Constrained Density-Functional Theory—Configuration Interaction

by

Benjamin James Kaduk

B.S. Chemistry, University of Illinois (2007)
B.S. Mathematics, University of Illinois (2007)

Submitted to the Department of Chemistry
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2012

Author .....
Department of Chemistry
May 10, 2012

Certified by .....
Troy Van Voorhis
Associate Professor
Thesis Supervisor

Accepted by .....
Robert W. Field
Chairman, Departmental Committee on Graduate Students

This doctoral thesis has been examined by a
Committee of the Department of Chemistry as follows:

Professor Jianshu Cao .....
Chairman, Thesis Committee
Professor of Chemistry

Professor Troy Van Voorhis .....
Thesis Supervisor
Associate Professor of Chemistry

Professor Robert W. Field .....
Member, Thesis Committee
Haslam and Dewey Professor of Chemistry

Constrained Density-Functional Theory—Configuration Interaction

by

Benjamin James Kaduk

Submitted to the Department of Chemistry
on May 10, 2012, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Chemistry

Abstract

In this thesis, I implemented a method for performing electronic structure calculations, "Constrained Density Functional Theory–Configuration Interaction" (CDFT-CI), which builds upon the computational strengths of Density Functional Theory and improves upon it by including higher level treatments of electronic correlation which are not readily available in Density-Functional Theory but are a keystone of wavefunction-based electronic structure methods. The method involves using CDFT to construct a small basis of hand-picked states which suffice to reasonably describe the static correlation present in a particular system, and efficiently computing electronic coupling elements between them. Analytical gradients were also implemented, involving computational effort roughly equivalent to the evaluation of an analytical Hessian for an ordinary DFT calculation. The routines were implemented within Q-Chem in a fashion accessible to end users; calculations were performed to assess how CDFT-CI improves reaction transition state energies, and to assess its ability to produce conical intersections, as compared to ordinary DFT. The analytical gradients enabled optimization of reaction transition-state structures, as well as geometry optimization on electronic excited states, with good results.

Thesis Supervisor: Troy Van Voorhis
Title: Associate Professor

Acknowledgments

I would like to thank my officemate for these past five years, Timothy Daniel Kowalczyk, for all the time we had together, and the great support and advice he has provided me with during our lasting friendship.

I also thank all of my family, for their support and understanding as my work occasionally grew to swallow up all of my time.

My coworker Jiahao Chen proved to be a valuable resource, providing solid answers and pointers to my sometimes-esoteric questions and musings.

Contents

List of Figures

List of Tables

Chapter 1

Theory

In this chapter, we introduce the relevant background in electronic structure theory, with a heavy focus on density-functional theory leading up to the constrained density-functional theory which features prominently in this thesis. We also discuss wavefunction theory as it leads to excited-state treatments, and close with an outline for the remainder of the thesis.

1.1 Electronic structure fundamentals

Electrons are light, fleeting particles, and as such require a wavelike treatment within quantum mechanics; in this work we only consider electrons bound by the Coulomb potential of nuclear charge into molecules. As such, the Born-Oppenheimer approximation with non-relativistic electrons interacting in a fixed nuclear potential is appropriate, and we begin with the time-independent molecular Schrodinger equation

$$\hat{H}\Psi = E\Psi \tag{1.1}$$

where the electronic wavefunction Ψ is an eigenstate of the molecular Hamiltonian $\hat{H}$ which includes the nuclear potential, electronic kinetic energy, and electron-electron repulsion terms. Textbook[1] electronic structure proceeds to Hartree-Fock (HF) theory, perhaps the conceptually simplest explicit wavefunction theory. A one-particle basis is introduced, from

which one-particle orbitals ϕ_i are constructed as linear combinations of basis functions. The full N-electron wavefunction is a Slater determinant (antisymmetrized product) of the single-particle orbitals, $\Phi = \mathcal{A}(\phi_1(\mathbf{r}) \cdots \phi_N(\mathbf{r}))$ to preserve the fermionic nature of the multi-electron wavefunction. Within the Born-Oppenheimer approximation, the (electronic) Hamiltonian is (in atomic units)

$$\hat{H} = \sum_i -\frac{1}{2}\nabla_i^2 + \hat{v}_n(\mathbf{r}) + \sum_i \sum_{j>i} \frac{1}{\hat{r}_{ij}} \tag{1.2}$$

where i and j index the electrons and $\hat{v}_n$ is the nuclear potential. Hartree-Fock theory assumes an orbital representation and applies a mean-field treatment to this Hamiltonian, folding the two-electron Coulomb interaction into an average potential experienced by each electron. Introducing the one-electron Hamiltonian $\hat{h} = -\frac{1}{2}\nabla^2 + \hat{v}_n$, $\hat{J}$ the classical Coulomb repulsion

$$\hat{J}\phi_i(\mathbf{r}_1) = \sum_j \int d\mathbf{r}_2 \phi_j(\mathbf{r}_2) \frac{1}{\hat{r}_{12}} \phi_j(\mathbf{r}_2)\phi_i(\mathbf{r}_1) \tag{1.3}$$

and $\hat{K}$ the quantum-mechanical exchange:

$$\hat{K}\phi_i(\mathbf{r}_1) = \sum_j \int d\mathbf{r}_2 \phi_j(\mathbf{r}_2) \frac{1}{\hat{r}_{12}} \phi_i(\mathbf{r}_2)\phi_j(\mathbf{r}_1) \tag{1.4}$$

we write to the Hartree-Fock equations for the orbitals and energies

$$\left(\hat{h} + \hat{J} - \hat{K}\right)\phi_i = \epsilon_i \phi_i \tag{1.5}$$

These equations are coupled to each other, and must be solved self-consistently. Within the Hartree-Fock framework, the total energy of the state Φ is the expectation value of the Hamiltonian, $E = \langle \Phi | \hat{H} | \Phi \rangle$. The exact Hamiltonian can be partitioned into kinetic energy, nuclear attraction, Coulomb repulsion, and exchange terms, leading to a natural partitioning

of the energy as $E_{HF} = E_T + E_n + E_J + E_k$, with

$$E_T = -\frac{1}{2}\sum_i \int \phi_i^\dagger(\mathbf{r}_1)\nabla^2\phi_i(\mathbf{r}_1)d\mathbf{r}_1 \tag{1.6}$$

$$E_n = \sum_i \int \phi_i^\dagger(\mathbf{r}_1)\hat{v}_n(\mathbf{r}_1)\phi_i(\mathbf{r}_1)d\mathbf{r}_1 \tag{1.7}$$

$$E_J = \frac{1}{2}\sum_{i,j} \int \phi_i^\dagger(\mathbf{r}_1)\phi_j^\dagger(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\phi_i(\mathbf{r}_1)\phi_j(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.8}$$

$$E_J = -\frac{1}{2}\sum_{i,j} \int \phi_i^\dagger(\mathbf{r}_1)\phi_j^\dagger(\mathbf{r}_2)\frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}\phi_j(\mathbf{r}_1)\phi_i(\mathbf{r}_2)d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.9}$$

Hartree-Fock is not of particular interest to this work in its own right, but rather as a structural framework in which other methods may be developed and understood.

We will return to such methods later; however, we now turn to density-functional theory (DFT). The Hohenberg-Kohn theorem[2] shows that the electronic energy E from equation (1.1) may be determined solely from the (exact) electronic density (by extracting the nuclear geometries and charges from the density and solving the Schrodinger equation), which reduces the dimensionality of the system from $3N$ to just 3. This theoretical breakthrough prompted interest in methods to determine an approximate energy E from an approximate electronic density, that is, density functional methods seeking density functionals $E[\rho(\mathbf{r})]$. There is some work in this space directly, working either by approximating ρ on a grid or by other methods, but we ignore it in favor of the Kohn-Sham (KS) flavor of density functional theory.

We again only repeat those portions of the textbook[3] material which will be useful for later discussions, eliding the importance of the kinetic energy approximations involved, among other things. KS DFT is strikingly similar to HF theory, again involving a one-particle basis set and one-particle orbitals $\phi_i(\mathbf{r})$ which are assembled into a Slater determinant Φ for evaluating various properties. The orbitals are determined from the Kohn-Sham equations, which bear a striking parallel to the Hartree-Fock equations

$$\left(-\frac{1}{2}\nabla^2 + v_n(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}d^3r' + v_{xc}(\mathbf{r})\right)\phi_i = \epsilon_i\phi_i \tag{1.10}$$

The orbitals again must be determined self-consistently, but other than the kinetic energy, the interactions involved arise from effective potentials of noninteracting particle averaged over all the electrons, including the one being acted upon (leading to the so-called "self-interaction error"). The orbitals from the Kohn-Sham equations are used to construct a density $\rho(\mathbf{r}) = \phi_i(\mathbf{r})\phi_i^\dagger(\mathbf{r})$, which is then used as input for the energy functional $E[\rho]$. All current density functionals are approximate, but seek to come close to the exact ("universal") functional $G[\rho]$ whose existence is guaranteed by the Hohenberg-Kohn theorem. The total energy can again be decomposed, now as $E = E_{T'} + E_n + E_J + E_{xc}$, with

$$E_{T'} = -\frac{1}{2}\sum_i \int \phi_i^\dagger(\mathbf{r})\nabla^2\phi_i(\mathbf{r})d\mathbf{r} \tag{1.11}$$

$$E_n = \int v_n(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} \tag{1.12}$$

$$E_J = \frac{1}{2}\int \frac{\rho(\mathbf{r}_1)\rho(\mathbf{r}_2)}{|\mathbf{r}_1 - \mathbf{r}_2|}d\mathbf{r}_1 d\mathbf{r}_2 \tag{1.13}$$

$$E_{xc}[\rho] = G[\rho] - E_{T'}[\rho] - E_n[\rho] - E_J[\rho] \tag{1.14}$$

which makes it clear that the exchange-correlation energy is merely defined to be the difference between the exact energy and the pieces that have a closed-form expression. The exchange-correlation potential v_{xc} which appears in equation (1.10) is just the functional derivative of the exchange-correlation energy, $v_{xc} = \delta E_{xc}/\delta\rho$; as such, KS DFT development therefore focuses on improving the approximations in the linked $E_{xc}[\rho]$ and $v_{xc}(\mathbf{r})$ functions, leading to an abundance of functionals to choose from. In KS-DFT, then, the electronic states are affected by the kinetic energy functional, the nuclear potential, Coulomb repulsion, this effective potential v_{xc}, and any other external potentials that may be applied.

We began this discussion by assuming a one-particle basis but have been referring extensively to orbitals ϕ_i throughout. These orbitals are of course just linear combinations of the one-particle basis functions, and the expansion coefficients are referred to as the MO coefficients. The potentials and interactions described above are assembled into a Fock matrix in the original one-particle basis, whose eigenvectors are the MO coefficient vectors to be

used for the next iteration of the self-consistency algorithm, with corresponding eigenvalues as orbital energies. The self-consistency problem is then written in matrix form,

$$\mathbf{Fc} = \mathbf{Sc}\epsilon \tag{1.15}$$

where $\mathbf{F}$ is the overall Fock matrix, $\mathbf{c}$ are the MO coefficients, $\mathbf{S}$ is the overlap matrix between the one-particle basis functions, and ϵ the orbital energies.

1.2 Constrained Density Functional Theory

CDFT is directly involved in the rest of this thesis, and so we give a substantially more thorough introduction than was needed for HF and KS-DFT.

We now proceed to outline the working equations of CDFT and describe how they can be solved efficiently. External applied potentials can come into play in many sorts of situations; in particular constrained density functional theory (CDFT) may be thought of as applying external potentials to achieve the target constraint values. Development of modern CDFT has benefitted greatly from the foresight of the original presentation of CDFT, which fully anticipated all manner of applications and formalisms.[4] In modern molecular usage, the theory of CDFT has been refined so that constraints are typically phrased in terms of the charge and spin on arbitrary molecular fragments, which are defined in terms of an atomic charge prescription.[5–9] This portrayal allows for multiple constrained fragments, analytical gradients, and efficient determination of the self-consistent constraint potential. In this section we introduce the general theory with emphasis on the formulation in terms of populations. We close the section with a few illustrations of best practices in using constraints to solve chemical problems.

1.2.1 Original CDFT Equations

The first presentation of a constrained DFT formalism is due to Dederichs *et al.*[4] and proceeds as follows. Suppose we seek the ground electronic state of a system subject to

the constraint that there are N electrons in a volume Ω. One can accomplish this by supplementing the traditional DFT energy functional, $E[\rho(\mathbf{r})]$, with a Lagrange multiplier:

$$E(N) = \min_{\rho} \max_{V} \left[E[\rho(\mathbf{r})] + V \left(\int_{\Omega} \rho(\mathbf{r}) d^3r - N \right) \right] \tag{1.16}$$

The addition of a single Lagrange multiplier term $V\left(\int_{\Omega} \rho(\mathbf{r}) d^3r - N\right)$ is sufficient to effect a constrained optimization that yields the lowest-energy state with exactly N electrons in the volume Ω. This would clearly be useful, for example, in looking at the localization of charge around an impurity. Continuing along these lines, one can easily come up with other interesting constraint formulations.[4] One could constrain local d (or f) charge variation in transition (or rare-earth) metals:

$$E(N) = \min_{\rho} \max_{V_d} \left[E[\rho(\mathbf{r})] + V_d \left(\int \rho_d(\mathbf{r}) d^3r - N_d \right) \right] \tag{1.17}$$

or the (net) magnetization:

$$E(N) = \min_{\rho} \max_{H} \left[E[\rho(\mathbf{r})] + H \left(\int_{\Omega} m(\mathbf{r}) d^3r - M \right) \right] \qquad \left[m(\mathbf{r}) \equiv \rho^{\alpha}(\mathbf{r}) - \rho^{\beta}(\mathbf{r}) \right]. \tag{1.18}$$

One could go even further and note that the magnetization in a given system need not have a uniform orientation throughout, so that one could partition the system into magnetization domains with different axes of magnetization. In this case, the magnetization on each domain would become an independent parameter, with the energy $E(\vec{M}_1, \ldots, \vec{M}_N)$ being a function of the constrained parameters.

All of the constraints above can be cast in a unified notation:[5]

$$W[\rho, V; N] \equiv E[\rho] + V \left(\sum_{\sigma} \int w^{\sigma}(\mathbf{r}) \rho^{\sigma}(\mathbf{r}) d^3r - N \right) \tag{1.19}$$

$$E(N) = \min_{\rho} \max_{V} W[\rho, V; N]. \tag{1.20}$$

Here, one introduces a (spin-dependent) weight function, $w^{\sigma}(\mathbf{r})$, that defines the property of

interest. For example, to match equation (1.16), $w^{\alpha}(\mathbf{r}) = w^{\beta}(\mathbf{r})$ would be the characteristic function of Ω. To match equation (1.18), $w^{\alpha}(\mathbf{r}) = -w^{\beta}(\mathbf{r})$ would again be the characteristic function of Ω. In this way, we think of the various constraints as specific manifestations of a single unified formalism.

These core equations have been widely used for determining the U parameter in LDA+U, Anderson, and Hubbard models,[10–25] frequently in combination with the Hund's rule exchange parameter J.[26–40] A survey of the results based upon CDFT finds that virial and Hellmann-Feynman theorems have been given for CDFT,[41] and the theory has been generalized for application to the inverse Kohn-Sham problem.[42] CDFT has found use examining charge localization and fluctuation in the d density of bulk iron,[43] studying localized excitons on the surface of GaAs(110),[44] and constraining core orbital occupations to obtain core excitation energies.[45] Combining Janak's theorem and its integrated version the Slater formula with CDFT yields an efficient method for determining the charge on quantum dots,[46] and using CDFT to constrain orbitals to a fixed atomic form provides a projection operator for use in self-interaction correction (SIC) calculations;[47] the CDFT equations have been reformulated for use with DFTB+ tight-binding models.[48] With this slew of varied applications, the theory of constraining properties of DFT states has proven quite versatile, being applied to study a wide variety of phenomena. In this thesis, we will focus on simultaneously constraining the charge and spin on a particular fragment or fragments of a molecular system, as needed in combination with the promolecule approach (section 2.2.3) for modifying constraint values.

1.2.2 Constrained Observables

There is a great deal of flexibility available for constraining the ground-state density in equation (1.19), since in an unrestricted KS DFT framework an arbitrary constraint may be applied to the integrated population of each spin, over any number of arbitrary regions of space, subject to an arbitrary weighting scheme. In practice, this degree of flexibility is simply overwhelming, and requires some way to streamline the choice of appropriate constraints.

In this spirit, real-space atomic charge schemes have driven much of the modern work with CDFT: they are flexible enough to define a variety of states in accord with chemical intuition, but at the same time compact enough that the number of reasonable constraints is not too large.

