression statistics in Table 4.3 are more favorable to RO-DFTB for the small molecule set (aided by the omission of water and the improved CO_2) but are less favorable for the dye set. While a systematic shift would not affect the regression, the poorer correlation is at least in part due to the different convergence of RO-DFTB and ΔDFTB. ΔDFTB converges for dye **5**, which has a Stokes shift outside of the range of the other reference values, while RO-DFTB converges for dye **2**, which has a Stokes shift that clusters with the other dyes. With these factors taken into account, it appears that both ΔDFTB and RO-DFTB can offer reasonable accuracy for Stokes shift calculations, which provides some indirect validation of the accuracy of the RO-DFTB geometry optimizations.

	Small Molecules			Dye Set		
method	MSE	MAE	RMSE	MSE	MAE	RMSE
HamqMTBroyden	0.40	2.31	3.99	0.12	0.13	0.14
HamqMTDIIS	0.40	2.31	3.99	0.13	0.14	0.16
HamqMTSimple	-0.01	2.65	4.22	0.26	0.26	0.26
ΔDFTB Broyden	-0.03	2.65	4.25	0.07	0.09	0.10

Table 4.4 Stokes shift error statistics for three RO-DFTB methods and ΔDFTB, for the small molecule set and the large dye set, against reference values from ROKS and Ref. [65]. (Reported in eV.)

Error statistics for the two Stokes shift sets are shown in Table 4.4, where the three different mixers are reported separately for HamqMT RO-DFTB. These errors are averaged over the molecules for which all mixers and ΔDFTB obtain Stokes shifts, which excludes MCP and toluene, since the simple mixer did not converge for these. We see that the Broyden and DIIS mixers are in close agreement even after geometry optimization, while the simple

mixer has slightly greater errors. Still, the relatively low and systematic errors for the large dye set support the usefulness of RO-DFTB, especially with the application of a correction factor.

It is more difficult to compare the RO-DFTB versions for the Stokes shifts than for the vertical excitation energies, as only three compounds converge for all versions: CO, N_2, and acetone. Still, four versions do have lower MAEs than the rest (by only 0.001 to 0.004 eV): Hamiltonian mixing with the Broyden mixer for all charge inputs, and the chargeq0Broyden method. This success of the chargeq0Broyden method is somewhat surprising, given both the Broyden mixer's poor performance with charge mixing and the fact that all other charge mixing versions had higher errors than any Hamiltonian mixing version.

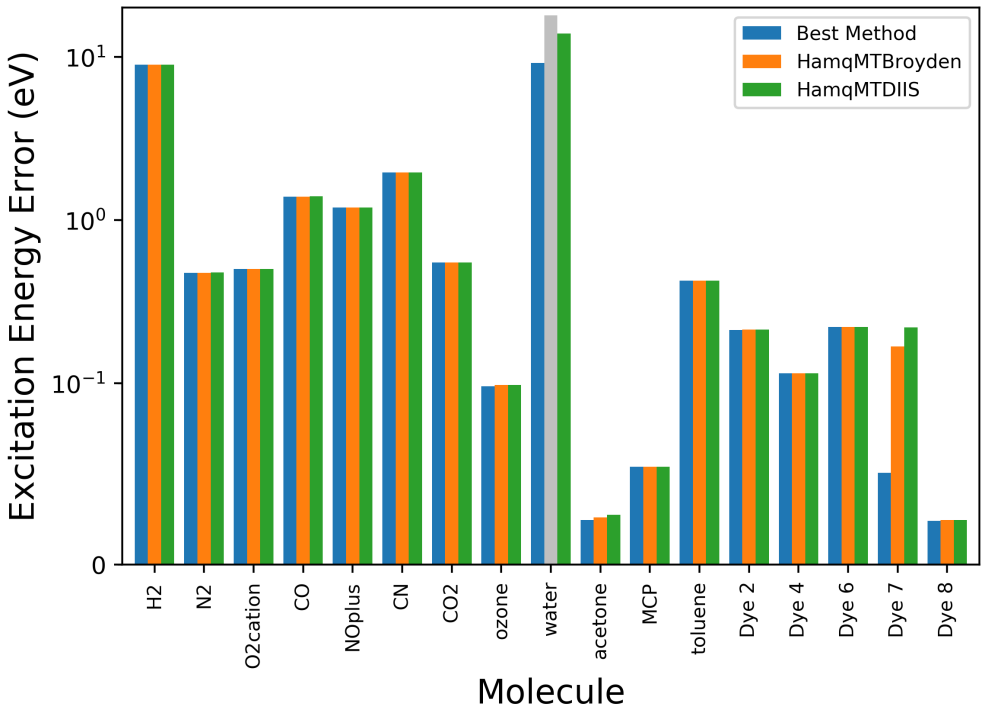

Figure 4.11 Stokes shift errors of the HamqMTBroyden and HamqMTDIIS RO-DFTB versions are compared to the version with the lowest Stokes shift error for each molecule. The grey bar represents the HamqMTBroyden Stokes shift which did not converge for water.

To look at the Stokes shift errors of the compounds where not all RO-DFTB versions converge, we plot the lowest RO-DFTB error alongside the errors of the HamqMTBroyden and HamqMTDIIS methods in Figure 4.11. In fact, there are at least three molecules which reach a better Stokes shift with a method other than HamqMTBroyden or HamqMTDIIS. The first is water, which is not difficult since the HamqMTDIIS method actually has the largest error. And since no RO-DFTB version converges excited-state water to a reasonable geometry, it is not particularly important which method is slightly less inaccurate. Acetone also reaches a better Stokes shift without a HamqMT version, a slight improvement of 0.0014 eV with chargeq0DIIS. But given that the typical errors are on the order of 0.1 to 1 eV, this does not seem highly significant.

The largest decrease in error is for dye **7**, for which the HamqSumDIIS version calculates a Stokes shift 0.118 eV lower than the HamqMTBroyden version. The HamqSumDIIS version is the only RO-DFTB version which reaches this more accurate Stokes shift, with an error of only 0.051 eV; the other versions have errors between 0.168 eV and 0.334 eV. While it is possible that this does suggest that the qSum method is more accurate than other RO-DFTB methods, such conclusions are probably better addressed with a deeper exploration of the variability of the RO-DFTB versions.

4.6 Hamiltonian convergence

At numerous times our results have indicated a higher level of variability among the RO-DFTB versions than expected. In particular, different mixers with the same mixing and charge input routine can reach energies that are much more different than any variations that occur with ΔDFTB. Additionally, different degrees of level shifting may also converge to different energies. In many cases, this appears to occur when charge convergence is reached but the error matrices of the RO Hamiltonian are not near zero. This motivated the inclusion of Hamiltonian-based convergence, with the hope that adhering to the necessary conditions for RO convergence would minimize errors and improve consistency.

Unfortunately, converging by means of the error matrices comes at a cost, which is decreased convergence. A full 16 molecules fail to converge even a single-point calculation, including all of the dyes, half of the REKS set, and two nucleotides, as well as naphthalene, octatetraene, hexatriene, and toluene.

Figure 4.12 summarizes the differences between charge and Hamiltonian convergence with the DIIS mixer, calculating the variance across the three degrees of level shifting. The three charge input methods are shown, but there does not appear to be significant trends among them. What is most apparent is that Hamiltonian convergence tends to fail mainly for the molecules where variance is greatest. It does achieve much lower variances for CO^- and NO^+, as well as pyridine and pyrimidine, but it is hard to know the significance of the convergence failures. Convergence might be reached with additional level shifting, and Hamiltonian mixing may more accurate because it rejects results obtained when near-degeneracy contaminates the optimization. Or this could be part of the approximate nature of DFTB, and ease of convergence may be more important than energy variations near the tolerance level—especially given that the RO-DFTB error analysis is already comparable to ΔDFTB.