First, it is important to note that a variety of commonly used prescriptions for computing the charge on atom A can be cast in the form

$$N_A \equiv \int w_A(\mathbf{r})\rho(\mathbf{r})d^3r. \tag{1.21}$$

Thus, constraining the charge or spin using one of these population prescriptions is just a special case of equation 1.19. The easiest to understand is probably the Voronoi method,[49] which partitions space up into cells Ω_I consisting of all points closest to atom I. The number of electrons on atom A is then

$$N_A \equiv \int_{\Omega_A} \rho(\mathbf{r})d^3r \tag{1.22}$$

which is obviously a special case of equation (1.16). The Becke population scheme is similar:[50] here one defines a weight function, w_A^{Becke}, that is nearly unity inside the Voronoi cell, nearly zero outside and smoothly connects the two limits. The number of electrons on atom A is then

$$N_A \equiv \int w_A^{\text{Becke}}(\mathbf{r})\rho(\mathbf{r})d^3r. \tag{1.23}$$

In a completely different fashion, the Hirshfeld (or Stockholder) partitioning can also be written in terms of atomic weight functions.[51] In the Hirshfeld scheme, one constructs a promolecule density, $\tilde{\rho}(\mathbf{r})$ that is just the sum of (usually spherically averaged) atomic densities, $\rho_A(\mathbf{r})$. One then defines an atomic weight function and number of electrons respectively by:

$$w_A^{\text{Hirshfeld}}(\mathbf{r}) \equiv \frac{\rho_A(\mathbf{r})}{\tilde{\rho}(\mathbf{r})} \qquad N_A \equiv \int w_A^{\text{Hirshfeld}}(\mathbf{r})\rho(\mathbf{r})d^3r. \tag{1.24}$$

Similar constructions apply to the variations on this theme — including Hirshfeld-I[52] and iterated Stockholder[53] — with mild adjustments to the definitions of w_A. It is also in

principle possible to phrase more sophisticated schemes — such as partition theory[54–56] and Bader's atoms-in-molecules approach[57] — in terms of a weight function w_A, although to our knowledge these connections have never been made in the context of CDFT. Finally, there are charge prescriptions (including the popular Mulliken,[58] Löwdin[59] and NBO[60] schemes) that can not be written in terms of the density. In these cases, the charge is defined by partitioning the one-particle density *matrix* (1PDM) which technically goes outside the scope of constrained *density* functional theory. However, in practice it is a simple matter to apply constraints to the 1PDM within the same formalism[5,6] and thus when one constrains Löwdin or Mulliken populations it is still colloquially referred to as CDFT.

With a prescription for atomic charges in hand, one can easily build up a weight, w_F, for the charge on a fragment F, consisting of any group of atoms within a molecule or solid. The charge on the fragment is just the sum of the atomic charges, so that

$$\begin{aligned} N_F &\equiv \sum_{I\in F} N_I = \sum_{I\in F} \int w_I(\mathbf{r})\rho(\mathbf{r})d^3r \\ &= \int \sum_{I\in F} w_I(\mathbf{r})\rho(\mathbf{r})d^3r \equiv \int w_F(\mathbf{r})\rho(\mathbf{r})d^3r \qquad [w_F(\mathbf{r}) \equiv \sum_{I\in F} w_I(\mathbf{r})]. \end{aligned} \tag{1.25}$$

We can thus constrain the number of electrons on any fragment by adding the Lagrangian term

$$V_F \left(\int w_F(\mathbf{r})\rho(\mathbf{r})d^3r - N_F \right) \tag{1.26}$$

to the energy expression. Here N_F is the total number of electrons on the fragment, though for practical calculations the nuclear charge is subtracted off and only the net number of electrons on the fragment ($-q_F \equiv N_F - Z_F$) need be specified as input to the calculation.

For magnetic systems, we would also like to be able to constrain the local spin using population operators. That is, we would like the equivalent of equation (1.18) for subsets of the entire system. To accomplish this, we note that the number of electrons of spin σ ($\sigma = \alpha, \beta$) on F is just

$$N_F^\sigma \equiv \int w_F(\mathbf{r})\rho^\sigma(\mathbf{r})d^3r. \tag{1.27}$$

The net spin polarization (i.e. the local M_S value) is $(N_\alpha - N_\beta)/2$, where the factor of $1/2$ reflects the fact that electrons are spin-$1/2$ particles. We can thus constrain the net magnetization of any fragment by adding the Lagrangian term:

$$H_F \left(\int w_F(\mathbf{r})(\rho^\alpha(\mathbf{r}) - \rho^\beta(\mathbf{r}))d^3r - M_F \right) \tag{1.28}$$

M_F is then the net number of spin up electrons on the fragment, which is the same as twice the M_S value.

Finally, we can apply any number of spin and charge constraints by adding a number of such terms:

$$\begin{aligned} W[\rho, V_F, H_{F'}; N_F, M_{F'}] \equiv E[\rho] &+ \sum_F V_F \left(\int w_F(\mathbf{r})\rho(\mathbf{r})d^3r - N_F \right) \\ &+ \sum_{F'} H_{F'} \left(\int w_{F'}(\mathbf{r})(\rho^\alpha(\mathbf{r}) - \rho^\beta(\mathbf{r}))d^3r - M_{F'} \right) \end{aligned} \tag{1.29}$$

$$E(N_F, M_{F'}) = \min_\rho \max_{V_F, H_{F'}} W[\rho, V_F, H_{F'}; N_F, M_{F'}]. \tag{1.30}$$

The actual form of w_F (and thus the constraint) will depend on the choice of target populations as described above. But it is a trivial matter to write the equations in a manner that is independent of the population, and we will maintain this level of abstraction in what follows.

1.2.3 Choosing a Constraint

Even if we restrict our attention only to charge and spin constraints, in any given application one still has several choices to make about how an appropriate constraint should be defined. What atomic population should be used? Which atoms should be included in the fragment? Does the basis set matter? For the most part, the answers to these questions must be determined on a case-by-case basis either by trial and error or using chemical intuition. However, the literature does contain a number of empirically determined guidelines that can be helpful in practice:

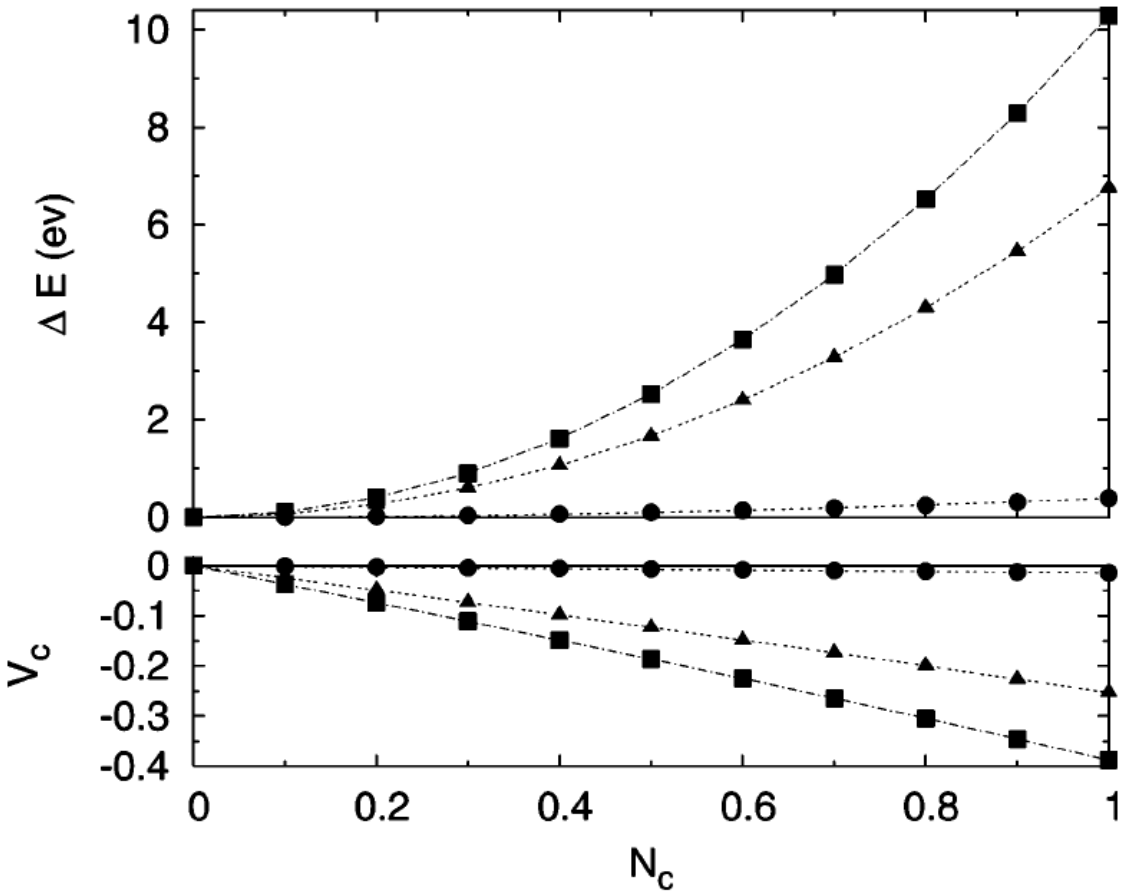

Figure 1-1: The energy and constraint potential as a function of charge separation in N_2 with different charge prescriptions. Squares: Becke population; triangles: Löwdin population; dots: Mulliken population. Calculations performed using B3LYP in a 6-31G* basis set. Reprinted with permission from reference 6. Copyright 2006 American Chemical Society.

- **Mulliken populations are not reliable.** One abiding rule is that Mulliken populations are unrealistic in CDFT. For example, in Figure 1-1, Mulliken populations spuriously predict that separating charge in dinitrogen to obtain the N^+N^- configuration should only require a fraction of an eV, whereas all other prescriptions predict energies on the order of 5-10 eV. This failure can be linked to the ability of Mulliken populations to become negative in some regions of space.[61]

- **When diffuse functions are involved, density based prescriptions are more stable.** Here again, the observation is tied to a known weakness of an atomic population scheme: AO-based schemes (like Löwdin, Mulliken or NBO) tend to get confused when diffuse functions are added.[62] In the worst cases, this fault keeps Löwdin-CDFT energies and properties from converging as the size of the basis set is increased. Such a case is illustrated in Table 1.1, which presents the electronic coupling (discussed in section 1.4.1) between benzene and chlorine at two different separations for a variety of basis sets. Clearly the Löwdin result shows an unreasonably large increase as the basis size increases, while the density-based Becke prescription shows fast convergence. Real-space population schemes such as the Becke weighting scheme and Hirshfeld partitioning correct for the broad spread of diffuse basis functions, giving good results for CDFT.[6,63]

Table 1.1: Diabatic coupling (in mHartree) for electron transfer from benzene to Cl.[7]

d (Å)	0.604		1.208	
	Löwdin	Becke	Löwdin	Becke
6-31G	21.3	58.2	30.1	65.9
6-31G(d)	21.0	56.9	29.9	64.8
6-31+G(d)	39.6	46.7	46.1	53.9
VDZ-ANO	95.3	48.8	94.0	56.1

- **Larger fragments give more consistent results.** This conclusion has mainly been drawn from the application of CDFT to predict exchange couplings in magnetic organometallic compounds, where there is a wealth of experimental data to compare to.[64] The qualitative picture is that all excess spin resides on the metal atoms. However, in practice, constraining the net spin of the metal atoms alone using any of the standard schemes gives unreasonable exchange couplings. The most reliable results are obtained if the fragments are made as large as possible; if there are two metals (A and B) then every atom in the molecule is assigned either to fragment A or fragment B, even if there is thought to be no net magnetization on that fragment. Likewise,

for charge transfer, making the fragments large helps stabilize the excess charge, e.g. constraining a metal center and its ligands (instead of just the metal), or not leaving an unconstrained "bridge" in a fully conjugated aromatic charge-transfer system. Making the constrained region too small can cause the constraint to be artificially too strong; a charged metal center really will delocalize charge to its ligands (Figure 1-2), and a charge-transfer state in a conjugated system will delocalize the electron and hole as much as possible to stabilize itself. By making the CDFT constraint region as large as possible, the minimum perturbation needed to enforce the constraint can be applied, with the system naturally seeking the correct level of localization. It is important to emphasize that adding "spectator" atoms to a fragment does not necessarily place any charge or spin on the spectator; adding the atom to the fragment merely means that the variational CDFT optimization *can* place additional charge or spin on that atom, not that it will. For example, in Figure 1-2, when half of the bridge is added to each fragment, not all of the bridge carbons will have extra charge.

- **When possible, constrain charge and spin together.** Suppose you were interested in charge transfer between C_{60} and C_{70} (i.e. $C_{60}^{+} \cdots C_{70}^{-}$). You could generate this state in one of two ways: either constrain only the charge (e.g. $q_{C_{60}} = +1$) or the charge and spin (e.g. $q_{C_{60}} = +1$ and $M_{C_{60}} = 1$). In many cases these two routes will give nearly identical answers (as long as the calculations are spin-unrestricted). However, in the cases where they differ significantly, it can often be the case that constraining the charge leads to a state that still has significant overlap with the ground state. This phenomenon is known as "ground state collapse" and generally leads to erroneous results for energetics.[66] Thus, to be on the safe side, it seems best to constrain both charge and spin rather than just charge alone.

- **There can be many equivalent ways of specifying the same state.** Returning to the $C_{60}^{+}...C_{70}^{-}$ example, because the overall charge on the system is fixed, specifying $q_{C_{60}} = +1$ or $q_{C_{70}} = -1$ would obtain exactly the same answer in CDFT. Alternatively, requiring that $q_{C_{60}} - q_{C_{70}} = +2$ would also give the same result. These observations are

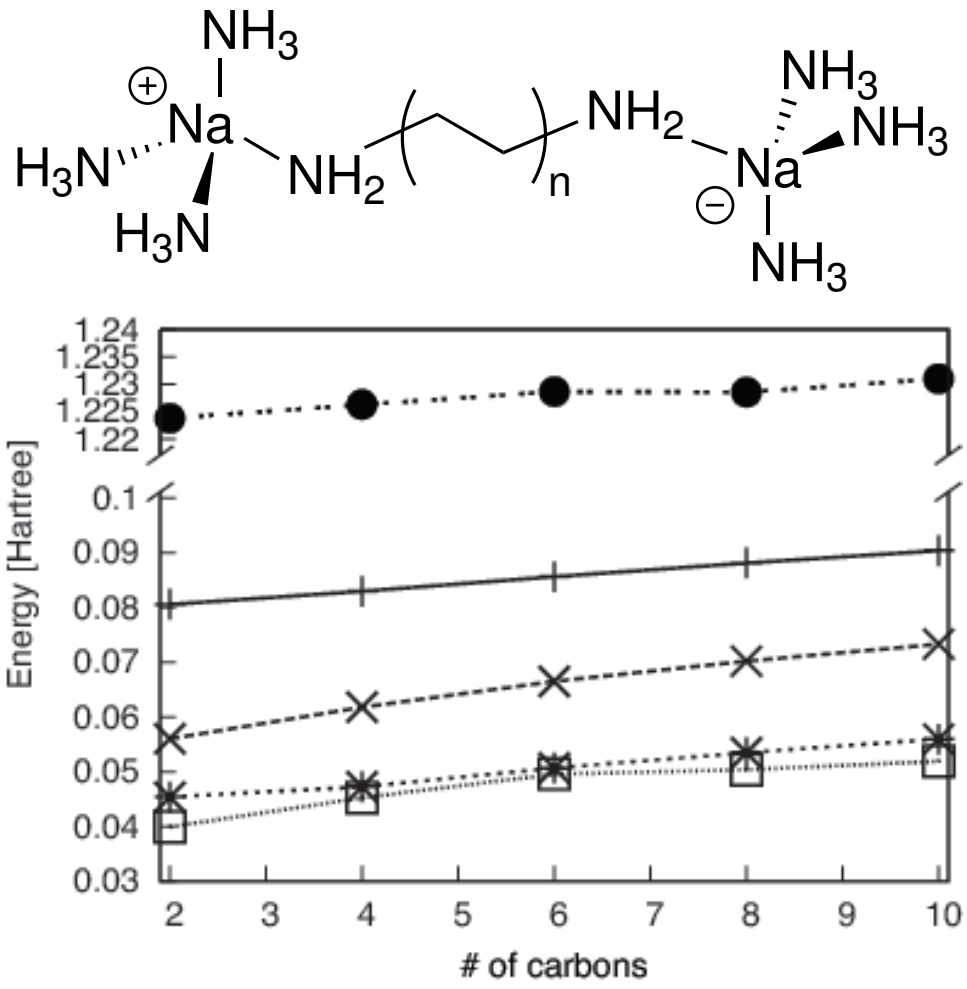

Figure 1-2: The energy behavior of $[Na(NH_3)_3]^+H_2N(CH_2)_nNH_2[Na(NH_3)_3]^-$ with the constraint applied to just the metal atoms (•), the $Na(NH_3)_3$ groups (+), the metal and ammonias and the amine group of the bridge (×), splitting the complex in two down the middle of the bridge (∗), or with a promolecule-modified constraint applied to the $Na(NH_3)_3$ groups (□). Energy differences are measured with respect to the ground-state DFT energy for each system, and plotted as a function of the number of carbons in the alkyl amine. Geometries are constructed with bond lengths and angles corresponding to the optimized geometry of the eight-carbon system. The metal-only constraint is comically overstrong (note the broken y-axis), while expanding the constraint region to include the ligands or the ligands plus bridge leads to physically plausible results. The constraint in (∗) is a weaker constraint than all the other curves except the promolecule-corrected constraint on the sodium and ammonias; this is because when the system is literally divided in two, only one constraint region is needed — the other partitionings require that one region is constrained to +1 charge and the other to −1, with an implicit constraint that the bridge is neutral. Reprinted with permission from reference 65. Copyright 2012 American Chemical Society.

general: it is always mathematically equivalent to describe the system with constraint N_A on A and with constraint $N_B \equiv N - N_A$ on a B defined as the set complement of A. The ability to add and subtract constraints in this manner is reminiscent of the elementary row operations of linear algebra, allowing for different presentations of equivalent physical constraints. The charge difference constraint illustrated above has been used rather extensively[6,63] because, in cases where the constraint regions do not cover all of space, the charge difference constraint is insensitive to fluctuations in the overall charge.