One question is whether Hamiltonian convergence actually improves the accuracy of RO-

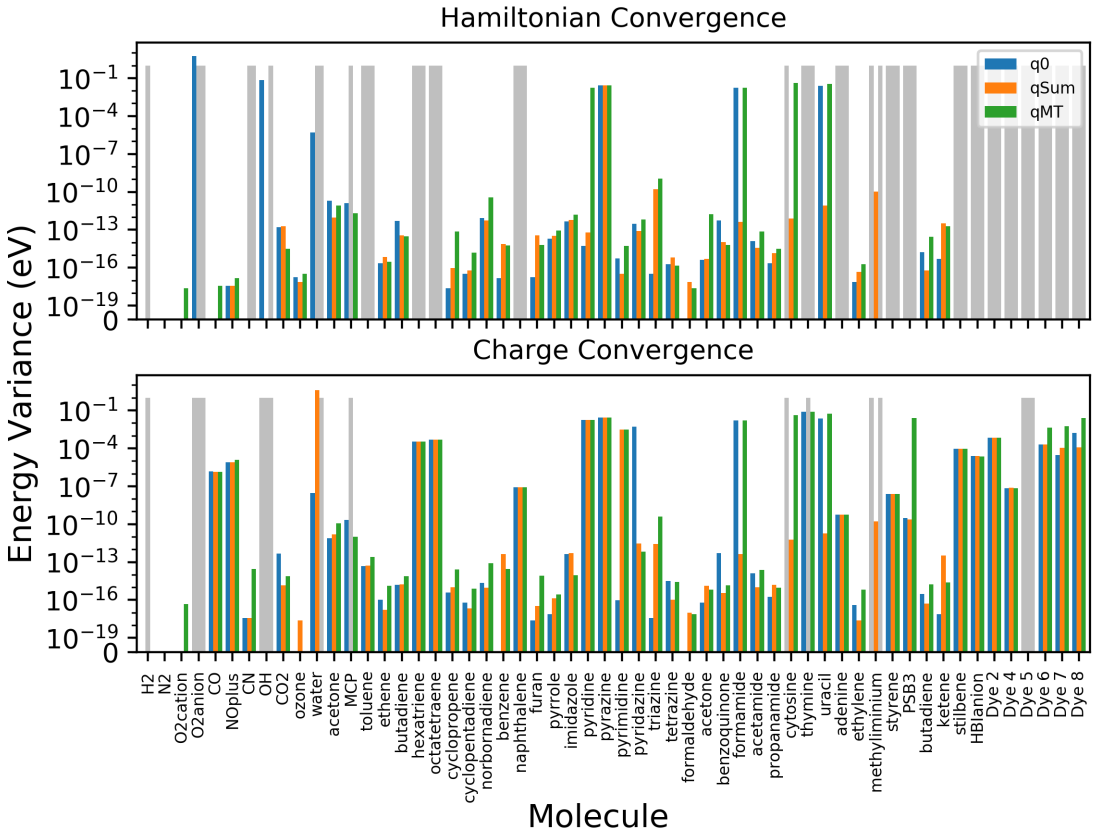

Figure 4.12 Energy variance of the Hamiltonian-based and charge-based convergence schemes, calculated with the DIIS mixer over the three degrees of level shifting. The three charge input methods are distinguished, and grey bars represent calculations that failed to geometry optimize for one or more level shiftings.

DFTB calculations. According to the average errors for the vertical excitation energies and Stokes shifts, it does not—the RMSEs over the molecules for which all mixers converge are identical to several decimal places. They are also identical between mixers; since Hamiltonian convergence generally only succeeds for molecules with low variance, their exclusion eliminated the largest inconsistencies.

This does give us confidence in the typical charge-based convergence, but it is worth paying special attention to the convergence of molecules that fail to reduce the error matrices to zero. To look at some of these compounds in more detail, we proceed to our analysis of RO-DFTB geometry optimizations and the energy gradient calculations they rely on.

4.7 RO-DFTB geometry optimization

As a first test of geometry optimized structures beyond the benchmarking of Stokes shifts, we can compare RO-DFTB converged structures with converged ΔDFTB structures using RMSDs. While ΔDFTB geometries themselves are not the most accurate reference, it is a well-established method with tested analytical gradients, and we mostly expect RO-DFTB to behave similarly.

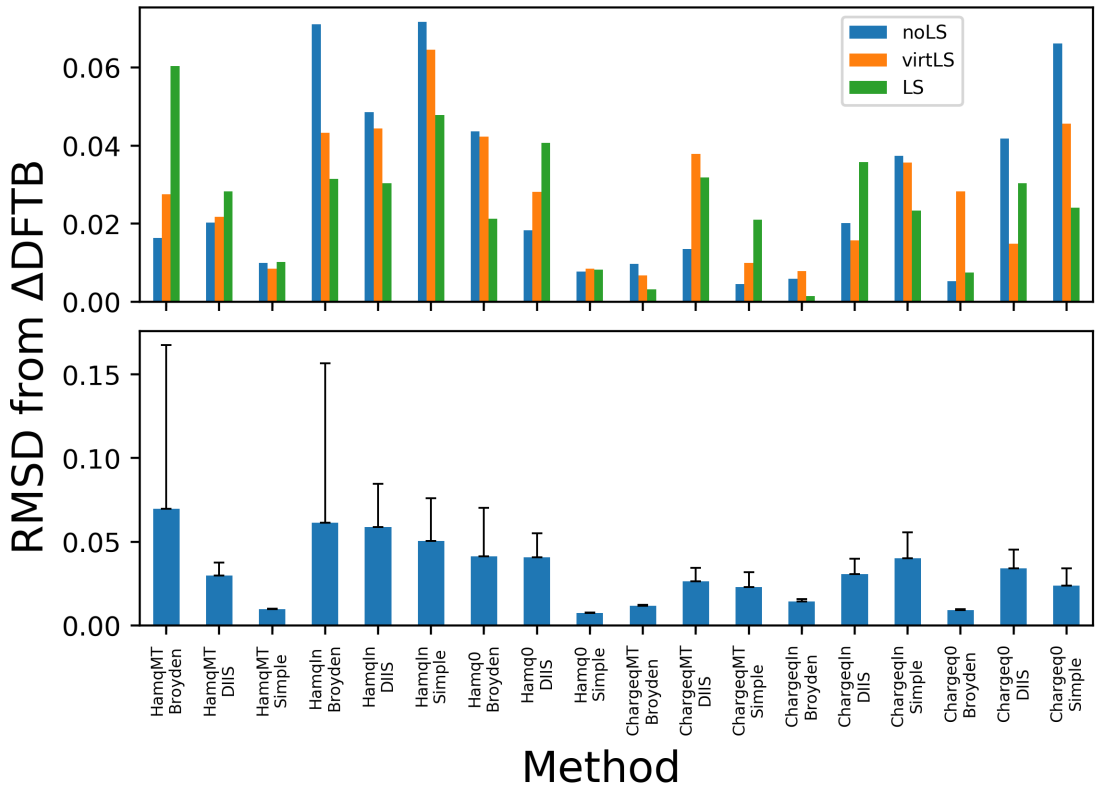

Figure 4.13 RMSD of converged RO-DFTB geometries against converged ΔDFTB geometries, averaged across all molecules for each RO-DFTB version. (Top) Showing the effects of level shifting, and (bottom) with level shifting included as necessary for convergence, showing only the positive range of the variance.

As shown in Figure 4.13, there is a wide range of average deviations from ΔDFTB geometries among the RO-DFTB versions, and an even wider range of variations in the

RMSD. The most inconsistent versions are HamqMTBroyden and HamqSumBroyden, while the HamqSumSimple version behaves much differently than the other Hamiltonian-mixing simple mixers. This variation is somewhat surprising for the simple mixer, and the qSum versions also stand out due to their high RMSD variance without level shifting, which seems to decrease as level shifting is added. It should be noted though that differences in convergence between versions have not been taken into account, so some differences are likely when methods fail to optimize tricky geometries. Nonetheless, it is interesting to see which methods are more apt to converge highly dissimilar geometries. Lower RMSD values are less meaningful since they could represent either non-convergence or close agreement with ΔDFTB.