- **When donor and acceptor are very close to one another, CDFT may fail.** When atoms are bound together in molecules, there is no perfect prescription for assigning atomic charges: at some point any method for dividing delocalized charge becomes arbitrary. It is particularly challenging when atoms are very close to each other, e.g. the two nitrogen atoms in N_2, illustrated in Figure 1-1. Here, even when two reasonable charge prescriptions (Löwdin and Becke) are used, the energy of the N^+N^- state varies by more than 3 eV. This is clearly an unacceptably large error for chemical purposes and trying more population prescriptions will not fix the problem. There is simply no unambiguous way to apportion the charge in N_2 to the different nitrogen atoms. In section 2.2.3 we will discuss how these problems can be mitigated somewhat by using fragment densities, but they cannot be entirely ignored.

Thus, while defining an appropriate constraint is not a trivial task, in practice we at least have some empirical guidelines of what to do and what not to do when we approach a new problem with CDFT.

1.2.4 Implementation

A full implementation of CDFT needs to find the density which obeys the specified charge/spin constraints at SCF convergence. That is to say, it needs to solve for the stationary points of the Lagrangian in equation (1.30). Ideally, we would like to solve these equations with

approximately the same computational cost as a regular KS-DFT calculation. Toward that end, we re-write equation 1.30 as

$$\begin{aligned} E(N_k) &= \min_{\rho} \max_{V_k} W[\rho, V_k; N_k] \\ &= \min_{\rho} \max_{V_k} \left[E[\rho] + \sum_k V_k \left(\int \sum_\sigma w_k^\sigma(\mathbf{r}) \rho^\sigma(\mathbf{r}) d^3r - N_k \right) \right] \end{aligned} \tag{1.31}$$

where the index k indexes charge and spin constraints: for charges $V_k \equiv V_F$ and $w_k^\alpha = w_k^\beta = w_F$, while for spins $V_k \equiv H_F$ and $w_k^\alpha = -w_k^\beta = w_F$. This notation obfuscates the meaning somewhat, but makes the equations uniform. Recall that the DFT energy expression is defined by

$$E[\rho] = \sum_\sigma \sum_i^{N_\sigma} \left\langle \phi_{i\sigma} \middle| -\frac{1}{2}\nabla^2 \middle| \phi_{i\sigma} \right\rangle + \int v_n(\mathbf{r}) \rho(\mathbf{r}) d^3r + J[\rho] + E_{xc}[\rho^\alpha, \rho^\beta] \tag{1.32}$$

where the terms on the right hand side are, in order, the electronic kinetic, electron-nuclear attraction, Coulomb and exchange-correlation energies. Requiring that equation (1.31) be stationary with respect to variations of the orbitals, subject to their orthonormality, yields the equations:

$$\left(-\frac{1}{2}\nabla^2 + v_n(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' + v_{xc}^\sigma(\mathbf{r}) + \sum_k V_k w_k^\sigma(\mathbf{r}) \right) \phi_{i\sigma} = \epsilon_{i\sigma} \phi_{i\sigma} \tag{1.33}$$

with Hermitian conjugate for $\phi_{i\sigma}^*$. These equations are just the standard Kohn-Sham equations with the addition of some new potentials, which may be thought of as the external applied potentials needed to enforce the constraints of CDFT. These potentials are proportional to the Lagrange multipliers, which illustrates the physical mechanism by which CDFT controls charges and spins: it alters the potential in such a way that the ground state in the new potential satisfies the desired constraint. Another way to say it is that the excited state of the unperturbed system can be approximated by the ground state of the system in the presence of the constraining potential. Thus, CDFT takes the fact that the KS approach is

exact for any potential and exploits it to obtain information about nominally inaccessible excited states.

However, these constraint potentials are not yet fully specified — though the w_k are given as parameters, the Lagrange multipliers V_k are only implicitly defined by the constraints on the fragment charges and spins. These constraints become clear when we attempt to make W stationary with respect to the V_i:

$$\frac{dW}{dV_k} = \sum_{\sigma} \sum_{i}^{N_\sigma} \left(\frac{\delta W}{\delta \phi_{i\sigma}^*} \frac{\partial \phi_{i\sigma}^*}{\partial V_k} + \text{cc} \right) + \frac{\partial W}{\partial V_k} \tag{1.34}$$

$$= \sum_{\sigma} \int w_k^\sigma(\mathbf{r}) \rho^\sigma(\mathbf{r}) d^3r - N_k \tag{1.35}$$

$$= 0 \tag{1.36}$$

where the eigencondition $\delta W / \delta \phi_{i\sigma}^* = 0$ has been used. Note that only the constraint with index k remains after differentiation, even when multiple constraints are imposed on the system, and the stationary condition of the derivative being zero enforces the desired charge/spin constraints.

The separate conditions of equations (1.33) and (1.36) imply that V_k and ρ must be determined self-consistently to make W stationary. This is somewhat daunting, as the Lagrangian optimization is typically only a stationary condition — that is, it is *not* typically a pure maximization or minimization. As a practical matter, it is much more difficult to locate indefinite stationary points than maxima or minima. For example, it is significantly harder to find a transition state (an indefinite stationary point) than an equilibrium structure (a minimum). However, even though the CDFT stationary point is not a maximum or a minimum, it is easy to locate, because one can show that the desired solution is a *minimum* with respect to ρ and a *maximum* with respect to V_k.[5,6] Thus, the stationary point can be solved for via alternating between minimization along one coordinate (the density) followed by maximization along the others (the potentials).

To see this, note that for any fixed V_k, equations (1.33) determine a unique set of orbitals,

$\phi_i[V_k]$. These orbitals define a density $\rho[V_k]$, which can then be used as input to W. In this manner, one can think of W as a function *only* of V_k: $W(V_k)$. We can work out the second derivatives of this function:[6]

$$\frac{\partial^2 W}{\partial V_k \partial V_l} = \sum_\sigma \sum_i^{N_\sigma} \int w_k^\sigma(\mathbf{r}) \phi_{i\sigma}^*(\mathbf{r}) \frac{\delta \phi_{i\sigma}(\mathbf{r})}{\delta[V_l w_l^\sigma(\mathbf{r}')]} w_l^\sigma(\mathbf{r}') d^3r d^3r' + \text{cc} \tag{1.37}$$

$$= \sum_\sigma \sum_i^{N_\sigma} \int w_k^\sigma(\mathbf{r}) \phi_{i\sigma}^*(\mathbf{r}) \sum_{a \neq i} \frac{\phi_{a\sigma}^*(\mathbf{r}') \phi_{i\sigma}(\mathbf{r}')}{\epsilon_{i\sigma} - \epsilon_{a\sigma}} \phi_{a\sigma}(\mathbf{r}) w_l^\sigma(\mathbf{r}') d^3r d^3r' + \text{cc} \tag{1.38}$$

$$= 2 \sum_\sigma \sum_i^{N_\sigma} \sum_{a > N_\sigma} \frac{\langle \phi_{i\sigma} | w_k^\sigma | \phi_{a\sigma} \rangle \langle \phi_{a\sigma} | w_l^\sigma | \phi_{i\sigma} \rangle}{\epsilon_{i\sigma} - \epsilon_{a\sigma}} \tag{1.39}$$

where first-order perturbation theory has been used in the evaluation of the functional derivative $\delta \phi_{i\sigma}(\mathbf{r}) / \delta[V_l w_l^\sigma(\mathbf{r}')]$. The index i only covers the occupied orbitals of the constrained state, whereas the index a need only cover the virtual orbitals, as the summand is antisymmetric in i and a. This Hessian matrix is nonpositive definite because[6]

$$\sum_{k,l}^m V_k \frac{\partial^2 W}{\partial V_k \partial V_l} V_l = 2 \sum_\sigma \sum_i^{N_\sigma} \sum_{a > N_\sigma} \frac{\left| \langle \phi_{i\sigma} | \sum_{k=1}^m V_k w_k^\sigma | \phi_{a\sigma} \rangle \right|^2}{\epsilon_{i\sigma} - \epsilon_{a\sigma}} \leq 0 \tag{1.40}$$

This holds because the KS method chooses the lowest-energy eigenstates as the occupied orbitals, so for every occupied orbital i and virtual orbital a, $\epsilon_{i\sigma} \leq \epsilon_{a\sigma}$. Thus, the overall Hessian product is nonpositive, as desired, giving a stationary point as a maximum.

Having worked out the second derivatives, we see two features that simplify the CDFT optimization procedure. First, the condensed version of W is globally concave in the V_k, giving a *unique* fixed point which satisfies all the applied constraints. Thus, there can be no confusion about local versus global maxima. Second, since both the first and second derivatives of $W(V_k)$ are easily computed, rapidly converging algorithms such as Newton's method can be used to locate its stationary point. Convergence to the constrained SCF minimum can thus be achieved by means of a nested-loop algorithm with outer SCF loop and inner constraint loop. The outer loop closely resembles a normal DFT calculation, with SCF iterations being performed to optimize the orbitals. Within each step of the outer loop, a second

loop of microiterations is performed to determine the Lagrange multipliers V_k that make the density satisfy the charge and spin constraints (equations (1.25) and (1.27)). Because the V_k contribute to the Fock matrix, the orbitals must be redetermined by diagonalization of the Fock matrix at each microiteration step. Fortunately, the V_k contribution to the Fock matrix is easy to calculate and a full build with exchange and correlation contributions is not necessary, making the microiterations relatively cheap for atom-centered basis sets. After the first few iterations of the outer loop, it is common for the inner loop to converge after only two or three microiterations. Essentially all available SCF codes use a convergence accelerator, such as direct inversion in the iterative subspace (DIIS).[67,68] Since CDFT introduces an extra layer of microiterations at each SCF step, care is needed in incorporating CDFT into existing SCF codes so as to not interfere with these accelerators. DIIS keeps historical Fock matrices for several SCF iterations, and extrapolates a new Fock matrix from them in order to generate MO coefficients for the next SCF iteration. Since the CDFT microiterations add a constraint potential to the Fock matrix to determine MO coefficients, but use the unconstrained Fock matrix for energy determination, both unconstrained and constrained Fock matrices must be retained. The extrapolation coefficients determined from the constrained Fock matrices are then applied to the unconstrained matrices to yield an initial unconstrained Fock matrix for the next round of CDFT microiterations.

It is important to note that at stationarity, the Lagrangian, W (equation (1.31)), is equal to the physical energy of the system, E (equation (1.32)). The energy in the presence of the constraining potentials $V_k w_k$ is then a form of free energy,

$$F = E + V_{\mathrm{tot}} N_{\mathrm{tot}} + V_{\mathrm{spin}} M_{\mathrm{spin}}. \tag{1.41}$$

In accord with this free energy picture, we obtain the thermodynamic relations

$$\frac{dE(N_k)}{dN_k} = -V_k \qquad \text{and} \qquad \frac{dF(V_k)}{dV_k} = N_k \tag{1.42}$$

reflecting that E is a natural function of N_k but F is a natural function of V_k. It also follows

that $d^2E/dN_k^2 = -(d^2W/dV_k^2)^{-1}$, so that the concavity of $W(V_k)$ implies convexity of $E(N_k)$, an important physical condition.

In addition to energy derivatives with respect to the internal parameters V_i and N_i, we may also wish to compute derivatives of the energy with respect to external parameters such as nuclear position. Such analytical gradients have been implemented for CDFT, making possible *ab initio* molecular dynamics on charge-constrained states and parameterizations of Marcus electron transfer theory therefrom.[5,6,63,65,69–71] Implementation of analytical gradients for the CDFT-CI method which underlies this thesis will be presented as chapter 4; CDFT-CI gradients of course rely on CDFT gradients. Consider the problem of computing the derivative of the electronic energy (equation (1.32)) with respect to the position of nucleus A. In addition to obeying $E[\rho_{\mathrm{CDFT}}] = W[\rho_{\mathrm{CDFT}}, V_k^{\mathrm{CDFT}}, N_{\mathrm{CDFT}}]$ at convergence, W has the additional property that it is variational with respect to both ρ and the V_i (in contrast to $E[\rho_{\mathrm{CDFT}}]$ which is not even a stationary point of the energy), which allows the use of the Hellmann-Feynman theorem, writing

$$\nabla_A W = \nabla_A E + \sum_i V_i \rho \nabla_A w_i \tag{1.43}$$

The first term is the standard gradient for unconstrained calculations, which includes the Hellmann-Feynman force, Pulay force, and terms from change in DFT integration grid with nuclear displacement; the second term represents the extra force due to the constraint condition on the density. The form of this term is necessarily dependent on the form of the population operator w used to define the constraint; for the Becke population scheme, these terms have been computed in reference 72. With a Mulliken or Löwdin treatment of population, which depends on the AO overlap matrix, this term has a more complicated form; reference 8 performs the calculation for the Löwdin scheme. Oberhofer and Blumberger's plane-wave CDFT implementation using Hirshfeld's population scheme has also implemented analytical gradients; their expressions for the weight constraint gradient is in Appendix B of reference 63.

Finally, we note that we have focused here on the implementation of CDFT in localized

orbital codes, but the method can equally well be implemented in plane-wave codes.[63] The primary difference is in the cost tradeoff — whereas diagonalization of the KS Hamiltonian is cheap in localized orbitals it is expensive for plane waves. Thus the relative cost of the microiterations is somewhat higher in a plane-wave-based scheme, but the SCF iterations can be significantly faster, particularly for pure functionals applied to condensed phase problems.

1.3 Configuration interaction

We previously gave some basic introduction to HF and KS-DFT, but of course there are many more-modern methods which have been developed (and are still being developed) on top of the basic structures. Since CDFT-CI is a configuration interaction method, we take some time to talk about the more standard CI methods and mention briefly other related techniques which we will use as references for comparing numerical results.

To give a sense of the evolution of these methods, we note that the difference between the exact energy and the Hartree-Fock energy is defined to be the "correlation energy", corresponding to the correlations between different electronic coordinates which are present in the exact wavefunction but not describable in the Slater determinant form of the Hartree-Fock wavefunction. This correlation energy is sometimes further divided into "static" and "dynamic" pieces as we describe shortly, though there is no rigorous separation or definition of these terms. The Hartree-Fock wavefunction is used as a starting point for higher-level methods which endeavor to recover more and more of the correlation energy; its structure as a Slater determinant Ψ of single-particle orbitals ϕ_i makes such extensions relatively straightforward. We introduce a slight change of notation, calling the Hartree-Fock ground state Ψ_0, for reasons that will shortly become clear, and note that this Slater determinant is just a single configuration; in some cases, a multiconfigurational wavefunction is necessary to describe the system correctly. This is the case during (homolytic) bond dissociation, and in general a multiconfigurational treatment is quite beneficial when there are multiple states low-lying in energy; such cases are generally considered to be examples of static correlation; dynamic correlation is then "everything else". Extensions to HF add more

configurations to the effective wavefunction, using the one-particle orbitals determined from the HF calculation to construct a new configuration basis of *many-particle* wavefunctions, and constructing a multiconfigurational wavefunction $\Psi = \sum_k c_k \Psi_k$ from that basis. For ground states, the variational theorem guarantees that this will result in an improved energy, though the situation for excited states is a bit more complicated.

In the extreme case the configuration basis would consist of all possible combinations of one-particle orbitals into N-particle Slater determinants; the number of such configurations is on the order of $\binom{M}{N}$ where M is the size of the original one-particle basis and N is the number of electrons. For a particular basis set, this quantity scales exponentially with the number of atoms, which imposes a strict limit on the size of systems for which this treatment is reasonable (a bit over 10 atoms, with current technology).

Given that the HF ground state is expected to be a reasonable approximation to the exact wavefunction (or at least not too bad), it is reasonable to use it as a reference state from which to describe other configurations in the multi-particle-wavefunction basis. These other configurations are then represented as the HF wavefunction with one or more orbitals ϕ_i replaced in the Slater determinant by previously unoccupied orbitals ϕ_a, a "substitution" from orbital i to orbital a. This configuration can then be written as Ψ_i^a, a "double substitution" with orbital a replacing i and b replacing j as Ψ_{ij}^{ab}, and so forth.[a] The set of all possible combinations of one-particle orbitals into N-particle configurations would then include the HF solution, as well as single, double, ... substitutions up to the total number of electrons in the system. This case is known as full configuration interaction, since it includes all possible configurations of the one-particle orbitals (precisely the $\binom{M}{N}$ from above); it will in this sense give the exact energy within a particular one-particle basis set.