A more robust means of evaluating geometry optimization accuracy is by comparing the analytical energy gradients computed in DFBT+ to numerical derivatives calculated from single-point energies around the geometry. For these numerical calculations we have used a step size of 0.0001 Å, which consistently yields gradients in good agreement with ΔDFTB. In fact, all numerical gradients for ΔDFTB are within 0.00001 H/bohr for all atoms of all molecules at all geometry steps—except for water with the DIIS mixer, which does not converge, and for dye **5** with both Broyden and DIIS mixers, although it does converge. Even for these molecules, the magnitudes of the gradient discrepancies are 1.2 H/bohr and 0.13 H/bohr, respectively.

The results are not as clean for RO-DFTB. One issue with the RO-DFTB numerical gradients is that the Broyden mixer can sometimes lead to charges which diverge to infinity. This results in “convergence" to an energy that is large and positive, but finite, which results in sporadic massive gradients. This behavior does not occur with either the DIIS or simple mixers, and does not seem to appear in geometry optimization with Hamiltonian mixing. However, these errors may well the cause of the poor Broyden mixer performance with the charge mixing RO-DFTB versions.

Yet the Broyden mixer divergences do not account for the majority of the numerical and

analytical gradient disagreements. There can also be smaller anomalies when a single-point calculation converges to a slightly different energy on one side of the derivative and not the other. This is most common with the DIIS mixer, although it can occur with the Broyden mixer as well. A single-point energy change of even 0.01 eV can still result in a sizeable spike in the numerical gradient.

Unfortunately, but perhaps not surprisingly, Hamiltonian convergence did not significantly improve numerical gradient agreement. A larger step size of 0.001 Å was also tested, but this had no effect.

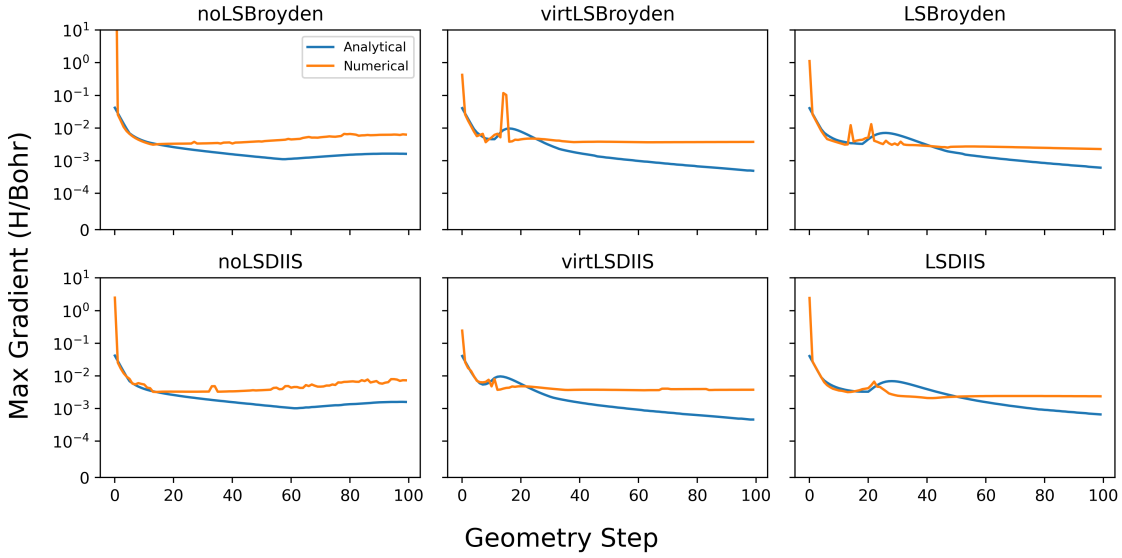

Figure 4.14 Analytical and numerical gradient maximums for dye **2** during the first 100 steps of the analytical optimization. The Hamq0 routine is used with both the Broyden and DIIS mixer, as well as the three degrees of level shifting.

Figure 4.14 shows the maximum analytical and numerical gradients for the first 100 steps of the geometry optimization of dye **2** with the steepest descent method and Hamq0 RO-DFTB. (The q0 input charges help in the direct comparison of gradients, since the single-point derivative calculations use q0 input charges as well.) While these plots are smoother than many, a few numerical spikes occur most prominently with the Broyden mixer when level shifting is added. It is also notable that the addition of level shifting changes the shape of the analytical gradient curve. Dye **2** is one of the molecules with high variance among

degrees of level shifting, although the extent to which level shifting can alter the optimized geometries has not been explored.

The most evident problems with these gradient plots, though, are that the magnitudes of the numerical and analytical gradients hardly ever coincide. On the initial geometry step they never match, and the Broyden mixer without level shifting calculates at least one numerical gradient several orders of magnitude too large. These optimizations do converge in DFTB+, but the numerical gradients do not reduce.

This occurs for a handful of molecules, where the numerical gradients seem to point to a significantly different geometry. These make up the majority of the set of molecules which fail to converge based on Hamiltonian convergence: all of the large dyes, stilbene, HBI^-, styrene, and PSB3 of the REKS set, thymine and adenine, hexatriene and octatetraene, and napthalene of the Thiel set, and also toluene.

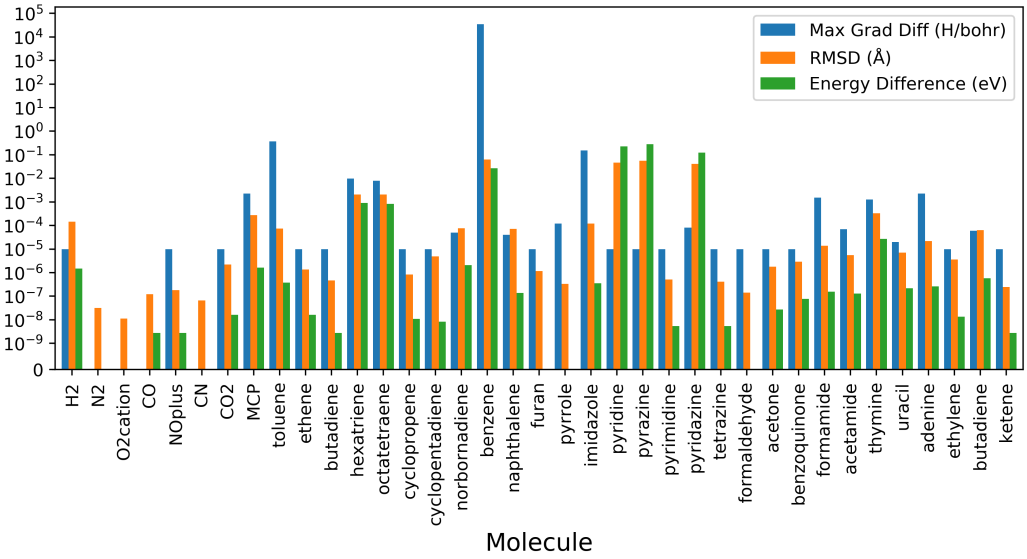

Figure 4.15 Comparison of numerical and analytical geometry optimizations, including the maximum difference between the numerical and analytical gradients during the analytical optimization, the RMSD between the analytically and numerically optimized structures, and the difference in energy at the converged geometry.

To examine whether numerical gradient discrepancies for RO-DFTB are due to errors in the single-point calculations or whether they suggest that the analytical gradients are

actually inaccurate, geometry optimizations have been carried out following the numerical gradients. Figure 4.15 compares the maximum difference between numerical and analytical gradients during the analytical optimization with the RMSD between the analytically and numerically optimized geometries. These are further compared against the magnitude of the difference in energies between the two optimized geometries.

This set comprises only the molecules which have converged so far with the numerical optimization. This does include several compounds with divergent numerical and analytical gradients during the analytical optimization, but their RMSDs from the analytically optimized structures are not significantly larger than for those where the both gradients converged. Interestingly, pyridine, pyrazine, and pyridazine all had relatively low disagreement between gradients, but their RMSDs suggest they optimized to significantly different geometries.