Having described the configuration portion, it remains to describe the interactions in the CI method. With a general form for the wavefunction being written as

$$\Psi = C_0 \Psi_0 + \sum_{ia} C_{ia} \Psi_i^a + \sum_{ijab} C_{ijab} \Psi_{ij}^{ab} + \cdots \tag{1.44}$$

[a] We are perhaps fortunate that methods involving octuple substitutions have not come into common use.

(here Ψ_0 is the HF wavefunction), the energy will be $\langle\Psi|\hat{H}|\Psi\rangle$ (presuming the expansion coefficients C are such to give a normalized wavefunction). The variational theorem guarantees that the minimum-energy wavefunction will be an eigenstate of the Hamiltonian; however, it is conceptually useful to rephrase the energy problem as an explicit eigenvalue problem. We already referred to the configurations Ψ_0, Ψ_i^a, ... as a "configuration basis", and thus it is natural to expect to write the Hamiltonian in this basis. This representation of the Hamiltonian has interesting sparseness properties in its off-diagonal elements; Ψ_0 only couples to states Ψ_{ij}^{ab} (but not Ψ_i^a), and in general two configurations will only couple if they differ by two or fewer orbitals (since the Hamiltonian is a two-electron operator and the orbital basis is orthonormal). That is, the element $\langle\Psi_i^a|\hat{H}|\Psi_{ijkl}^{abcd}\rangle$ is always zero, and so forth. Thinking of the Hamiltonian as having block structure, with the Ψ_0 block, single substitutions, double substitutions, triples, etc., this makes the Hamiltonian a banded matrix with two bands aside the diagonal. With these interactions in place, the CI energy is just the lowest eigenvalue of the Hamiltonian in this basis of configurations.

Taking a brief excursion away from the ground state, we note that since the single substitutions do not couple directly to the Ψ_0, one can then approximate the energy of the first excited state as the lowest eigenvalue of just the singles-singles block of the Hamiltonian. This is known as configuration-interaction singles (CIS), and is an old and fairly inexpensive method for approximating excited-state energies — we will refer back to CIS results when we investigate excited-state PESs in chapter 4. Returning to the ground state, the non-interaction of the singles with Ψ_0 means that the first correction to the ground-state energy comes when double substitutions are included. ($\langle\Psi_{ij}^{ab}|\hat{H}|\Psi_i^a\rangle$ is nonzero, though, so the single substitutions are included as well.) This is known as CISD, with obvious extension to CISDT, CISDTQ, etc.. However, since only the double substitutions couple directly to Ψ_0 and including the triple substitutions is more computationally expensive, CISD remains the most popular.

Since this class of CI methods includes *all* substitutions of a given form, the computational burden can be quite substantial. One might question the physical relevance of

substitutions from orbitals deep in the core, or substitutions to very-high energy orbitals. As such, an alternate CI method to those mentioned above exists, which labels some orbitals as being always occupied and some orbitals as being always unoccupied, and considers only configurations within the remaining orbitals, which are said to form an "active space". Because the set of possible substitutions is thus drastically reduced, it becomes possible to consider all possible configurations within the active space, thus making this a "complete active space" method. With the limited active space, though, the orbitals should be optimized simultaneously with the expansion coefficients, leading to the CASSCF (or just CAS) method. Calculations are described by the number of electrons and the number of orbitals in the active space, as in a CAS(3,4) calculation of three electrons in four orbitals. Because CAS only includes a relatively small number of configurations, it treats static correlation better than dynamic correlation; therefore a perturbation theory correction to the CASSCF energy (very similar to Møller-Plesset perturbation theory) is frequently added to include dynamic correlations, leading to the CASPT2 method.

However, all of these CI methods (except full CI) share a common systematic deficiency: they are not size consistent. That is, if the energy of system A is E_A, then considering a system B which consists of two identical copies of A at infinite separation from each other, the energy E_B will *not* adhere to $E_B = 2E_A$ for these methods! Though this particular situation does not actually occur in practice (and if it did, CAS could be adapted to use an appropriately scaled active space and remain size consistent), it is still enough to indicate a systematic problem as these methods are applied to larger and larger systems. To address the issue, coupled-cluster methods have been developed[73] which are similar to CI methods in including various levels of substitution but have the necessary size-consistency property. The particular structure of coupled-cluster methods is not relevant to the rest of this thesis, merely that they should be more accurate versions of CI methods, so we do not further treat coupled-cluster theory in this introduction.

1.4 CDFT–configuration interaction

These configuration-interaction methods which originate from HF theory all involve a common set of orbitals, and as such they are obliged to use a fairly large number of N-electron basis states in order to extract accurate energies from the CI matrix. CDFT provides a route for the chemist to explicitly specify particular N-electron states to be included as a basis for the CI matrix, and to phrase these N-electron states in terms of chemically relevant charge and spin distributions. These basis states include dynamic correlation directly through their DFT XC functional, and because each state has its own distinct set of orbitals, they may be fully relaxed in a self-consistent fashion above and beyond the linear response of orbital substitution.

More concretely, CDFT-CI is constructed as follows: Given a system with inherent multiconfigurational nature, e.g., a dissociating system with significant ionic and covalent character such as LiF, we introduce multiple configurations Φ_{ionic} and Φ_{neutral} corresponding to those different limiting cases. The intermediate regime might then be described as

$$\Psi_{\text{stretched}} = C_1\Phi_{\text{ionic}} + C_2\Phi_{\text{neutral}} \tag{1.45}$$

$$C_1^2 + C_2^2 = 1 \tag{1.46}$$

The C_i form the CI vector of a Hamiltonian matrix, in the basis of the (as-yet unspecified) Φ_{ionic} and Φ_{neutral}. CDFT provides an easy framework to produce these chemically intuitive states by enforcing a combination of charge and spin constraints to the system. The CDFT energies then form the diagonal elements of this Hamiltonian, whilst the off-diagonal elements are populated by the coupling elements between the states, section 1.4.1.[7,9,74–76]

1.4.1 Evaluating CDFT Couplings

Given two electronic states $|\Psi_1\rangle$ and $|\Psi_2\rangle$, the coupling between them is just the matrix element of the Hamiltonian,

$$H_{12} = \langle\Psi_1|\hat{H}|\Psi_2\rangle \tag{1.47}$$

It is a bit difficult to see how equation (1.47) can be computed in the context of CDFT. The exact expression is written in terms of the wavefunctions for donor and acceptor, but KS theory only gives us the *density* of each state. The KS wavefunctions, $|\Phi_I\rangle$, are fictitious determinants that are constructed to give the correct density. Hence, in practice we need some approximate (but hopefully accurate) prescription for computing the coupling between two CDFT states. The most common prescription for this task was provided in reference 7. Here, we note that if the diabatic states are defined by constraints, then neither of the exact wavefunctions, $|\Psi_I\rangle$, is an eigenstate of the Hamiltonian. Rather they are eigenstates of $\hat{H}$ plus the relevant constraining potential $V_I w_I(\mathbf{r})$:

$$\left[\hat{H} + V_I w_I(\mathbf{r})\right] |\Psi_I\rangle = F_I |\Psi_I\rangle. \tag{1.48}$$

where F_i are considered as the free energy in the presence of the constraining potential per equation (1.41). We can therefore manipulate equation (1.47) to read:

$$\begin{aligned}
H_{12} &= \left\langle \Psi_1 \middle| \hat{H} \middle| \Psi_2 \right\rangle && (1.49)\\
&= \left\langle \Psi_1 \middle| \frac{\hat{H} + V_1\hat{w}_1(\mathbf{r}) + \hat{H} + V_2\hat{w}_2(\mathbf{r})}{2} - \frac{V_1\hat{w}_1(\mathbf{r}) + V_2\hat{w}_2(\mathbf{r})}{2} \middle| \Psi_2 \right\rangle && (1.50)\\
&= \left\langle \Psi_1 \middle| \frac{F_1 + F_2}{2} - \frac{V_1\hat{w}_1(\mathbf{r}) + V_2\hat{w}_2(\mathbf{r})}{2} \middle| \Psi_2 \right\rangle && (1.51)\\
&= \frac{F_1 + F_2}{2} S_{12} - \left\langle \Psi_1 \middle| \frac{V_1\hat{w}_1(\mathbf{r}) + V_2\hat{w}_2(\mathbf{r})}{2} \middle| \Psi_2 \right\rangle && (1.52)
\end{aligned}$$

Thus, the coupling only requires the free energies F_I, the overlap between the states and the matrix elements of a one-electron potential between the states. This is certainly simpler than the many electron matrix element we started with, but it still requires the (unknown) wavefunctions $|\Psi_I\rangle$. Hence, at this point we approximate the true wavefunctions by their KS surrogates ($|\Psi_I\rangle \approx |\Phi_I\rangle$) to arrive at a formula for the CDFT diabatic coupling:

$$H_{12} \approx \frac{F_1 + F_2}{2} S_{12}^{\mathrm{KS}} - \left\langle \Phi_1 \middle| \frac{V_1\hat{w}_1(\mathbf{r}) + V_2\hat{w}_2(\mathbf{r})}{2} \middle| \Phi_2 \right\rangle \tag{1.53}$$

Approximating the exact wavefunction with the appropriate KS determinant is an uncontrolled approximation and it must be tested in practice. As a whole the approximation holds up well, but it is obvious that a more rigorous definition of the diabatic coupling in CDFT would be a significant discovery.

CDFT is not unique in providing a prescription for the diabatic coupling, and we can use some of these other coupling prescriptions to attempt to assess the accuracy of the CDFT coupling formula. The square of the electronic coupling is proportional to the rate of transitions between the two diabatic states via Fermi's golden rule. When considering charge transfer between a spatially separated donor and acceptor, the rate is primarily governed by tunneling, which is expected to decay exponentially with the separation. Thus, one sanity check for any coupling prescription is to verify that it decays exponentially at large separations.

Directly computing the coupling element between diabatic states is far from a unique route to couplings; on the other end of the spectrum are methods that compute diabatic couplings directly from a set of adiabatic states. There are a large number of such methods;[77–83] the generalized Mulliken-Hush (GMH) prescription[84,85] is perhaps the most widely used, seeing heavy use for electron transfer problems. It is often taken as the definitive reference method for computing diabatic couplings (see Figure 1-3). The GMH method defines diabatic states as the eigenstates of the dipole moment operator in the basis of the low-lying adiabatic states. This makes physical sense: the eigenvalues of the dipole will be the extreme values and the desired neutral and CT states will have very small and very large dipole moments, respectively. Thus, given a set of N adiabatic states (e.g. from CASSCF theory), GMH first constructs matrix elements of the dipole operator and then diagonalizes this matrix in the basis of the adiabatic states.[84,85] GMH also allows for multiple diabats to have charge localized on a given site, which are forced to be locally adiabatic with respect to each other as the assumption of zero transition dipole moment is not reasonable in that case.[85] The GMH diabats are automatically orthogonal, and so the diabatic coupling is obtained by transforming the diagonal adiabatic Hamiltonian into the basis of dipole eigenstates; the

physical coupling(s), H_{ab}, are just the off-diagonal elements of the transformed Hamiltonian.

The GMH and CDFT couplings are very different in execution: the former requires some pre-existing route to adiabatic excited states (e.g. TD-DFT) while the latter only requires ground state calculations in alternative potentials. In GMH, diabatic states and their couplings are deduced from the adiabats, while in CDFT diabatic states are constructed directly. Finally, in GMH there is a clear route toward exact diabatic couplings (by improving the adiabatic excited states) while in the latter the route toward exact couplings is somewhat murky. The primary reason GMH (and related methods) deserve mention here is that, like CDFT, GMH defines the diabatic states with relation to an operator. That is to say, in GMH one chooses diabatic states as eigenstates of the dipole moment, much as in CDFT one chooses diabats as states with defined charge. Thus, while the technical operations involved are quite distinct, the two methods share a common picture of diabatic states as being "special" with respect to some physical operator.

In contrast to GMH, the fragment orbital (FO-DFT) method does not use adiabatic states to produce diabatic states and couplings; it computes diabatic couplings directly. However, it uses a very severe approximation that only the HOMO and LUMO participate in the coupling, and furthermore it requires non-interacting fragments for those HOMO and LUMOs to be meaningful concepts. The requisite fragment calculations are quite similar to the "promolecule" fragments of section 2.2.3. In practice, the LUMO of the acceptor is computed as the HOMO of the reduced acceptor, so that occupied orbitals are used on both sides of the matrix element.[76] Figure 1-3 shows the behavior of the FO-DFT coupling as a function of internuclear separation for the standard electron transfer Zn/Zn^+ dimer cation system, as well as CDFT couplings and the Generalized Mulliken-Hush (GMH) method described previously. The FO-DFT method performs somewhat more poorly than the other methods, and the CDFT and GMH couplings are comparable.[76]

It is important to note that the diabatic picture does not always predict an exponential decay — it also lends itself to the Condon approximation that the coupling is insensitive to (transverse) nuclear motion (e.g. relaxation within the donor or acceptor fragment).[86] The

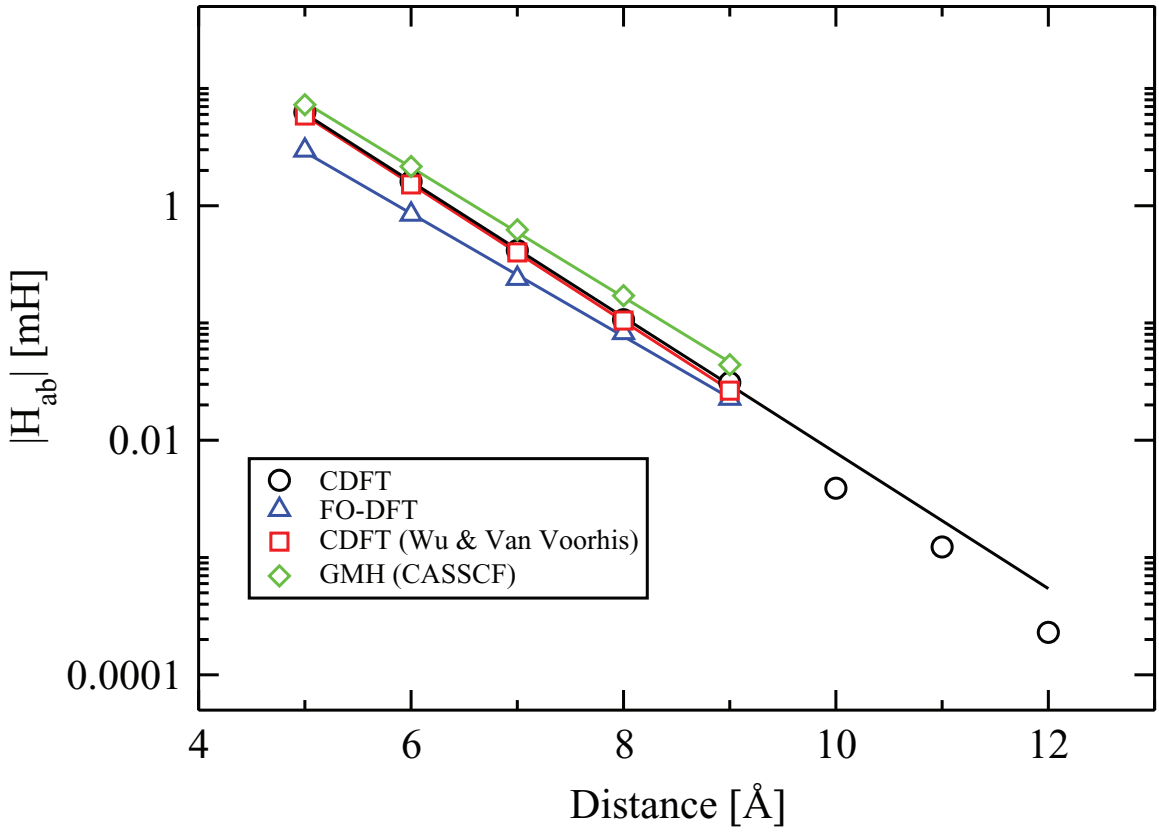

Figure 1-3: GMH, CDFT (plane-wave), CDFT (AO), and FO-DFT methods compared for diabatic coupling elements decaying with separation for the zinc dimer cation. Reprinted with permission from reference 76. Copyright 2010 American Institute of Physics.

availability of CDFT couplings permits investigation of the validity of this approximation for intramolecular electron transfer, by computing the coupling element as a function of the reaction coordinate. Table 1.2 shows the variation in the electronic coupling along the reaction coordinate for intramolecular charge transfer in the mixed-valence tetrathiafulvalene-diquinone (Q-TTF-Q) anion.[7] In the anion, the excess electron localizes on one of the quinone rings, causing some out-of-plane distortion of the structure. Here, as the reaction coordinate moves from q=1 to q=-1 the conformation changes from "electron on left" to "electron on right". As the data make clear, the electronic coupling changes very little over the full domain of the reaction coordinate, showing that the Condon approximation is reasonable for this system.

$q(\pm)$	$\lvert H_{ab}\rvert$
1.0	2.89
0.8	2.95
0.6	3.00
0.4	3.03
0.2	3.05
0.0	3.06
-0.2	3.05
-0.4	3.03
-0.6	3.00
-0.8	2.95
-1.0	2.89

Table 1.2: The electronic coupling element $|H_{ab}|$ (kcal/mol) for Q-TTF-Q anion as a function of the charge-transfer reaction coordinate. $q = -1$ corresponds to charge fully localized on the left quinone, and $q = 1$ to charge localized on the right.7.

Having given ourselves some reassurance that the CDFT coupling of equation (1.53) produces reasonable coupling values, we will move forward and use it as the off-diagonal coupling elements in our CDFT-CI Hamiltonian. Combined with the diagonal CDFT state energies, this gives the matrix to be diagonalized, and the CDFT-CI energies as the eigenvalues.

1.4.2 The CDFT-CI Equations

This formalism can be readily generalized to the case of N states generated by arbitrary constraints:

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & & H_{2N} \\ \vdots & & \ddots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = E \begin{pmatrix} 1 & S_{12} & \cdots & S_{1N} \\ S_{21} & 1 & & S_{2N} \\ \vdots & & \ddots & \vdots \\ S_{N1} & S_{N2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \tag{1.54}$$

where the S terms incorporate the non-orthogonality of generic CDFT states. By analogy to conventional configuration-interaction (CI) methods which build and diagonalize an interaction matrix between Hartree-Fock determinants, this method is termed CDFT-CI, using

interactions between CDFT Kohn-Sham determinants to produce better approximations to the true energy eigenvalues of the Hamiltonian.[9] CDFT-CI is quite remarkable in its generality — it is a framework for constructing custom models for particular systems. There is flexibility to use an arbitrary set of constrained states as the basis for the Hamiltonian of the system in question. Choosing these basis states (tailored for the particular system of interest) then defines the Hamiltonian, which CDFT-CI computes and diagonalizes to yield energies, CI vectors, and other one-electron properties.

The reasons why CDFT-CI is so successful are relatively clear: by using CDFT states as the basis for the CI we are able to effectively control the impact of SIE on the calculations and include dynamic correlation through the CDFT energies. By performing a CI calculation on top of the CDFT states, we add back in the static correlation that is artificially missing from the localized CDFT solutions. As a result, CDFT-CI seems like a well-balanced tool for the description of static correlation in molecular systems.

In a sense, the CDFT-CI basis is a diabatic basis, holding the charge/spin on various portions of the system to be fixed irregardless of the nuclear configuration. Diabatic states such as these have a long history in chemistry, being incorporated into valence bond theories of bonding and models for molecular energy surfaces, with a strong continuing presence in the diverse spread of methods for their determination.[77–85,87–97] The diabatic picture proves itself quite useful in various circumstances, producing PESs for dynamics that vary slowly with nuclear coordinates and thus avoid sharp changes where errors can accumulate.

To confirm with actual calculations our expectations from theory, we return to the system whose discussion began this section, LiF. The dissociation curve for LiF using CDFT-CI is shown in Figure 1-4, within a four-state basis of $Li^{+}F^{-}$, $Li^{-}F^{+}$, $Li^{\uparrow}F^{\downarrow}$, and $Li^{\downarrow}F^{\uparrow}$. Results are presented for CDFT-CI using two different functionals; both perform well, with the hybrid B3LYP yielding the best results. As expected, passing through the dissociation region shows a smooth transition from ionic to neutral for all three curves, as tracked quantitatively by following the CI vectors, shown in Figure 1-5. The crossover between ionic and neutral basis states occurs at 6.6 Å, as expected from where the Coulombic attraction of the ions

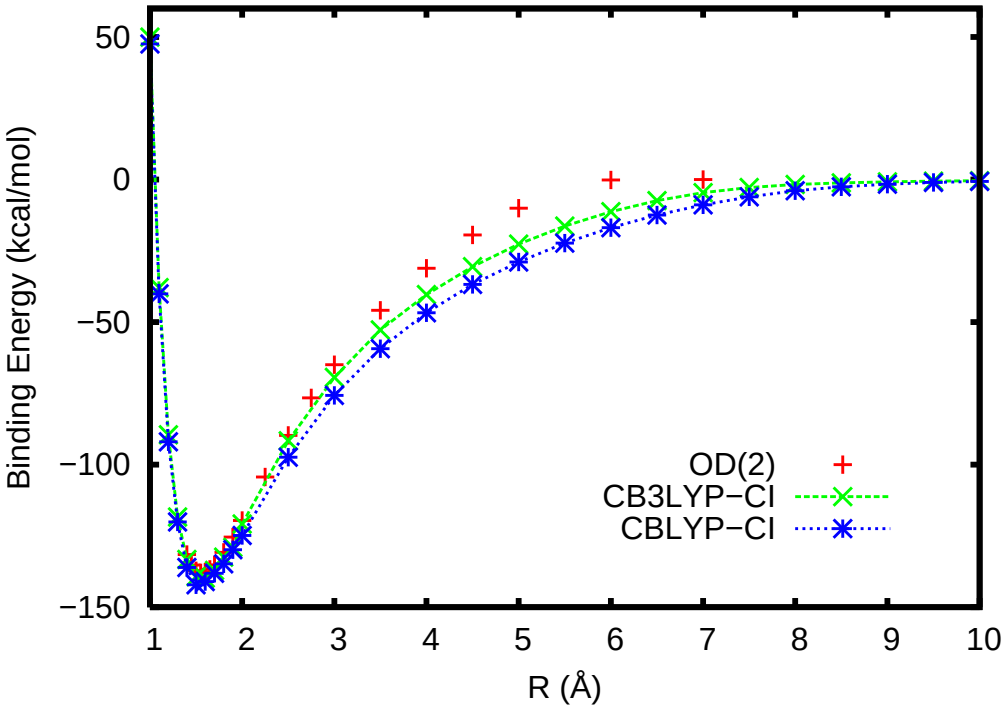

Figure 1-4: Dissociation curve of LiF as computed with various CDFT-CI prescriptions in a 6-311G++(3df,3pd) basis set. Optimized-orbital coupled cluster doubles calculations[98,99] with a second-order correction [OD(2)] results are included as a reference. Reproduced with permission from reference 9. Copyright 2007 American Institute of Physics.

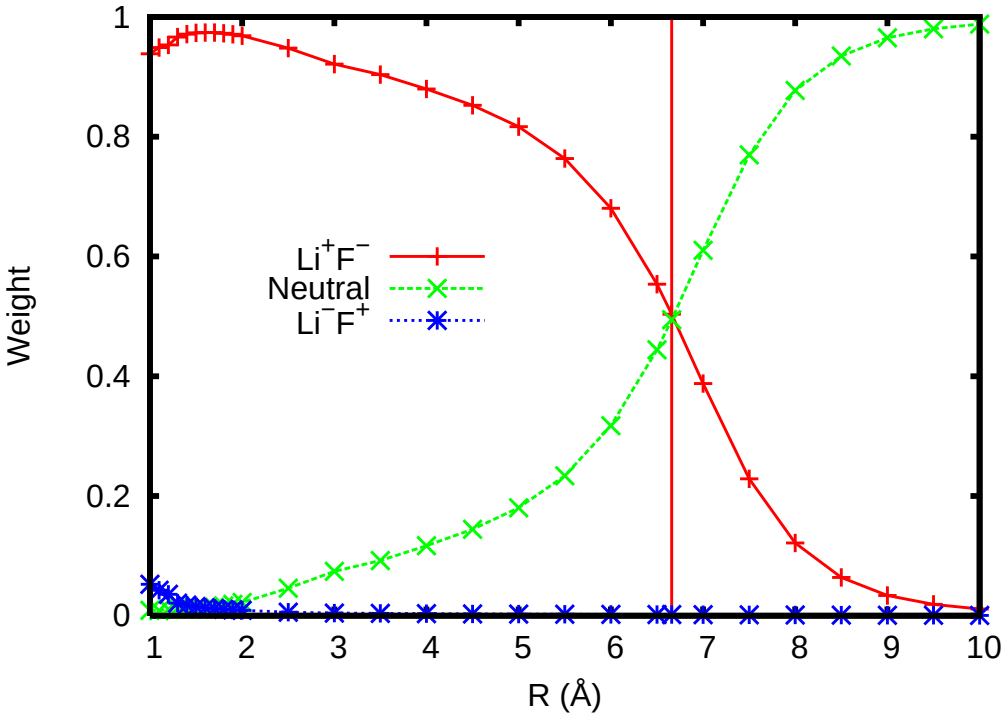

Figure 1-5: Weights of configurations in the final ground state of LiF. Reproduced with permission from reference 9. Copyright 2007 American Institute of Physics.

equals the difference of electron affinity and ionization potentials.[9] Unlike conventional DFT, all the CDFT-CI curves show the correct dissociation limit in Figure 1-4. The accuracy of BLYP and B3LYP around the equilibrium geometry is preserved, indicating that the CDFT-CI prescription does not spoil conventional DFT in regions with little static correlation. The PESs are accurately described everywhere — at the equilibrium geometry, at infinite separation, and in the troublesome region in-between where static correlation is strongest.

Having introduced the methods and basic equations needed and shown that the CDFT-CI couplings, energies, and descriptions of systems are accurate, we are ready to build a thesis atop of CDFT-CI. As such, the outline of the rest of this thesis is as follows. First, we will will show how CDFT-CI can be used at reaction transition states to improve the barrier heights calculated by DFT by a factor of three for a benchmark test suite. We then proceed to use CDFT-CI for electronic excited states, locating conical intersections between the ground and first excited state of some test systems. The next chapter returns to the reaction barrier height test suite, implementing analytical gradients for CDFT-CI and using them to optimize the transition-state geometries of the reactions in the test suite. The analytical gradients are also used on the electronic excited state, locating optimized geometries for the excited states of a few molecules. We conclude with a summary of the applications of CDFT-CI and a perspective on its future.

The bulk of this chapter has been published in reference 65.

Chapter 2

Transition States

In this chapter, CDFT-CI is applied to calculating transition-state energies of chemical reactions that involve bond forming and breaking at the same time. At a given point along the reaction path, the configuration space is spanned by two diabatic-like configurations: reactant and product. Each configuration is constructed self-consistently with spin and charge constraints to maximally retain the identities of the reactants or the products. Finally, the total energy is obtained by diagonalizing an effective Hamiltonian constructed in the basis spanned these two configurations. By design, this prescription does not affect the energies of the reactant or product species, but *will* affect the energy at intermediate points along the reaction coordinate, most notably by modifying the reaction barrier height. When tested with a large set of reactions that include hydrogen transfer, heavy-atom transfer and nucleophilic substitution, CDFT-CI is found to improve the prediction of barrier heights by a factor of 2 to 3 for some commonly used local, semi-local and hybrid functionals. Thus, just as CDFT can be used to cure energy errors in charge localized states, so CDFT-CI can recover the correct energy for charge *delocalized* states by approximating the true wavefunction as a linear combination of localized configurations (e.g. reactant and product). The well-defined procedure and the promising results of CDFT-CI suggest that it could broaden the applicability of traditional DFT methods for reaction barrier heights.

2.1 Introduction

Commonly used density functional theory (DFT) approximations have enjoyed great success in predicting equilibrium thermochemical properties, but have proven less reliable in predicting the energies of transition states. A chemical reaction often involves the breaking of one chemical bond and the formation of another. Its transition state has the characteristics of strong electron correlation due to bond breaking and electron delocalization, where the same electron is shared between different molecular fragments. It is understood that local or semi-local functionals cannot describe the transition states at the same accuracy as the equilibrium states because these functionals have a magnified self-interaction error (SIE) for systems with delocalized electrons.[100,101] The potential curve of H_2^+ exemplifies this defect most clearly, as the DFT energy of the stretched bond configuration is pathologically too low[102] due to the fractional charges on the two hydrogen atoms. Recent studies have also analyzed the SIE in many-electron systems.[103,104] The conclusion is that, even though DFT gives a reasonable electron distribution for the transition state, the energy is severely wrong.

Since Hartree-Fock (HF) is known to overestimate barrier heights, the errors in transition-state energy can be largely reduced by including a fraction of the HF exchange into the exchange-correlation functional. However, to obtain accurate kinetics, over 50% of HF exchange is often needed, which forms hybrid functionals that unfortunately deteriorate thermochemistry results.[105] To make a fixed fraction of HF exchange to work for all cases, one can empirically fashion complementary local exchange-correlation functionals that cover various situations.[106,107] While giving excellent kinetics by design, the large number of fitted parameters in these functionals can make it difficult to grasp the underlying physics. In a very different fashion, HF exchange can be built into the functional nonlinearly.[108–110] These novel functionals have only a few parameters to fit and can offer improved understanding of fundamental issues in DFT.[111–113] They also have excellent performance for both thermochemistry and kinetics.[111,114,115] The disadvantage of these functionals is the difficulty of implementing them self-consistently. Taken together, recent progress in functional developments gives great hope that a unique functional can be found to apply to all chemical

problems,[116] though the best choice for this purpose is still far from clear. For that reason, traditional functionals remain in widespread use.

In addition to ongoing functional development, there is a steady effort to combine the success of DFT and traditional wave function based quantum chemistry methods, such as valence bond (VB)[117] , perturbation theory,[118] and configuration interaction (CI).[119–121] The challenge here is three-fold. First, one wants to retain the simplicity of DFT methods so as not to significantly increase the computational cost. Second, one must cope with the issue of double-counting, so that the component present in the wavefunction calculation is absent in the DFT simulation. Finally, one must be careful not to use DFT in situations where it is known to be inaccurate. For example, as detailed above, one should avoid using DFT in situations with strong electron delocalization.

CDFT-CI has shown promise to meet these challenges by successfully correcting some fundamental DFT issues in the dissociation energy curves of diatomic molecules. The basic idea is to avoid unreliable DFT calculations on delocalized electronic states by forcing the bonding electrons to localize using constraint potentials.[5] After obtaining reliable localized states, the delocalized nature of the original states is restored with a CI calculation based on only a few CDFT configurations. As long as the number of configurations is small, the computational cost of this method is not significantly different from normal DFT methods and thus it can be applied to very large systems.

In this section, we extend the application of CDFT-CI to the prediction of reaction barrier heights in simple chemical reactions. In particular, we set up a simple scheme for the specification of the localized states based on a user-specified separation of the atoms into reactant and product. We test this automated scheme on a set of reactions that include hydrogen transfer, heavy atom transfer and nucleophilic substitution using a variety of local, semi-local and hybrid functionals. We find the new method improves barrier heights by factors of as much as two to three. We close the section with some discussion of the prospects of this prescription for other reaction properties.

2.2 Method

2.2.1 CDFT-CI

Our implementation of constrained DFT and CDFT-CI has been described in detail previously[5–7,9] and in chapter 1. We need only repeat the few key equations used herein.

Like all CI techniques, CDFT-CI produces eigenstates by diagonalizing a Hamiltonian in a basis of configurations, i.e.,

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = E \begin{pmatrix} 1 & S_{12} \\ S_{21} & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \tag{2.1}$$

Here H is the Hamiltonian matrix, S_{12} is the overlap, c_1, c_2's are coefficients and we have assumed for simplicity that there are only two important configurations. In our method, configurations $|1\rangle$ and $|2\rangle$ are the *exact* ground states of the system under two different potentials, V_1 and V_2, and thus are not orthogonal to each other. All the matrix elements $(H_{11}, H_{12}, S_{12}, \cdots)$ are obtained from CDFT calculations.

CDFT determines lowest energy electronic state that satisfies certain constraints on the electron density.[4] Each constraint is generally expressed as

$$\int w_C(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} - N_C = 0, \tag{2.2}$$

where w_C is a weighting function, ρ is the electron density, and N_C is the constraint value. Such constraints can be applied to both α and β densities. If $w_C^\alpha = w_C^\beta$, equation 2.2 puts a charge constraint on the system; if $w_C^\alpha = -w_C^\beta$, the constraint is on the spin density. Thus we can choose w_C and N_C to constrain the molecule to be in different local charge or spin configurations. Minimizing the Kohn-Sham[122] energy functional under the constraint of equation 2.2 is equivalent to solving an effective KS equation

$$[-\frac{1}{2}\nabla^2 + v_{\mathrm{n}}(\mathbf{r}) + \int \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d\mathbf{r}' + v_{\mathrm{xc}}(\mathbf{r}) + V_C w_C(\mathbf{r})]\phi_i = \varepsilon_i\phi_i, \tag{2.3}$$

where an extra external potential $V_C w_C$ is applied. Here V_C is a unique variable to be determined simultaneously with KS orbitals to satisfy equation 2.2. We[5,6] and other groups[123–125] have developed similar methods to solve these coupled equations efficiently.