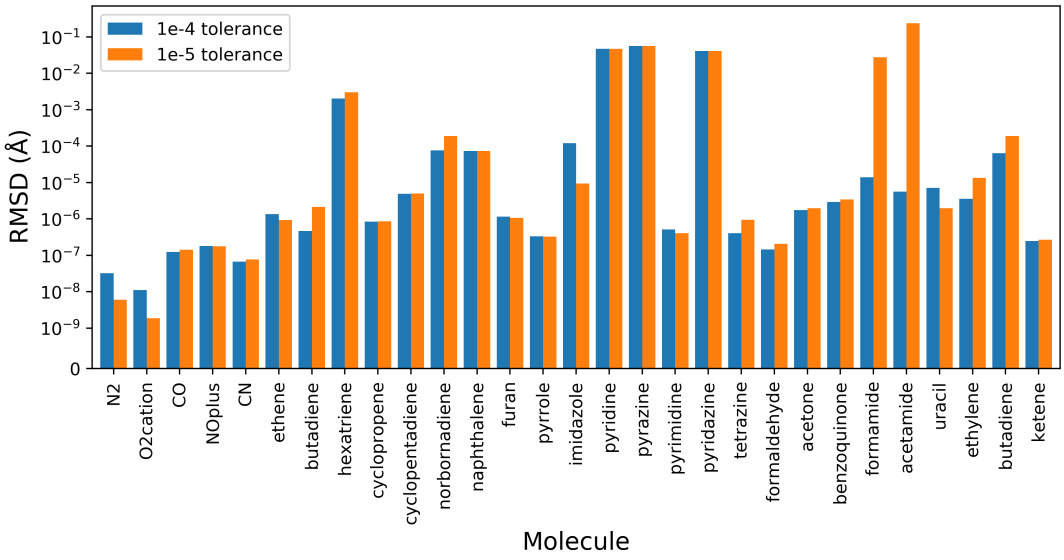

Figure 4.16 RMSDs between numerically and analytically optimized geometries, converged to both a 1e-4 tolerance and a 1e-5 tolerance.

To explore these differences in numerical and analytical optimization further, we increased the convergence tolerance for both optimizations and compared RMSDs again. As shown in Figure 4.16, the RMSD does not change significantly for most molecules. It increases

dramatically for formamide and acetamide, but these are the only two molecules which fail to reach the tighter tolerance with the analytical optimization. Their RMSDs are large because they have bent out of plane; a closer examination of their respective potential energy surfaces at a higher level of theory would be useful in determining if this is a reasonable optimization trajectory.

This further optimization to gradients below 1e-5 did not decrease most RMSDs as we would expect if the analytical and numerical gradients were converging to the same geometry. But a closer look at some of the structures where the gradients are most divergent may help understand the differences we've observed. For dye **2**, analytical and numerical gradients with Hamq0 RO-DFTB are shown in Figure 4.17 at the first step of both the analytical and numerical optimizations. Notably, the magnitudes of these gradients are not directly comparable between the different types of gradient calculation; the numerical gradients are more than an order of magnitude larger than the analytical gradients, so they have been scaled independently.

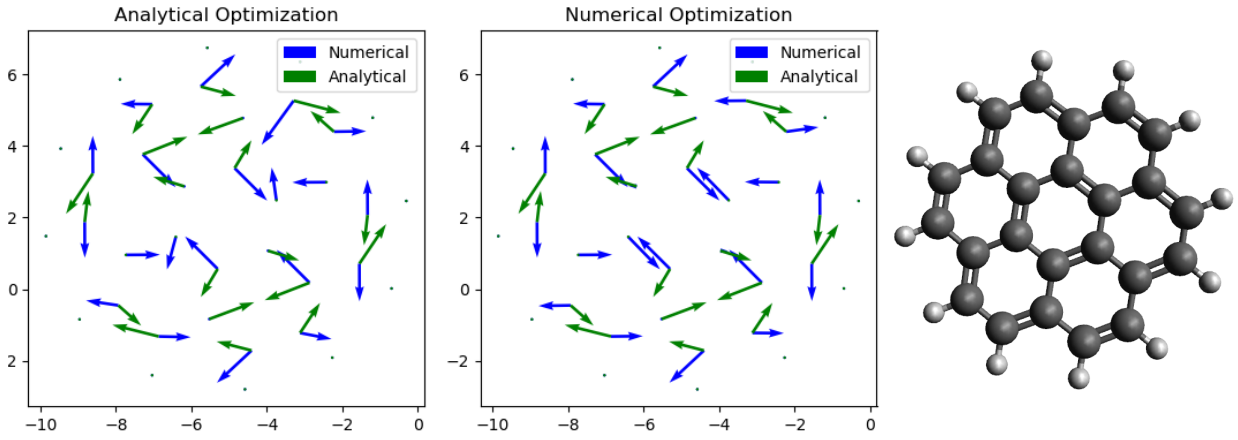

Figure 4.17 (Left) Vectors representing the analytical and numerical gradients (in H/bohr) from both the analytical and numerical optimizations of dye **2** at the ground state geometry. Gradients in the z-axis are negligible. (Right) The structure of dye **2**.

The analytical gradients are identical, as we would expect, but the numerical gradients differ for several atoms. Additionally, we see in Figure 4.18 that after three geometry steps

the analytical and numerical gradients are in close agreement at least on the direction of the gradients. But the geometries have clearly diverged, as both gradients take on a significantly different pattern between optimizations.

One concern is that the numerical geometry optimization still has occasional spikes in gradients, as shown for two of the atoms at geometry step 3. This sort of instability does not occur during analytical gradient calculations, so it is possible that these occasional spikes could disrupt a numerical optimization that would otherwise agree with its analytical counterpart.

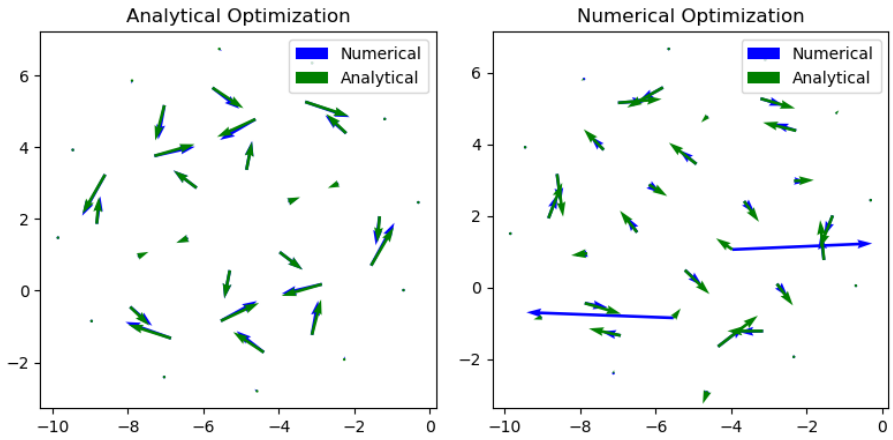

Figure 4.18 Numerical and analytical gradients for both optimizations of dye **2** at geometry step 3.

However, stilbene is one example of a numerical optimization that does not seem to have any single-point errors in its gradient calculations. At the initial geometry, its analytical and numerical gradients align reasonably well, although the maximum numerical gradient is slightly larger than the maximum analytical gradient. But by geometry step 20 (shown in Fig. 4.19) it is clear that the optimizations have diverged; the analytical gradients point towards a longer center C=C bond length, while the numerical gradients favor a shorter bond length. (Recalling that atoms move in the direction opposite the gradient, towards a more negative energy.) This discrepancy between numerical and analytical is maintained until both optimizations converge. At convergence, the optimized C=C bond lengths are

1.450 Å for the analytical optimization and 1.445 Å for the numerical optimization, while ΔDFTB converges to the lowest bond length, at 1.438 Å.