In CDFT-CI, the configurations $|1\rangle$ and $|2\rangle$ are constrained states. If the exact exchange-correlation functional were known, one could compute the energies of these states (E_1 and E_2) exactly. Their interaction, i.e. the off-diagonal elements of the Hamiltonian matrix, cannot be directly calculated from DFT because the true wavefunctions of these states are not available. We elect to use the following approximation for the off-diagonal elements. An exact expression for H_{12} can be written as:

$$\begin{aligned} H_{12} &= \frac{1}{2}[\langle 1|H + V_{C1}w_{C1} - V_{C1}w_{C1}|2\rangle + \langle 1|H + V_{C2}w_{C2} - V_{C2}w_{C2}|2\rangle] \\ &= \frac{1}{2}[(E_1 + V_{C1}N_{C1} + E_2 + V_{C2}N_{C2})\langle 1|2\rangle - V_{C1}\langle 1|w_{C1}|2\rangle - V_{C2}\langle 1|w_{C2}|2\rangle]. \quad (2.4) \end{aligned}$$

Equation 2.4 is exact because the constrained state $|i\rangle$ is obtained as the ground state of $H + V_{Ci}w_{Ci}$ with an eigenvalue of $E_i + V_{Ci}N_{Ci}$, where i =1 or 2. But the wavefunction overlap $S_{12} = \langle 1|2\rangle$ and the coupling of weighting functions $\langle 1|w_{Ci}|2\rangle$ in equation 2.4 still cannot be directly calculated from DFT. However, the coupling of the two-electron operator in the original Hamiltonian is now absent. Thus equation 2.4 should be less sensitive to the presence (or absence) of correlation in the wavefunctions, which might make it easier to apply approximations. We therefore use the Slater determinants of Kohn-Sham orbitals in equation 2.4 to obtain S_{12}, $\langle 1|w_{Ci}|2\rangle$ and H_{12}. Our previous experience shows that this approximation is successful.[7,9]

2.2.2 Configurations and constraints

The configurations in CDFT-CI are not based on orbital occupations, but on local charges and spins of different molecular fragments. Previous work[9] has made the connection between the configurations in CDFT-CI and valence bond (VB) configurations. This connection means that CDFT-CI can borrow many well-developed ideas from VB, which will be very

helpful in the choice of configurations. At the simplest level, a VB model for a chemical reaction includes two states R and P, corresponding to the reactant and product states, respectively. This idea forms the basis, for example, of the empirical VB (EVB) approach commonly used in molecular dynamics.[87,126] This picture can be directly adopted in CDFT-CI.

The reactions we focus on in this work are generally expressed as

$$\mathrm{AX + B} \quad \leftrightarrow \quad \{\mathrm{A \cdots X \cdots B}\} \quad \leftrightarrow \quad \mathrm{A + XB} \tag{2.5}$$

Hence the whole system is divided into three fragments A, B and X, where the bonds that are breaking and forming in the reaction are between A and X, and B and X, respectively. The electrons in X participate in both bonds, thus are delocalized at the transition state. Note that A, B or X represents the nuclear conformation of an entire molecular fragment. We identify reactant and the product electronic states according to the *bonding patterns* of the reactant and product configurations. These states are diabatic-like configurations: regardless of the nuclear configuration, the reactant (product) state retains the electronic structure appropriate for AX+B (A+XB). We illustrate this point by using the following nucleophilic substitution reaction:

$$\mathrm{ClCH_3 + F^-} \quad \leftrightarrow \quad \{\mathrm{[Cl \cdots CH_3 \cdots F]^-}\} \quad \leftrightarrow \quad \mathrm{Cl^- + CH_3F}. \tag{2.6}$$

According to the reactants and the products, the two configurations used to span the transition state are $\mathrm{ClCH_3}$(N=0,S=0)+F(N=-1,S=0) and Cl(N=-1,S=0)+$\mathrm{CH_3F}$(N=0,S=0), where the charge N and the total spin S are defined for each fragment and will be enforced with constrained DFT.

The charge and spin values for fragments in each configuration are well-defined formally. But they are only exact when the fragments are completely separate. At the transition state, the density overlap between fragments makes the formal charges and spins less realistic as constraint values. This problem has been discussed in our previous work, and addressed

with a promolecule density approach.[9] We will use this approach again to obtain practical constraint values for reactant and product.

2.2.3 Promolecules

CDFT is designed to construct electronic states of fixed character at arbitrary geometries, even those where fragments overlap, but sometimes it does not perform as well as might be hoped in such close-contact geometries. One of the sorest impediments to its ability to do so is the choice of atomic population prescription. In cases of close approach, the real-space constraint potentials must distinguish between fragments in regions where the density is nonzero, so that assignment of density in that transition region to fragment "A" or fragment "B" is ambiguous, a clear weakness of the available charge prescriptions. The simplest example that shows this ambiguity is H_2^+, with a single electron and two protons. Formally, we can constrain the electron to lie only on one proton ("A"), but when the two protons begin to approach each other, any real-space-based atomic population scheme will begin to assign some fraction of this electron to the second nucleus ("B"), for any physically reasonable density corresponding to a constrained $H - H^+$ configuration. Thus, the formal charge constraint putting a full electron on atom A is unattainable with present charge prescriptions, and the numerical value of the constraint must be adjusted to accommodate their failings.

This failure of formal charge/spin constraint values extends to the case of arbitrary fragments, coming into play when constrained molecular fragments come in close approach, as in nucleophilic substitution reactions; the promolecule formalism was pioneered to allow CDFT to be used in precisely these types of situations.[74] The sequence of steps involved in the promolecule formalism is a bit complex and probably best understood with our example of equation (2.6). The overlap between states is strongest at the reaction transition state, and the formal charges and spins simply do not represent realistic constraint values for any population scheme at that closest approach. The promolecule treatment corrects for these errors by modifying the formal integer charge (spin) constraints into values that are

appropriate for a given charge prescription. The basic steps involved are illustrated in Figure 2-1.

The first step in the calculation is to break the system into the appropriate fragments — e.g. F^- and CH_3Cl for the reactant configuration — maintaining the internal geometry of each fragment (Figure 2-1a). One then performs separate calculations on each fragment with the relevant total charge and total spin (Figure 2-1b). These converged fragment densities are the fragment promolecule densities — they approximate what an F^- or CH_3Cl density should look like. The fragment promolecule densities are then arranged in the original molecular geometry and summed to obtain the total promolecule density, $\tilde{\rho}$ (Figure 2-1c). For the example S_N2 system in the reactant configuration, we obtain

$$\tilde{\rho}_r^\sigma(\mathbf{r}) = \rho_{\mathrm{ClCH_3}}^\sigma(\mathbf{r}) + \rho_{\mathrm{F^-}}^\sigma(\mathbf{r}) \qquad (\sigma = \alpha \text{ or } \beta) \tag{2.7}$$

With this full promolecule density $\tilde{\rho}(\mathbf{r})$, the actual constraint values used for the final reactant CDFT calculation are

$$N_{\mathrm{tot}} = \int w(\mathbf{r}) \left[\tilde{\rho}_r^\alpha(\mathbf{r}) + \tilde{\rho}_r^\beta(\mathbf{r})\right] d\mathbf{r} \tag{2.8}$$

$$M_{\mathrm{spin}} = \int w(\mathbf{r}) \left[\tilde{\rho}_r^\alpha(\mathbf{r}) - \tilde{\rho}_r^\beta(\mathbf{r})\right] d\mathbf{r} \tag{2.9}$$

as depicted in Figure 2-2. These modified constraint values reflect the expectation that a molecule constrained to be in the "reactant state" should, *within the limits of the charge prescription in use*, look as much as possible like the superposition of the reactants brought from infinite separation to the geometry in question. In many cases, the correction from the promolecule density is small and can safely be omitted, particularly when the constrained regions are on different molecules (as for charge transfer in organic semiconductors[71,127–129]) or widely separated (as for molecular sensors[130]). In other cases, though, the correction is essential, as for the very small fragments illustrated in Figure 1-1, or for reaction transition states that enter into the CDFT-CI barrier height calculations under consideration in this chapter. As the fragments come from being well-separated into closer approach, the effect of

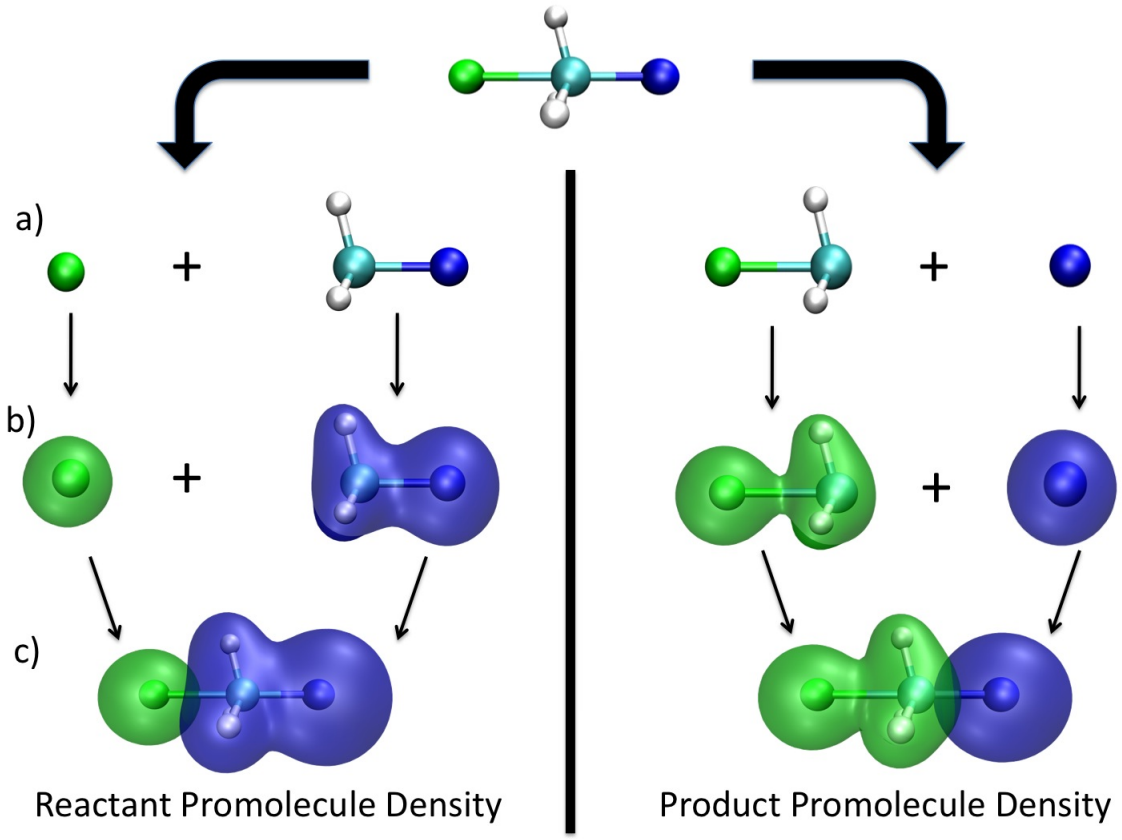

Figure 2-1: Construction of reactant and product promolecule densities for $[F \cdots CH_3 \cdots Cl]^-$. (a) The system is divided into fragments, with atoms being apportioned to the fragments corresponding to reactant (product) and the internal fragment geometry held fixed at the transition-state values. (b) The ground-state density of the isolated reactant (product) fragments is determined. (c) The fragment densities are superimposed to form the reactant (product) promolecule density. Reproduced with permission from reference 74. Copyright 2009 American Institute of Physics.

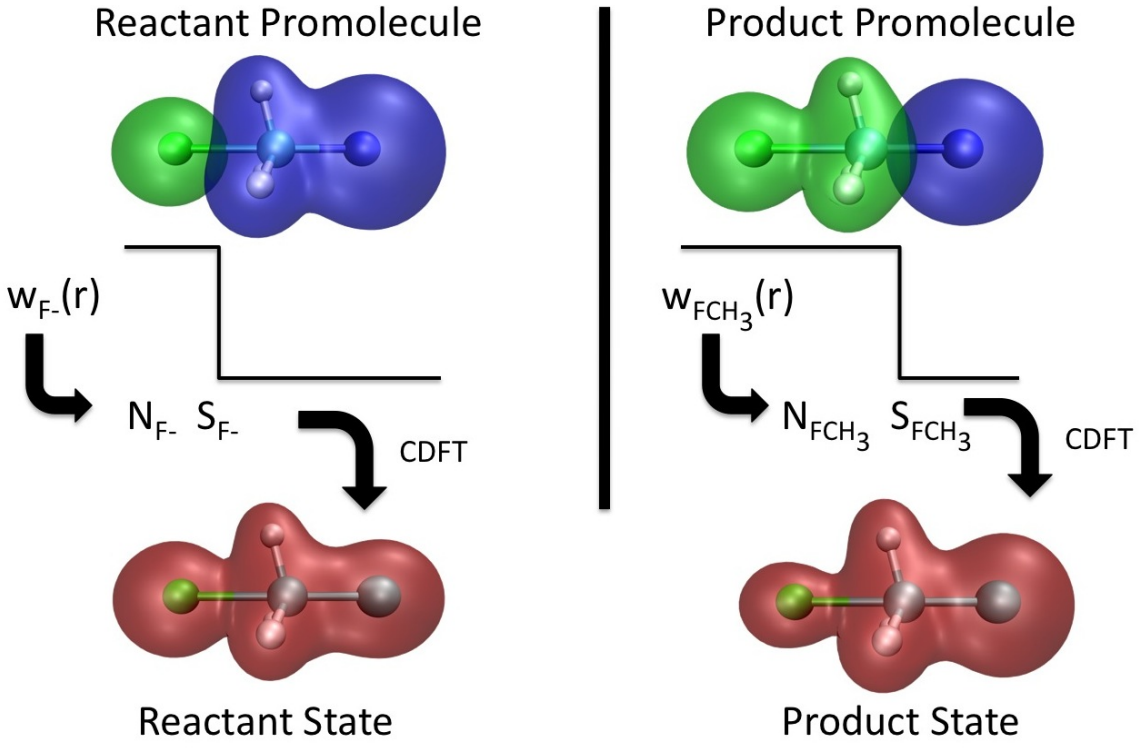

Figure 2-2: Computation of reactant- and product-like states for $[\mathrm{F}\cdots\mathrm{CH_3}\cdots\mathrm{Cl}]^-$. The total promolecule density is integrated against the charge prescription function $w(\mathbf{r})$ of the reactant (product) fragment to obtain target constraint values. CDFT calculations with these updated constraints produce the final reactant and product states. Reproduced with permission from reference 74. Copyright American Institute of Physics.

the correction grows smoothly, owing to the continuity of all functions involved. By the time the reacting fragments reach the transition-state geometry, the correction can be larger than half an electron! Nonetheless, the constraining potentials continue to enforce a consistent physical picture throughout the entire reaction, and allow higher-level methods to be built atop that picture.

To describe the product, an analogous procedure is used: the promolecule density $\tilde{\rho}_p$ is formed by superimposing the densities of Cl^- and CH_3F,

$$\tilde{\rho}_p^\sigma(\mathbf{r}) = \rho_{Cl^-}^\sigma(\mathbf{r}) + \rho_{CH_3F}^\sigma(\mathbf{r}) \qquad (\sigma = \alpha \text{ or } \beta)$$

where the calculation on CH_3F is performed at the internal geometry of CH_3F in the full calculation, and the final N_{tot} and M_{spin} values determined accordingly.

Now consider applying CDFT-CI to the equilibrium of the reactant state. At this point, for a bimolecular reaction, the reactants are well separated from each other. Using the reactant promolecule density will effectively put no constraints on the system. Thus with the two-state model, the ground-state energy at the reactant equilibrium will simply be that of the unconstrained reactants. The analogous argument also applies the product. Therefore, a CDFT-CI calculation of this type will not affect the energies of the reaction endpoints and will have no effect on thermochemistry. However, the energy at intermediate points along the reaction coordinate (such as the transition state) *will* be affected by this prescription, and thus CDFT-CI shows promise for kinetics.

2.3 Tests

2.3.1 Computational details

We have selected for our tests a set of reactions from the HTBH38/04 and NHTBH38 databases prepared by Truhlar and coworkers.[131] Table 2.1 lists all the reactions. Specifically, we have used 18 hydrogen transfer (HT), 6 heavy atom transfer (HAT) and 8 nucleophilic

substitution (SN) reactions. We have limited our test to intermolecular group transfer reactions in this work because they have well-defined subsystems for reactants and products. (Note that the SN reactions are viewed as group transfer, here.) Hence Reaction 19 in the HTBH38/04 set,[131] i.e., the intramolecular HT of *cis*-1,3-pentadiene, and the unimolecular and association reactions in the NHTBH38 set[131] are not used. Geometries for all reactants, products and transition states are taken from the databases, and so are the reference barrier heights, which represent the best estimates from experimental and theoretical kinetics data.

All our calculations are performed with a development version of Q-Chem. Four commonly used exchange and correlation functionals are tested; they are Slater exchange[132] with Vokso-Wilk-Nusair (VWN)[133] correlation, Becke exchange[134] with Lee-Yang-Parr (LYP) correlation,[135] Perdew-Burke-Ernserhof (PBE) exchange and correlation,[136] and the hybrid functionals B3LYP[137] and B97-2.[138,139] These functionals represent the spectrum of exchange-correlation functionals from the local spin density approximation (LSDA), to generalized gradient approximation (GGA), to the classical hybrid functional, and new parameterized functionals. The 6-311+G(3*df*,2*p*) basis set and the SG-1 standard quadrature grids[140] are used in all calculations. The original Becke weighting function[50] without atomic size adjustment is used for integration as well as for the definition of w_C in constraint DFT.

For each reaction, the energies of the reactant and product states are calculated in normal SCF procedures. In this work, the transition-state energy is calculated in both the traditional way and with the CDFT-CI method.