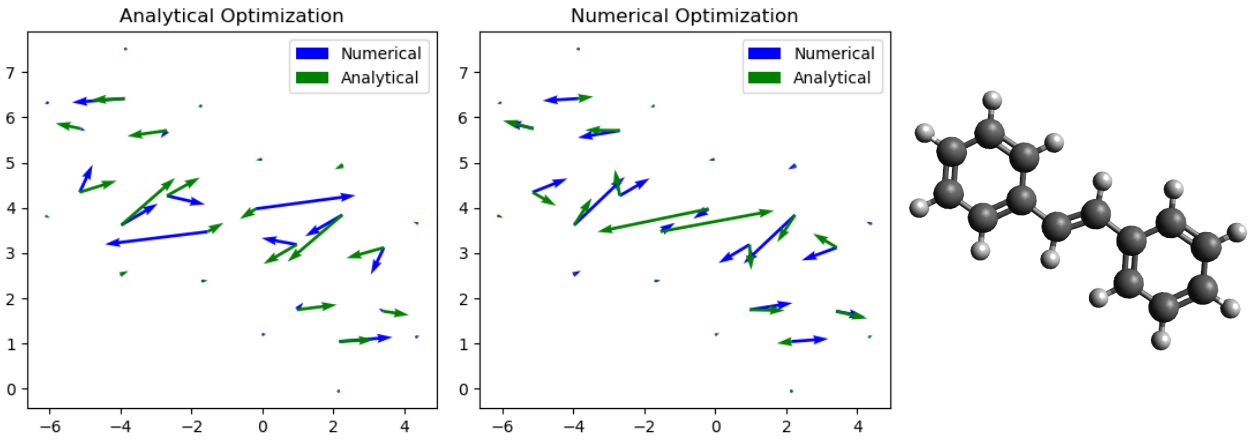

Figure 4.19 (Left) The analytical and numerical gradients for stilbene at the 20th geometry step. (Right) The structure of stilbene.

With this relatively small difference between geometries, it is hard to say whether one set of gradients is more correct than the other, though it is fairly clear that something is different between the numerical and analytical gradients beyond single-point convergence errors. The q0 charge inputs that we have used offer the best chance of maintaining agreement between single-point calculations and a geometry optimization, but some gradients still consistently diverge.

We have not yet looked at the qMT or qSum method geometry optimizations though, and doing so may offer additional insight both into the differences between the charge input methods and in the analytical gradients. An interesting example is CO_2, which we know the HamqMTBroyden method is able to converge to a bent geometry. In fact, potential energy surfaces from both ΔDFTB and RO-DFTB along the O-C-O bond angle degree of freedom show that the linear geometry is a local maximum, as in Figure 4.20. The improved convergence rate of RO-DFTB is evident in the two geometries that it converges below 100 ° which ΔDFTB does not, and any increased energy variability is negligible at this scale.

As we noted in the Stokes shifts results, ΔDFTB does not converge to a bent geometry.

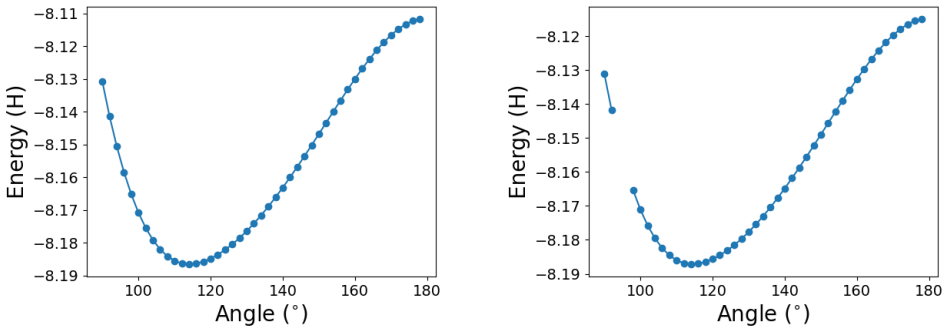

Figure 4.20 The potential energy surface of CO_2, calculated with ΔDFTB (left) and HamqMTBroyden RO-DFTB (right). The bond lengths for each method correspond to the excited-state bond lengths from ΔDFTB.

Instead, it converges in only 7 geometry steps to a linear structure. By starting with an initial geometry bent to 179 °, though, ΔDFTB does optimize to a symmetric bent structure for all mixers when the conjugate gradient geometry driver is used. Optimizations from the ground and bent initial geometries are shown in Figure 4.21, where we can see that the bond lengths asymmetrize significantly during the optimization but eventually return to symmetrical. The optimizations are more nuanced for RO-DFTB, shown in Figure 4.22 across charge input methods for the Broyden mixer with full level shifting.

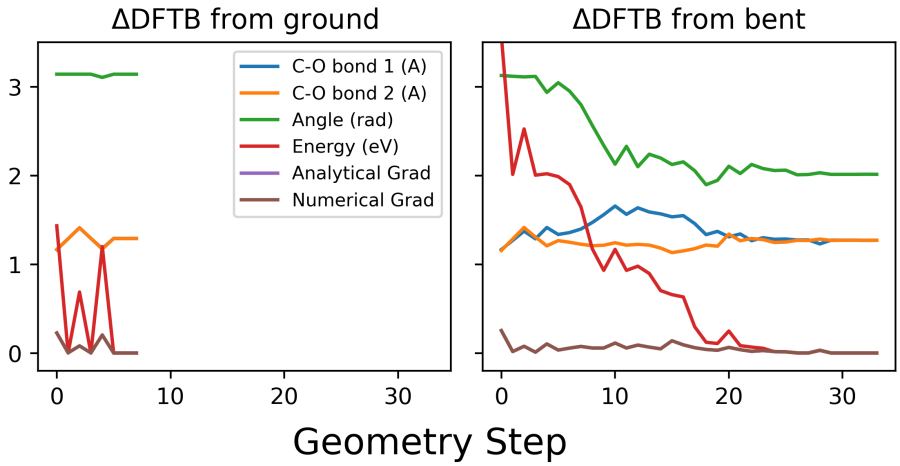

Figure 4.21 Geometry optimization of CO_2 with ΔDFTB, which compares bond lengths, the O-C-O bond angle, the energy difference from the minimum, and both analytical and numerical gradients on the same scale. Gradients are in H/bohr.

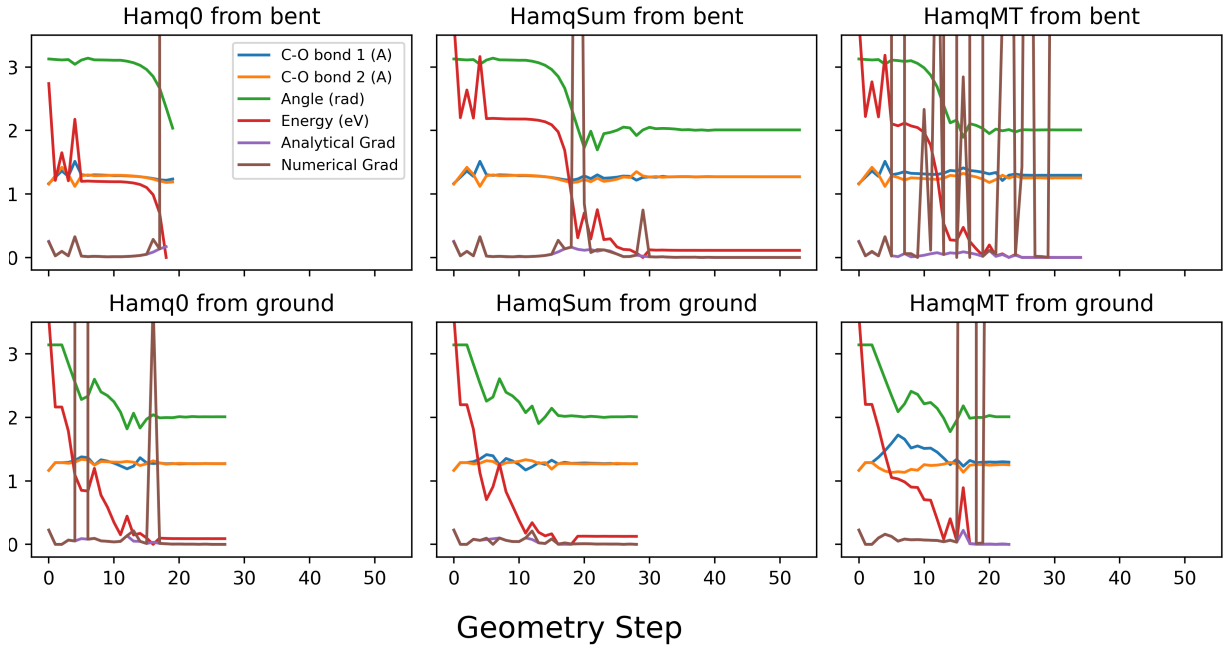

Figure 4.22 Geometry optimization of CO_2 with Hamiltonian mixing versions of RO-DFTB, all using full level shifting and the Broyden mixer. Energies shown are subtracted from the minimum energy, and the gradients are in H/bohr.