2.3.2 Results and discussion

Our main computational results can be found in Table 2.1 and 2.2. In Table 2.1, we have listed the transition-state energy shifts of CDFT-CI relative to traditional DFT with the same functional for all the tested reactions. In Table 2.2, the mean errors (ME) and mean absolute errors (MAE) in barrier heights are enumerated for CDFT-CI and traditional DFT.

It is clear from Table 2.1 that CDFT-CI consistently shifts the transition-state energy upwards for all functionals. There are two causes for this shift. The first relates to SIE.

Table 2.1: Energy change (in kcal/mol) of the transition state for each reaction due to CDFT-CI.

		LSDA	PBE	BLYP	B3LYP	B97-2
1.	$H+HCl \leftrightarrow C1+H_2$	11.7	11.9	12.4	12.2	11.8
2.	$OH+H_2 \leftrightarrow H+H_2O$	14.0	13.4	13.6	9.8	10.5
3.	$CH_3+H_2 \leftrightarrow H+CH_4$	4.4	4.1	4.3	4.0	4.5
4.	$OH+CH_4 \leftrightarrow CH_3+H_2O$	11.0	10.9	11.1	11.3	12.0
5.	$H+H_2 \leftrightarrow H_2+H$	3.6	3.3	3.8	3.4	3.8
6.	$OH+NH_3 \leftrightarrow H_2O+NH_2$	7.1	5.8	5.9	5.8	5.9
7.	$HCl+CH_3 \leftrightarrow Cl+CH_4$	8.5	7.3	7.2	7.7	8.4
8.	$OH+C_2H_6 \leftrightarrow H_2O+C_2H_5$	10.6	10.7	10.9	11.4	11.9
9.	$F+H_2 \leftrightarrow H+HF$	12.7	11.0	11.2	9.4	9.6
10.	$O+CH_4 \leftrightarrow OH+CH_3$	9.3	9.3	9.4	10.0	8.7
11.	$H+PH_3 \leftrightarrow PH_2+H_2$	4.3	3.8	4.3	4.2	4.3
12.	$H+HO \leftrightarrow H_2+O$	27.5	8.4	8.8	5.4	6.4
13.	$H+H_2S \leftrightarrow H_2+HS$	13.4	6.2	6.8	6.7	6.7
14.	$O+HCl \leftrightarrow OH+Cl$	3.2	3.6	3.5	4.6	4.8
15.	$NH_2+CH_3 \leftrightarrow CH_4+NH$	4.6	4.3	4.4	4.2	4.6
16.	$NH_2+C_2H_5 \leftrightarrow C_2H_6+NH$	4.5	4.3	4.4	4.3	4.6
17.	$C_2H_6+NH_2 \leftrightarrow NH_3+C_2H_5$	4.7	4.5	4.5	4.5	4.8
18.	$NH_2+CH_4 \leftrightarrow CH_3+NH_3$	4.8	4.6	4.6	4.5	4.9
19.	$H+N_2O \leftrightarrow OH+N_2$	8.8	8.1	8.4	8.6	8.8
20.	$H+FH \leftrightarrow HF+H$	6.3	6.1	6.4	6.0	5.9
21.	$H+ClH \leftrightarrow HC1+H$	7.0	6.9	7.3	7.1	7.5
22.	$H+FCH_3 \leftrightarrow HF+CH_3$	6.0	5.5	5.8	5.4	5.5
23.	$H+F_2 \leftrightarrow HF+F$	10.5	9.3	9.8	8.8	8.3
24.	$CH_3+FCl \leftrightarrow CH_3F+Cl$	11.6	8.4	8.5	7.7	8.8
25.	$F^-+CH_3F \leftrightarrow FCH_3+F^-$	6.5	6.1	5.6	5.2	5.6
26.	$F^-\cdots CH_3F \leftrightarrow FCH_3\cdots F^-$	6.5	6.1	5.6	5.2	5.6
27.	$Cl^-+CH_3Cl \leftrightarrow ClCH_3+Cl^-$	3.6	3.4	3.0	2.8	3.0
28.	$Cl^-\cdots CH_3Cl \leftrightarrow ClCH_3\cdots Cl^-$	3.6	3.4	3.0	2.8	3.0
29.	$F^-+CH_3Cl \leftrightarrow FCH_3+Cl^-$	4.6	4.3	3.9	3.5	3.7
30.	$F^-\cdots CH_3Cl \leftrightarrow FCH_3\cdots Cl^-$	4.6	4.3	3.9	3.5	3.7
31.	$OH^-+CH_3F \leftrightarrow HOCH_3+F^-$	5.4	5.1	4.6	4.3	4.6
32.	$OH^-\cdots CH_3F \leftrightarrow HOCH_3\cdots F^-$	5.4	5.1	4.6	4.3	4.6

In CDFT-CI only energies of localized electron configurations are directly calculated. Thus, the large overstabilization of delocalized transition states is avoided, resulting in an overall

Table 2.2: Summary of the mean error (ME) and mean absolute error (MAE) of barrier heights. The numbers in parenthesis represent the total number of barrier heights in each data set. Numbers in bold are CDFT-CI results. All energies are in kcal/mol.

	LSDA		PBE		BLYP		B3LYP		B97-2	
hydrogen transfer (36)										
ME	-18.2	**-9.3**	-9.7	**-2.6**	-8.0	**-0.7**	-4.6	**2.3**	-3.4	**3.7**
MAE	18.2	**10.9**	9.7	**3.8**	8.0	**3.1**	4.6	**3.0**	3.6	**4.0**
heavy atom transfer (12)										
ME	-23.5	**-15.1**	-14.9	**-7.6**	-14.7	**-7.0**	-8.5	**-1.2**	-3.1	**4.3**
MAE	23.5	**15.1**	14.9	**7.6**	14.7	**7.0**	8.5	**2.3**	3.4	**4.7**
nucleophilic substitution (16)										
ME	-8.4	**-3.4**	-6.9	**-2.2**	-7.3	**-3.0**	-3.4	**0.6**	-1.4	**2.9**
MAE	8.4	**4.2**	6.9	**2.3**	7.3	**3.0**	3.4	**1.3**	1.4	**2.9**
all (64)										
ME	-16.7	**-8.9**	-10.0	**-3.4**	-9.1	**-2.5**	-5.0	**1.2**	-2.8	**3.6**
MAE	16.7	**10.0**	10.0	**4.2**	9.1	**3.8**	5.1	**2.5**	3.0	**3.9**

upward shift in the barrier. Clearly this systematic trend will tend to improve the quality of CDFT-CI reaction barriers, as it removes one of the well-known deficiencies of commonly used functionals. However, there is a second factor at work here. CDFT-CI makes the *assumption* that the electronic state at any point along the reaction coordinate can be represented exactly as a linear combination of reactant-like and product-like configurations. Ignoring any other approximations made in constructing E_{xc} or computing the coupling between reactant and product, this assumption will introduce some bias into the results. Further, it is easy to see

Table 2.3: Factors of improvement that are calculated as the MAE ratio between the traditional DFT and constrained DFT results. The columns list different types of reactions , where HT is hydrogen transfer; HAT is heavy atom transfer; SN is nucleophilic substitution.

	LDA	PBE	BLYP	B3LYP	B97-2
HT	1.7	2.5	2.6	1.5	0.9
HAT	1.6	2.0	2.1	3.7	0.7
SN	2.0	3.0	2.4	2.7	0.5
all	1.7	2.4	2.4	2.1	0.8

that this bias will tend to push the transition state higher in energy. One can always get the correct answer out of CI by including enough configurations, and any approximation that reduces the number of configurations will always result in a higher energy. Thus, this trend will introduce an uncontrolled error in our calculations. One must hope that either this error is small or that it cancels with some other error in the calculation (e.g. due to the choice of functional).

Examining the results from Table 2.1 shows that the CDFT-CI energy shifts of a given transition state for different functionals are generally in a close range, although there is a clear decrease in the shift in the order LSDA > GGA > Hybrid. This is consistent with the observations above: the functional-independent shift can be largely associated with the approximations inherent to CDFT-CI, while the functional dependence must be associated with the reduction of SIE as the quality of the functional is improved. The large functional-independent piece suggests that the pragmatic success of CDFT-CI (in the simplest two-state model) will rely in part on cancellation of errors and that a particular approximate functional will have the best performance when coupled with CDFT-CI. As we shall see shortly, this turns out to be the case as CDFT-CI/B3LYP gives the best barrier heights, despite the fact that B3LYP does not give the best kinetic data among the unconstrained functionals.

Table 2.2 displays the mean errors (ME) and mean absolute errors (MAE) of CDFT-CI and traditional DFT for various reaction barrier heights. As a more direct comparison of the two approaches, we also list in Table 2.3 the factors of improvement, which is calculated as the ratio of the MAE between traditional DFT and CDFT-CI. Clearly, the CDFT-CI calculations offer significant improvements over their unconstrained counterparts. We note, for example, that C-PBE-CI gives reaction barrier heights competitive with unconstrained hybrid functionals. This is significant because in many program packages — most notably, those formulated in a plane-wave basis — Hartree-Fock exchange is several times more expensive than semilocal DFT exchange. PBE contains no Hartree-Fock exchange, and thus the ability to obtain relatively accurate barrier heights from PBE without modifying the functional is intriguing. The error reduction is especially dramatic for functionals from older

generations, such as LSDA, PBE and BLYP. Even for the arguably most popular functional — the simple hybrid B3LYP — there is significant improvement that brings its mean absolute errors for barrier heights to approximately 2 kcal/mol. The situation is different for the newer parameterized functional B97-2, in which case its error in barrier heights is already small and the CDFT-CI method gives slightly worse results. In this latter case, one assumes that the inherent errors of the CDFT-CI procedure are large enough that they outstrip the errors present in the functional, resulting in slightly worse predictions. Nonetheless, the CDFT-CI procedure seems promising, in particular for the simpler functionals.

The result presented above were all performed using the 6-311+G(3df,2p) basis set. We have verified that, for the properties considered here, this is functionally a complete basis. For example, in the much larger aug-cc-pVQZ basis the MAE of B3LYP for our test set is 4.9 kcal/mol, as compared to 5.1 kcal/mol in the smaller basis. Meanwhile, CB3LYP-CI in the larger basis gives an MAE 2.5 kcal/mol, precisely the same as in the smaller basis. Thus the results above are not appreciably affected by the finiteness of our basis.

2.4 Conclusions

CDFT-CI[9] has been successfully applied to calculating energies of transition states that have bonds breaking and forming simultaneously. Traditional DFT approximations are known to underestimate the barrier heights for these reactions due to their large errors for electron delocalization. With CDFT-CI, the transition state is represented by a simple two-state model, involving the reactant and product configurations as in EVB theory.[87,126] Each state represents a localized electronic configuration, whose DFT energy is easily predicted using semilocal functionals. Thus, the transition state and equilibrium states are treated more consistently in CDFT-CI than in traditional DFT. When tested over a suite of 32 reactions, CDFT-CI is found to consistently raise the transition-state energy, and the energy shift for the same reaction is very similar for different functionals. Among the functionals tested, B3LYP is found to give the best overall performance in the CDFT-CI method, giving a mean absolute error of 2.5 kcal/mol for barrier heights. This good accuracy is very promising, but

probably relies in part on a cancellation of errors between the inherent approximations in CDFT-CI and the limitations of the B3LYP functional. Our results suggest that comparable accuracy for kinetics as for thermochemistry is now possible using the traditional functionals within the CDFT-CI framework. In this sense, the self-consistent CDFT-CI has the desired features discussed in a recent paper[141] where Hartree-Fock orbitals are used in local and semi-local functionals to obtain better barrier heights.

Within the two-state model, the procedures for a CDFT-CI calculation are well defined. The entire calculation can be started from a single input file like a usual DFT routine. The only additional required information is the nature of the fragments, including their charges and multiplicities, defined by the reactants and products. Subsequent work has made CDFT-CI calculations straightforward (chapter 3) and implemented analytic gradients (chapter 4), enabling access to a fairly wide range of chemistry. Notably, the results presented here take as input transition-state structures obtained at higher levels of theory, and it is not *a priori* clear that these structures are close to the saddle point structures that would be predicted by CDFT-CI. In many chemical applications, these transition-state structures can be more valuable than the barrier height, and determining to what extent CDFT-CI improves the prediction of transition-state geometries is an important part of the work presented in chapter 4.

Looking forward in additional directions, there are two more obvious paths for extending this work. On the one hand, one would like to assess this CDFT-CI prescription to more complex situations, involving, for example, more than two fragments or more than two active electronic states. Steps towards this have been taken, e.g., in chapter 3, but the field remains only minimally explored. Such improvements could potentially mitigate the overestimation of energies that arises from restricting attention to two electronic configurations. We have shown previously that CDFT-CI can be applied to more than two state systems,[9,75] but the challenge lies in selecting larger configuration spaces in an *automated* way. The selection of "reactant" and "product", while somewhat arbitrary, has a certain amount of chemical appeal to it, and some further intuition and experimentation were used to obtain correct

results in reference 75. As one pushes toward larger active spaces in CDFT-CI, one would like to maintain this intuitive feel.

At the same time, one would like to extend CDFT-CI to a broader range of chemical applications. CDFT-CI presents a unique paradigm to study chemical reactions. It is a molecular orbital (MO) method. However, its connection to valence bond theory can offer insights that are not obvious from MO results but nonetheless are familiar to chemists.[142] In fact it is exactly CDFT-CI's capability of incorporating chemical intuition readily into quantum chemistry computations that makes it so effective in describing interesting chemistry such as bond breaking and forming. We will see it used to study excited states in chapters 3 and 4. Yet there are many venues left to explore: what about using CDFT-CI to describe localized excited states at a defect, or near an interface? Further applications of CDFT-CI will elucidate its advantages and disadvantages as a framework for describing all sorts of systems.

The bulk of this chapter has been published in reference 74.

Chapter 3

Conical Intersections

Excited states, whether at an interface or just general low-lying accessible excited states, are of great interest to chemistry as we seek to make matter do our bidding. Time-Dependent Density Functional Theory (TD-DFT) is now well-established as an efficient method to calculate excited-state energies of many chemical systems, being frequently used to investigate photochemistry and compute vertical excitation energies.[143–146] The manifold of ground and excited electronic states for almost all molecular systems is rife with conical intersections (CIs) — seams of true degeneracy between electronic states. These intersections are frequently important for photochemical dynamics, as an intersection provides an easy pathway for nonradiative decay, and CIs are usually quite accessible to photoexcited systems.[147] Even on the ground electronic state, and even in cases where the actual intersection is not energetically accessible, CIs can still have a dramatic effect on quantum dynamics, as evidenced in the phenomenon of geometric phase.[147] Over the entire many-dimensional manifold of states, CIs are actually quite prevalent, and accurate excited-state or dynamics treatments must account for them.[147] Unfortunately, despite its many successes, TD-DFT completely fails to describe conical intersections between excited states and the ground state.[148] In most cases, TD-DFT only produces one direction of degeneracy-splitting, and the S_1 state frequently has an ill-behaved $\frac{\partial E}{\partial R}$ in the vicinity of the intersection.[148] In order to get a qualitatively correct treatment of CIs, multi-reference methods such as Complete Active

Space (CAS), MRCI, and CASPT2 must be used.[148] Unfortunately, existing multi-reference treatments are almost universally wavefunction-based, and the computationally reasonable methods (such as CAS) have been shown to give only qualitatively-correct descriptions of multi-state energy-surface manifolds;[149] more accurate methods such as MRCI are frequently too expensive for use on real systems. We have developed a multi-reference DFT method, Constrained Density Functional Theory–Configuration Interaction (CDFT-CI), in previous works;[6,9] it has been shown to be effective at calculating ground-state energies and barrier heights,[74] but its treatment of electronic excited states has yet to be presented. In this chapter, we present CDFT-CI as a method for obtaining qualitatively-correct energy-surface manifolds encompassing both ground and excited electronic states, producing well-behaved conical intersections at appropriate geometries.

3.1 Methods

CDFT-CI is designed to be a DFT-based method that can robustly treat systems with both dynamic and static correlation. Existing DFT functionals can perform well for systems with dynamic correlation, but tend to do poorly when static correlation is present (such as when multiple near-degenerate states are accessible).[150–153] Other techniques combining DFT with Configuration Interaction have been proposed;[121] these and related techniques are discussed in relation to CDFT-CI in our previous work.[9,74] CDFT-CI works by introducing an active space whose states originate from distinct (constrained) SCF calculations, constructing a CI matrix of fully-*ab-initio* energies and couplings, and diagonalizing that matrix to obtain energies and coefficient vectors for the adiabatic states. This explicitly includes static correlation due to the multi-reference nature of the eigenstates, and dynamic correlation is treated through the DFT functional used in the formation of the basis states. In particular, these states are produced using the Constrained DFT (CDFT) method, which introduces an additional constraint potential to each state's Hamiltonian, enforcing a charge or spin constraint on some subset of the system of interest.[5] The CDFT equations rely on a partitioning of the system into multiple fragments, and a means for assigning (spin-)density to individual

atoms.[5]

For each state I in the CDFT-CI active space, one minimizes the DFT energy subject to the constraint that the average spin and charge on each fragment takes on the value specified for that state/fragment pair. The CDFT algorithm enforces the constraint by applying a potential, $\hat{V}_I(\vec{r})$, to the system. A more detailed description of the working equations is presented in the previous chapters, and also references 5 and 6.