Qualitatively, the accepted geometry for the first excited state of CO_2 is a bent symmetrical configuration,[68,69] which is in agreement with our ROKS reference calculations. But HamqMT RO-DFTB predicts asymmetrical bond lengths for every mixer and degree of level shifting—except when starting from the 179 ° bent geometry. Without level shifting, HamqMTBroyden RO-DFTB converges back to a linear geometry. HamqMT optimizations from the bent geometry also have the most erratic numerical gradients, although it's still unclear if this reflects poorly on the analytical optimization. The performance of the qSum and q0 methods in Figure 4.22 is similar to qMT, but all of their converged optimizations arrive at symmetrical bond lengths. It is interesting that all of the RO-DFTB methods have a harder time converging with the bent initial geometry; the Hamq0 optimization from bent is the only Broyden method with level shifting that does not converge, and both of the other charge input methods take many more SCC iterations to converge than with a linear initial geometry. But both Broyden and DIIS mixers do consistently and efficiently converge to a

bent geometry from the ground state geometry, which is a noteworthy improvement from ΔDFTB.

Overall, the status of the analytical gradients looks promising for RO-DFTB. The Stokes shifts are in close agreement with ΔDFTB and are reasonably accurate when compared to reference values; the geometry convergence rates are excellent for Hamiltonian mixing RO-DFTB; and the numerical and analytical gradients do match closely for most of the molecules in our test sets. For the instances where numerical and analytical gradients and optimizations do diverge, and where single-point energy calculations are erratic, there is a potential for future investigation into the source of these discrepancies.

Chapter Five

CONCLUSIONS

We have here presented our recent work on the implementation and evaluation of a restricted open-shell method for singlet excited states in the DFTB framework. Through rigorous analysis of the different mixing methods, charge inputs, and convergence methods, we have found that one RO-DFTB routine matches or outperforms the others in nearly every test. HamqMT RO-DFTB has the best convergence rate, and compute time of any other method, both with the Broyden and DIIS mixers. It has comparable vertical excitation energy and Stokes shift accuracy to the other RO-DFTB methods, though there is some variation between the DIIS and Broyden mixers. The Hamq0 and HamqSum methods also perform very well, and have been shown to be advantageous in certain cases, such as the Stokes shift of dye **7** and the geometry optimization of CO_2. In general these methods are highly similar, and any of them are capable of calculating approximate excited-state energies and gradients in DFTB.

The poor performance of the charge mixing and Hamiltonian convergence methods is in a way advantageous, making RO-DFTB calculations easier to run by reducing the number of choices necessary. Level shifting, and input charges provide a reasonable number of options if convergence fails, in addition to the choice between the DIIS and Broyden mixers. By automating the addition of level shifting, these calculations would be even easier.

We have shown RO-DFTB with Hamiltonian mixing to be highly comparable to ΔDFTB in both speed and accuracy, while offering better overall convergence rates. This is a signif-

icant result, and could have a real impact on the success of molecular dynamics simulations and large screening studies, where an improvement in convergence of 5-10 % could save significant time and resources. Although RO-DFTB is not twice as fast as ΔDFTB, it does still offer a modest reduction in SCC steps in most cases, as well as reduced runtimes for more than half of our test molecules.

The remaining work to further validate RO-DFTB should focus on a few key areas, mainly on the gradient calculations and the Hamiltonian convergence. Continuing the evaluation of the numerical and analytical gradients may yet be able to identify a source of the consistent disagreements that occur for some molecules, which would increase our confidence in RO-DFTB geometry optimization. A more thorough investigation of single-point energy instabilities in the numerical gradient calculations would also be useful, especially in determining the reason for the infinite charge divergence of the Broyden mixer; this may point to why the charge mixing methods are typically incompatible with the Broyden mixer. As for Hamiltonian convergence, it would be interesting to monitor the magnitudes of the error matrices during geometry optimization, and to see if this is related to the numerical and analytical gradient calculations.

Another future project could automate the use of level shifting in RO-DFTB: this could be added either by restarting after a calculation has failed or by identifying cases where orbital energies suggest imminent non-convergence (i.e., degeneracy in the orbitals between shells). With these minor adjustments to the current implementation, the results we have presented indicate that RO-DFTB should prove to be a useful tool for future electronic excited-state applications.

REFERENCES

(1) Kundu, S.; Patra, A. *Chem. Rev.* **2017**, *117*, 712–757.

(2) Yu, J.; Li, X.; Deria, P. *ACS Sustainable Chem. Eng.* **2019**, *7*, 1841–1854.

(3) Nelson, T. R.; White, A. J.; Bjorgaard, J. A.; Sifain, A. E.; Zhang, Y.; Nebgen, B.; Fernandez-Alberti, S.; Mozyrsky, D.; Roitberg, A. E.; Tretiak, S. *Chem. Rev.* **2020**, *120*, 2215–2287.

(4) Li, W.; Long, R.; Tang, J.; Prezhdo, O. V. *J. Phys. Chem. Lett.* **2019**, *10*, 3788–3804.

(5) Quant, M.; Hamrin, A.; Lennartson, A.; Erhart, P.; Moth-Poulsen, K. *J. Phys. Chem. C* **2019**, *123*, 7081–7087.

(6) Cembran, A.; Bernardi, F.; Garavelli, M.; Gagliardi, L.; Orlandi, G. *J. Am. Chem. Soc.* **2004**, *126*, 3234–3243.

(7) Quant, M.; Lennartson, A.; Dreos, A.; Kuisma, M.; Erhart, P.; Börjesson, K.; Moth-Poulsen, K. *Chem. - Eur. J.* **2016**, *22*, 13265–13274.

(8) Liu, Y.; Grossman, J. C. *Nano Lett.* **2014**, *14*, 7046–7050.

(9) Börjesson, K.; Lennartson, A.; Moth-Poulsen, K. *ACS Sustainable Chem. Eng.* **2013**, *1*, 585–590.

(10) Lennartson, A.; Moth-Poulsen, K. In *Molecular Devices for Solar Energy Conversion and Storage*, Tian, H., Boschloo, G., Hagfeldt, A., Eds.; Springer Nature Singapore: 2018; Chapter 9, pp 327–352.

(11) Jevric, M.; Petersen, A. U.; Mansø, M.; Kumar Singh, S.; Wang, Z.; Dreos, A.; Sumby, C.; Nielsen, M. B.; Börjesson, K.; Erhart, P.; Moth-Poulsen, K. *Chem. - Eur. J.* **2018**, *24*, 12767–12772.

(12) Muždalo, A.; Saalfrank, P.; Vreede, J.; Santer, M. *J. Chem. Theory Comput.* **2018**, *14*, 2042–2051.

(13) Lennartson, A.; Roffey, A.; Moth-Poulsen, K. *Tetrahedron Lett.* **2015**, *56*, 1457–1465.

(14) Beharry, A. A.; Sadovski, O.; Woolley, G. A. *J. Am. Chem. Soc.* **2011**, *133*, 19684–19687.

(15) Kuisma, M.; Lundin, A.; Moth-Poulsen, K.; Hyldgaard, P.; Erhart, P. *ChemSusChem* **2016**, *9*, 1786–1794.

(16) Durgun, E.; Grossman, J. C. *J. Phys. Chem. Lett.* **2013**, *4*, 854–860.

(17) Jorner, K.; Dreos, A.; Emanuelsson, R.; El Bakouri, O.; Galván, I. F.; Börjesson, K.; Feixas, F.; Lindh, R.; Zietz, B.; Moth-Poulsen, K.; Ottosson, H. *J. Mater. Chem. A* **2017**, *5*, 12369–12378.

(18) Montavon, G.; Rupp, M.; Gobre, V.; Vazquez-Mayagoitia, A.; Hansen, K.; Tkatchenko, A.; Müller, K. R.; Anatole Von Lilienfeld, O. *New J. Phys.* **2013**, *15*, DOI: 10.1088/1367-2630/15/9/095003.

(19) Koerstz, M.; Christensen, A. S.; Mikkelsen, K. V.; Nielsen, B.; Jensen, J. H. *ChemRxiv* **2019**, 1–16.

(20) Schrier, J. *J. Chem. Inf. Model.* **2020**, *60*, 3804–3811.

(21) Ha, J.-K.; Kim, K.; Min, S. K. *J. Chem. Theory Comput.* **2021**, DOI: 10.1021/acs.jctc.0c01261.

(22) David Sherrill, C.; Schaefer, H. F. In *Advances in Quantum Chemistry*; Academic Press Inc.: 1999; Vol. 34, pp 143–269.

(23) Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

(24) Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

(25) Silva-Junior, M. R.; Thiel, W. *J. Chem. Theory Comput.* **2010**, *6*, 1546–1564.

(26) Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974–12980.

(27) Pople, J. A.; Gill, P. M.; Johnson, B. G. *Chem. Phys. Lett.* **1992**, *199*, 557–560.

(28) Oliveira, A. F.; Seifert, G.; Heine, T.; Duarte, H. A. *J. Braz. Chem. Soc.* **2009**, *20*, 1193–1205.

(29) Seifert, G.; Joswig, J. O. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 456–465.

(30) Koskinen, P.; Mäkinen, V. *Comput. Mater. Sci.* **2009**, *47*, 237–253.

(31) Elstner, M.; Porezag, D.; Jungnickel, G.; Elsner, J.; Haugk, M.; Frauenheim, T.; Suhai, S.; Seifert, G. *Phys. Rev. B* **1998**, *58*, 7260–7268.

(32) Goringe, C. M.; Bowler, D. R.; Hernández, E. *Rep. Prog. Phys.* **1997**, *60*, 1447–1512.

(33) Elstner, M.; Seifert, G. *Philos. Trans. R. Soc., A* **2014**, *372*, 20120483.

(34) Park, Y. C.; Senn, F.; Krykunov, M.; Ziegler, T. *J. Chem. Theory Comput.* **2016**, *12*, 5438–5452.

(35) Zhao, G. J.; Han, K. L. *J. Phys. Chem. A* **2009**, *113*, 4788–4794.

(36) Prlj, A.; Curchod, B. F.; Fabrizio, A.; Floryan, L.; Corminboeuf, C. *J. Phys. Chem. Lett.* **2015**, *6*, 13–21.

(37) Kowalczyk, T.; Yost, S. R.; Voorhis, T. V. *J. Chem. Phys.* **2011**, *134*, DOI: 10.1063/1.3530801.

(38) Ziegler, T.; Rauk, A.; Baerends, E. J. *Theor. Chim. Acta* **1977**, *43*, 261–271.

(39) Ess, D. H.; Johnson, E. R.; Hu, X.; Yang, W. *J. Phys. Chem. A* **2011**, *115*, 76–83.

(40) Gilbert, A. T. B.; Besley, N. A.; Gill, P. M. W. *J. Phys. Chem. A* **2008**, *112*, 13164–13171.

(41) Kowalczyk, T.; Tsuchimochi, T.; Chen, P.-T.; Top, L.; Van Voorhis, T. *J. Chem. Phys.* **2013**, *138*, 164101.

(42) Filatov, M.; Shaik, S. *J. Chem. Phys.* **1999**, *110*, 116–125.

(43) Trani, F.; Scalmani, G.; Zheng, G.; Carnimeo, I.; Frisch, M. J.; Barone, V. *J. Chem. Theory Comput.* **2011**, *7*, 3304–3313.

(44) Hourahine, B. et al. *J. Chem. Phys.* **2020**, *152*, 124101.

(45) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover: Mineola, New York, 1989.

(46) Roothaan, C. C. J. *Rev. Mod. Phys.* **1960**, *32*, 179–185.

(47) Tsuchimochi, T.; Scuseria, G. E. *J. Chem. Phys.* **2010**, *133*, DOI: 10.1063/1.3503173.

(48) Plakhutin, B. N.; Gorelik, E. V.; Breslavskaya, N. N. *J. Chem. Phys.* **2006**, *125*, DOI: 10.1063/1.2393223.

(49) Fægri, K.; Manne, R. *Mol. Phys.* **1976**, *31*, 1037–1049.

(50) Hirao, K.; Nakatsuji, H. *J. Chem. Phys.* **1973**, *59*, 1457–1462.

(51) Edwards, W. D.; Zerner, M. C. *Theoretica Chimica Acta* **1987**, *72*, 347–361.

(52) Kollmar, C. *Int. J. Quantum Chem.* **1997**, *62*, 617–637.

(53) Guest, M. F.; Saunders, Y. R. *Molecular Physics* **1974**, *28*, 819–828.

(54) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(55) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(56) Filatov, M.; Shaik, S. *Chem. Phys. Lett.* **1998**, *288*, 689–697.

(57) Schulte, M.; Frank, I. *Chem. Phys.* **2010**, *373*.

(58) Gaus, M.; Cui, Q.; Elstner, M. *J. Chem. Theory Comput.* **2011**, *7*, 931–948.

(59) Gaus, M.; Goez, A.; Elstner, M. *J. Chem. Theory Comput.* **2013**, *9*, 338–354.

(60) Johnson, D. D. *Phys. Rev. B* **1988**, *38*, 12807–12813.

(61) Kovalenko, A.; Ten-No, S.; Hirata, F. *J. Comput. Chem.* **1999**, *20*, 928–936.

(62) Silva-Junior, M.; Schreiber, M.; Sauer, S.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, DOI: 10.1063/1.2973541.

(63) Filatov, M.; Shaik, S. *Chem. Phys. Lett.* **1999**, *304*, 429–437.

(64) Lee, I. S.; Filatov, M.; Min, S. K. *J. Chem. Theory Comput.* **2019**, *15*, 3021–3032.

(65) Kowalczyk, T.; Le, K.; Irle, S. *J. Chem. Theory Comput.* **2016**, *12*, 313–323.

(66) Niehaus, T.; Elstner, M.; Frauenheim, T.; Suhai, S. *J. Mol. Struct.: THEOCHEM* **2001**, *541*, 185–194.

(67) Deshaye, M. Development of TI-DFTB: Transition Dipole Moment Calculations in a Time-Independent Density Functional Tight-Binding Framework, MA thesis, Bellingham, Washington: Western Washington University, 2020.

(68) McDiarmid, R.; Doering, J. *J. Chem. Phys.* **1984**, *80*, 648–656.

(69) Ma, Y.; Peng, L.; Zhang, H.; Yu, J. *Russ. J. Phys. Chem. A* **2014**, *88*, 2239–2347.

APPENDIX

Appendix A

A.1 The Ziegler sum expression for the singlet excited state energy

We start with the definition of the triplet, singlet, and mixed wavefunctions:

$$\psi_t = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \tag{A.1}$$

$$\psi_s = \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{A.2}$$

$$\psi_m = |\uparrow\downarrow\rangle \tag{A.3}$$

Now, adding the triplet and singlet wavefunctions, we get

$$\begin{aligned}\psi_t + \psi_s =& \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) + \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \\ =& \frac{2}{\sqrt{2}}\,|\uparrow\downarrow\rangle \\ =& \frac{2}{\sqrt{2}}\psi_m \end{aligned} \tag{A.4}$$

Rearranging for the singlet wavefunction,

$$\psi_s = \frac{2}{\sqrt{2}}\psi_m - \psi_t \tag{A.5}$$

and then taking the expectation value of the Hamiltonian, we obtain the singlet energy:

$$\begin{aligned} E_s = \langle\psi_s|H|\psi_s\rangle &= \langle(\frac{2}{\sqrt{2}}\psi_m - \psi_t|H|\frac{2}{\sqrt{2}}\psi_m - \psi_t\rangle \\ &= \langle\frac{2}{\sqrt{2}}\psi_m|H|\frac{2}{\sqrt{2}}\psi_m\rangle - \langle\psi_t|H|\psi_t\rangle \\ &= 2\,\langle\psi_m|H|\psi_m\rangle - \langle\psi_t|H|\psi_t\rangle \\ &= 2E_m - E_t \end{aligned} \tag{A.6}$$

A.2 Derivation of the singlet excited-state energy expression for ROHF

From unrestricted Hartree-Fock, we can write down the Fock operators for the α and β spins of both the mixed-spin and triplet determinants just by observing the spin configuration of the determinants. In general, the unrestricted Fock operator for spin α is

$$F^{\alpha} = h + \sum_{a}^{N^{\alpha}} (J_a^{\alpha} - K_a^{\alpha}) + \sum_{a}^{N^{\beta}} J_a^{\beta} \tag{A.7}$$

with a corresponding expression for β spin electrons. The sums are over all occupied orbitals a, where N^{α} and N^{β} are the number of electrons of α and β spin, respectively.

To write the expressions for the mixed and triplet determinants, we need to choose a spin configuration, although the final expression will be independent of this. We do not have to account for all four spin configurations.

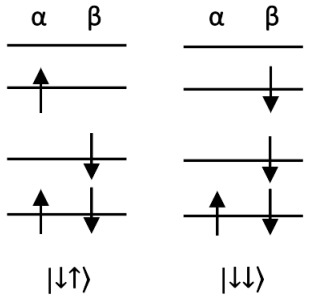

Figure A.1 The mixed and triplet spin configurations used for RO-DFTB.

Choosing the $|\downarrow\uparrow\rangle$ mixed determinant and $|\downarrow\downarrow\rangle$ determinant for the triplet state, as shown in Fig. A.1, we will write the Fock operators by separating out the terms for the closed and open shells. N^c is the number of closed-shell orbitals, which are indexed by k and l, and 1 and 2 are the indices of the first and second orbitals of the open shell. The four Fock

operators are then

$$F_m^\alpha = h + \sum_k^{N^c}(J_k^\alpha - K_k^\alpha) + \sum_k^{N^c} J_k^\beta + J_2^\alpha - K_2^\alpha + J_1^\beta \tag{A.8}$$

$$F_m^\beta = h + \sum_k^{N^c}(J_k^\beta - K_k^\beta) + \sum_k^{N^c} J_k^\alpha + J_1^\beta - K_1^\beta + J_2^\alpha \tag{A.9}$$

$$F_t^\alpha = h + \sum_k^{N^c}(J_k^\alpha - K_k^\alpha) + \sum_k^{N^c} J_k^\beta + J_1^\beta + J_2^\beta \tag{A.10}$$

$$F_t^\beta = h + \sum_k^{N^c}(J_k^\beta - K_k^\beta) + \sum_k^{N^c} J_k^\alpha + J_1^\beta - K_1^\beta + J_2^\beta - K_2^\beta \tag{A.11}$$

Using these four Fock operators, we can write the energy expressions for the mixed and triplet determinants separately. We have only covered the energy expressions of closed-shell restricted systems written in terms of spatial orbitals (summed over $N/2$ doubly-occupied orbitals) in equations 2.3 and 2.23, written in terms of spin orbitals the same closed-shell energy expression is

$$\begin{aligned} E &= \sum_a^N \langle a|h|a\rangle + \frac{1}{2}\sum_a^N\sum_b^N \langle ab|ab\rangle - \langle ab|ba\rangle \\ &= \sum_a^N h_{aa} + \frac{1}{2}\sum_a^N\sum_b^N (J_{ab} - K_{ab}) \end{aligned} \tag{A.12}$$

Whereas the typical energy expression of an unrestricted system is the same but the α and β terms are written separately:

$$\begin{aligned} E = &\sum_a^{N^\alpha} h_{aa}^\alpha + \sum_a^{N^\beta} h_{aa}^\beta + \frac{1}{2}\sum_a^{N^\alpha}\sum_b^{N^\alpha}(J_{ab}^{\alpha\alpha} - K_{ab}^{\alpha\alpha}) \\ &+ \frac{1}{2}\sum_a^{N^\beta}\sum_b^{N^\beta}(J_{ab}^{\beta\beta} - K_{ab}^{\beta\beta}) + \sum_a^{N^\alpha}\sum_b^{N^\beta} J_{ab}^{\alpha\beta} \end{aligned} \tag{A.13}$$

Similarly, the energy of the mixed determinant is

$$\begin{aligned} E_m =& \sum_k^{N^c} h_{kk}^{\alpha} + \sum_k^{N^c} h_{kk}^{\beta} + h_{11}^{\beta} + h_{22}^{\alpha} \\ &+ \frac{1}{2}\sum_{kl}^{N^c}(J_{kl}^{\alpha\alpha} - K_{kl}^{\alpha\alpha}) + \frac{1}{2}\sum_{kl}^{N^c}(J_{kl}^{\beta\beta} - K_{kl}^{\beta\beta}) \\ &+ \sum_k^{N^c}(J_{k1}^{\beta\beta} - K_{k1}^{\beta\beta}) + \sum_k^{N^c}(J_{k2}^{\alpha\alpha} - K_{k2}^{\alpha\alpha}) \\ &+ \sum_{kl}^{N^c} J_{kl}^{\alpha\beta} + \sum_k^{N^c} J_{k1}^{\alpha\beta} + \sum_k^{N^c} J_{k2}^{\alpha\beta} + J_{12}^{\alpha\beta} \end{aligned} \tag{A.14}$$

And the energy of the triplet determinant is

$$\begin{aligned} E_t =& \sum_k^{N^c} h_{kk}^{\alpha} + \sum_k^{N^c} h_{kk}^{\beta} + h_{11}^{\beta} + h_{22}^{\beta} \\ &+ \frac{1}{2}\sum_{kl}^{N^c}(J_{kl}^{\alpha\alpha} - K_{kl}^{\alpha\alpha}) + \frac{1}{2}\sum_{kl}^{N^c}(J_{kl}^{\beta\beta} - K_{kl}^{\beta\beta}) \\ &+ \sum_k^{N^c}(J_{k1}^{\beta\beta} - K_{k1}^{\beta\beta}) + \sum_k^{N^c}(J_{k2}^{\beta\beta} - K_{k2}^{\beta\beta}) \\ &+ \sum_{kl}^{N^c} J_{kl}^{\alpha\beta} + \sum_k^{N^c} J_{k1}^{\alpha\beta} + \sum_k^{N^c} J_{k2}^{\alpha\beta} + J_{12}^{\beta\beta} - K_{12}^{\beta\beta} \end{aligned} \tag{A.15}$$

Since we're going to use the energy expression for a restricted calculation where α and β spin orbitals of the same spatial orbital are identical, we don't need to keep track of spin in the energy expression. Then the only difference between the mixed and triplet determinant energies is the $-K_{12}$ in the triplet determinant, since the open shells have electrons of the same spin. The energy of the singlet state, $E_s = 2E_m - E_t$, is

$$\begin{aligned} E_s =& 2\sum_k^{N^c} h_{kk} + h_{11} + h_{22} + \sum_{kl}^{N^c}(2J_{kl} - K_{kl}) + \sum_k^{N^c}(2J_{k1} - K_{k1}) + \\ &+ \sum_k^{N^c}(2J_{k2} - K_{k2}) + J_{12} + K_{12} \end{aligned} \tag{A.16}$$