We take the time to note once again that the integer constraints that we might naïvely apply to molecular fragments from our chemical intuition (e.g. $N = \pm 1$, $S = \frac{1}{2}$) are not always reasonable. This can be due to charge- or spin-localized states that are inherently diffuse, or just to the inability of the charge model to describe bonded systems. (We recall that the notion of atomic charge within a molecular system is not well-defined, so we inherently must use an arbitrary scheme.[52]) This leads to a need to modify the given (intuitive) constraint values to account for the contribution to the computed "charge" on one fragment from the density tails of other fragments. This is accomplished by means of a "promolecule" formalism, in which the fragments of the constrained system are treated as independent systems, with integer charge and spin (as prescribed by the integer constraint values). A self-consistent converged density is obtained for each such fragment, and the sum of these fragment densities is integrated against the charge/spin prescription to determine the values of N and S that are used in the final calculation for the constrained state. A more in-depth discussion of the need for modified constraint values may be found in reference 9.

Once we have obtained a set of CDFT states (using the corrected N and S), we can then proceed to construct our CI matrix and the corresponding nonorthogonal secular equation:

$$\begin{pmatrix} H_{11} & H_{12} & \cdots & H_{1N} \\ H_{21} & H_{22} & & H_{2N} \\ \vdots & & \ddots & \vdots \\ H_{N1} & H_{N2} & \cdots & H_{NN} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} = E \begin{pmatrix} 1 & S_{12} & \cdots & S_{1N} \\ S_{21} & 1 & & S_{2N} \\ \vdots & & \ddots & \vdots \\ S_{N1} & S_{N2} & \cdots & 1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \tag{3.1}$$

The diagonal elements of **H** are just the energies of the constrained states that form the

basis for the active space; the off-diagonal elements are constructed as:[9,74]

$$H_{IJ} = H_{JI} = \frac{F_I + F_J}{2} S_{IJ} - \left\langle \Phi_I \left| \frac{\hat{V}_I + \hat{V}_J}{2} \right| \Phi_J \right\rangle \tag{3.2}$$

where Φ_I is the Kohn-Sham determinant for the Ith CDFT state, F_I is the energy of the Ith CDFT state in the presence of the constraining potential $\hat{V}_I$, and S_{IJ} is just $\langle \Phi_I | \Phi_J \rangle$. In the context of this configuration-interaction calculation, it is very natural to think of the CDFT states as being diabatic states, and these H_{12} matrix elements as the diabatic couplings between them. (Note that these couplings as written are computed in a non-orthogonal basis, and are only useful in their own right after transformation to an orthogonalized basis; unfortunately, such an orthogonalization does lose some information about the nature of the diabatic states.) The diabatic nature of the states is a consequence of how they are formed — they are explicitly constructed to have charge/spin distributions that are independent of nuclear position.

The CDFT-CI matrix diagonalization thus produces adiabatic states, and we expect that the true character of the ground *and excited* states of the system will be equally well-represented in them. Both ground and excited states arise from the same CI diagonalization, and are thus treated on an equal footing. This is in contrast to methods like TD-DFT, which generate an SCF ground state as a reference state and then seek to treat excited states as (single) excitations from that reference.[154,155] Such single-reference calculations are particularly prone to failure in the vicinity of conical intersections, where the nature of the exact ground state changes rapidly in a fashion that is very difficult for DFT methods to reproduce — the procedure for obtaining the DFT ground state has no mechanism to respond to low-lying excited states. Furthermore, the response state essentially can only account for single excitations, and this is frequently insufficient to describe the full conical nature of the intersection.[148] CDFT-CI can fully treat any number of different excitations, provided that the appropriate diabatic states are included in the configuration interaction. CDFT-CI may even prove superior to methods such as MR-CISD and CASPT2 in providing a consistent treatment of excited states of different characters, without invoking the extreme

computational expense of coupled-cluster calculations.

3.2 Results

We have implemented CDFT-CI in a development version of Q-Chem 3.2; the calculations described in this work were performed using the B3-LYP functional with the cc-pVDZ basis for water and the 6-31G basis for trihydrogen. TD-DFT and CASSCF calculations were performed using Gaussian 03. For the CDFT calculations underlying the CDFT-CI framework, atomic charges (and thus constraining potentials) were determined by applying Becke's multicenter integration scheme[50] against the (spin) density and attributing the results to the corresponding atomic centers.

3.3 H_3

The first system we consider is the simplest system to possess a conical intersection — trihydrogen. An intersection is symmetry-constrained to occur at all equilateral triangular geometries; to choose a particular one, we scanned over the symmetric "breathing" mode to find the lowest-energy such state. This was found to be at $R = 1.104$ Å for ground-state B3-LYP, $R = 1.198$ Å for full CI, and $R = 1.336$ Å for CDFT-CI. The well is rather shallow, and depends fairly strongly on the size of the basis set. We then held fixed two hydrogens on the y-axis at $\pm\frac{R}{2}$ and scanned over the x- and y-coordinates of the third hydrogen. As shown in Figure 3-1, TD-DFT does locate an intersection of electronic states with two splitting coordinates (three such intersections, actually!), but they are offset from the equilateral geometry, and furthermore are qualitatively incorrect — the upper state at the intersection is not cone-like, being instead a sharp cusp. For CDFT-CI, we considered three diabatic states in our CI matrix. In each state, we forced all of the excess spin density to localize on a single H atom; we did this for each H atom in turn. Figure 3-1 shows how the CDFT-CI surfaces meet in a well-formed dual-lobed cone that mirrors the full CI result.

3.4 H_2O

Like H_3, the water system has only three atoms. However, we now seek an intersection at a (again, symmetry-constrained) linear geometry. Since we know that a seam of conical intersections exists at linear geometries,[156] we can accurately locate an intersection by scanning over symmetric linear geometries; in this case, we find an intersection at $R_{\mathrm{O-H}} = 1.484$ Å for TD-DFT, $R_{\mathrm{O-H}} = 1.355$ Å for CAS(6,9), and $R_{\mathrm{O-H}} = 1.355$ Å for CDFT-CI. Using this geometry as the center for our scans over the internal coordinates for the symmetric stretch and bend, we produce the data plotted in Figure 3-2. Clearly, TD-DFT fails to describe the conical intersection, as only one splitting direction is found, instead of the correct two. The excitation energies away from the intersection are also too small (less than 1 eV) for larger R (and $\theta \approx 180°$). The TD-DFT method is not flexible enough to fully describe the excited state in the vicinity of the intersection. For CDFT-CI, in constructing our diabatic basis, we consider an active space of four states; in particular, we can make a covalent state with the oxygen atom a triplet ($S = \pm 1$) which is paired with a triplet "H_2" ($S = \mp 1$). We also include the two ionic states OH^-/H^+ and H^+/OH^- to fill out the four. The CDFT-CI surfaces meet at a well-defined cone, and smoothly vary away from the intersection; the comparison with the CAS surfaces is quite favorable.

3.5 Discussion

We note that attaining the correct qualitative behavior does rely on some amount of chemical intuition in the selection of constrained states for the CI matrix. Some preliminary CDFT-CI calculations on water with the ionic states in the active space replaced by states with $S = 0$ constraints on both O and "H_2" produced a conical intersection at the unphysical $R_{\mathrm{O-H}}$ of 0.94 Å! This active-space dependence is similar to the behavior of Complete Active Space wavefunction methods — a CAS(2,2) calculation on water also finds an intersection in the vicinity of 0.9 Å, whereas CAS(4,4) is in the correct area, near 1.4 Å. Poorly-chosen active spaces sometimes fail to yield an intersection at all. Experimentation and/or prior

knowledge of the nature of the states of interest is needed in order to perform reliable CDFT-CI calculations, but we expect that to some extent this can be avoided by using a larger active space as can be done with MCSCF. However, large active spaces have not been necessary for the small systems considered in this chapter.

3.6 Conclusions

We find that CDFT-CI is an effective DFT method for computing qualitatively-correct excited states, even for difficult cases such as the vicinity of conical intersections. This is a dramatic improvement over TD-DFT, which completely fails to give a proper description of the intersection (and thus the surfaces themselves in the vicinity of the intersection). We find this to be a very promising result, and plan for future work to assess the quantitative accuracy of CDFT-CI excited state energetics against reference wavefunction-based calculations. In light of the accuracy of CDFT-CI ground-state energies and barrier heights,[9,74] we think that the method is very promising for excited states as well. Future work should be performed to test the robustness of the CDFT-CI method to the size of the AO basis set used and the exchange-correlation functional used for the underlying CDFT calculations. Deeper questions that remain include the sensitivity of the method to the mechanism for enforcing the constraints that define the diabatic states: does the promolecule prescription for modifying the density constraints produce universally better results than the constraints given from naïve chemical intuition? The CDFT-CI method also makes an additional use of the weight (charge) prescription that is not present in ordinary CDFT, since it computes off-diagonal matrix elements of the constraining potential. The sensitivity of the couplings to the weight prescription should be further explored. Finally, if the predicted excited state surfaces computed by CDFT-CI prove to be accurate and robust, it will be useful to implement analytic gradients for the CDFT-CI states, which will facilitate excited-state dynamics and the location of minimal-energy conical intersections.

The bulk of this chapter has been published in reference 75.

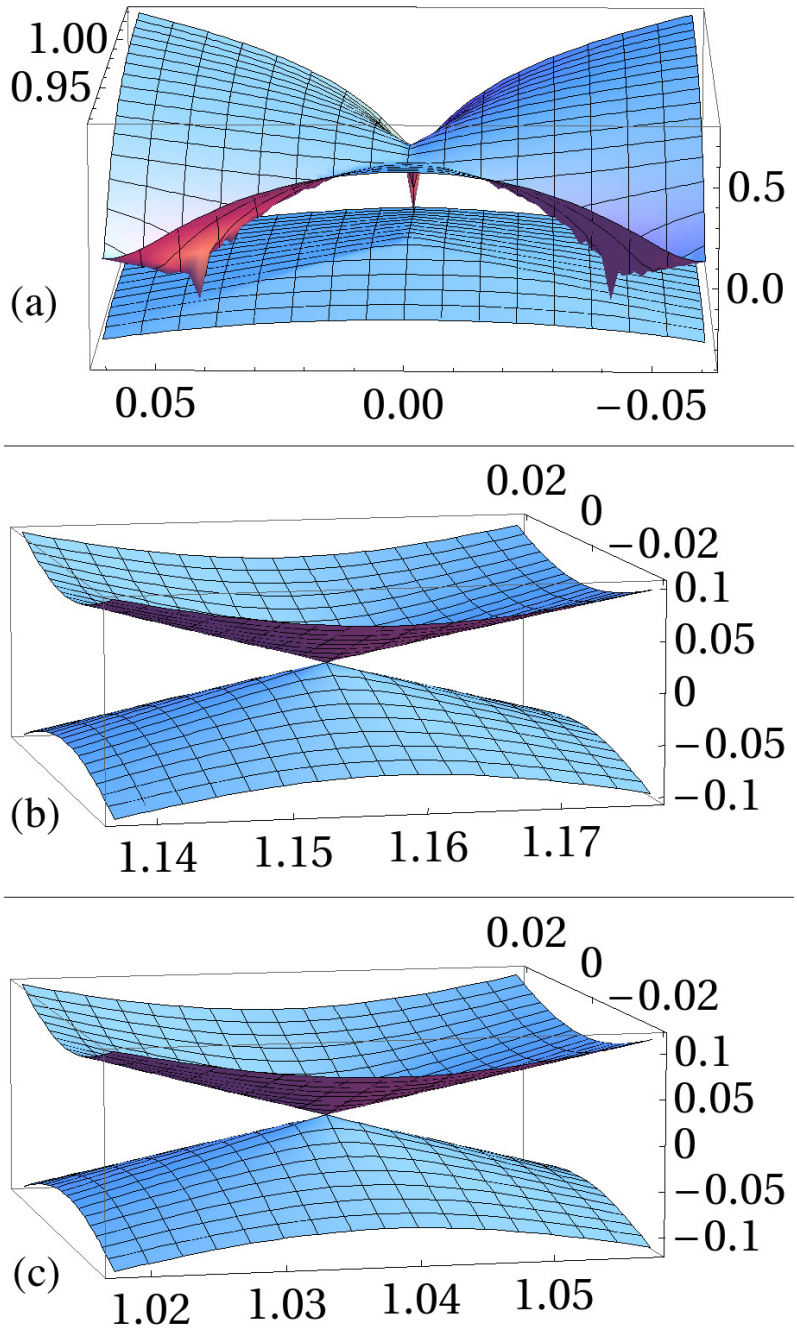

Figure 3-1: Triangular trihydrogen energy manifolds, as computed by (a) TD-DFT, (b) CDFT-CI, and (c) full CI. Note the different energy scale for TD-DFT. Reprinted with permission from reference 75. Copyright 2010, American Institute of Physics.

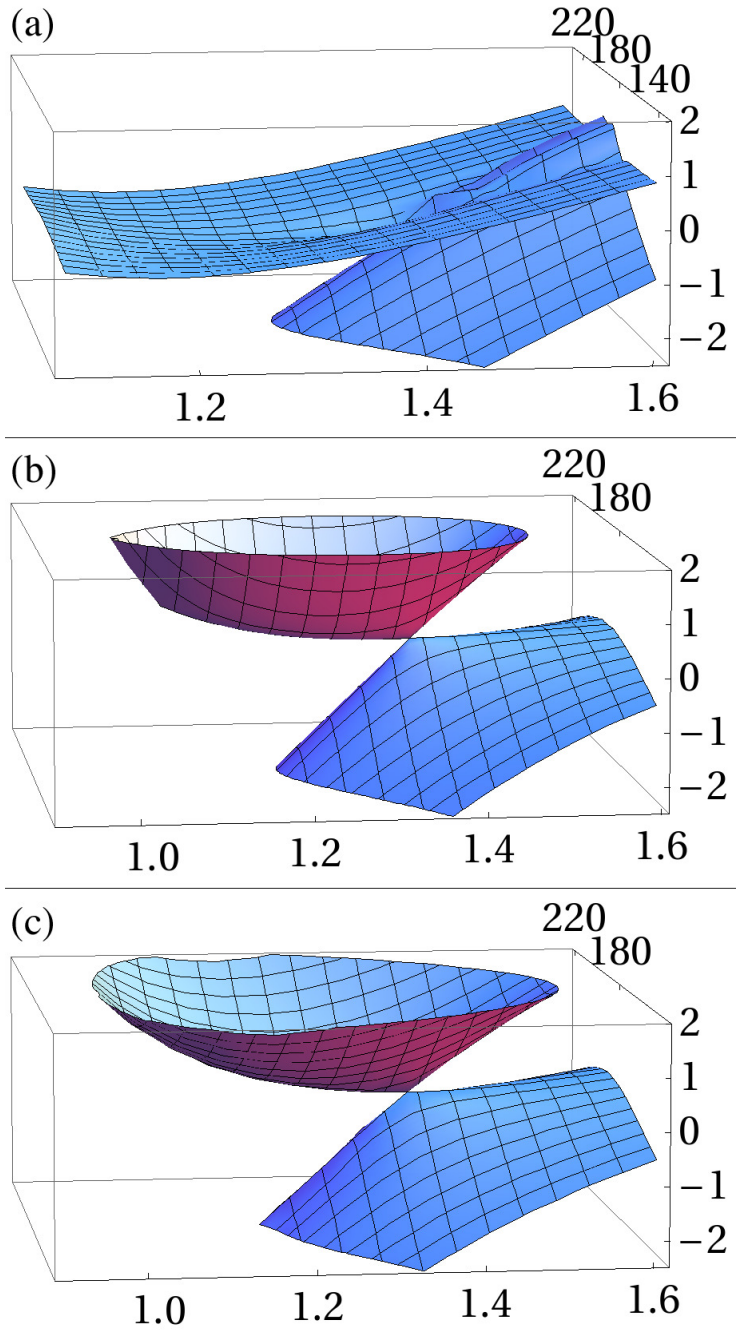

Figure 3-2: Symmetric water energy manifolds, as computed by (a) TD-DFT, (b) CDFT-CI, and (c) CAS(6,9). Reprinted with permission from reference 75. Copyright 2010, American Institute of Physics.

Chapter 4

Efficient Geometry Optimization

CDFT-CI has previously been used to calculate ground-state energies[9] and barrier heights,[74] and to describe electronic excited states, in particular conical intersections.[75] However, the method has previously been limited to evaluating the electronic energy at just a single nuclear configuration, with the gradient of the energy being available only via finite difference. In this chapter, we present analytic gradients of the CDFT-CI energy with respect to nuclear coordinates, which allows the potential for geometry optimization on both the ground and excited electronic states, a realm which is currently quite challenging for electronic structure theory.

4.1 Introduction

Electronic excited states are of interest in a great many chemical systems, being of relevance to photochemistry,[157–164] photodamage to DNA,[165–174] organic semiconductors,[127,175–178] and more.[179–181] Of particular interest is the *dynamics* on the excited state, after an excitation event has occurred; in order to study the geometric relaxation of electronic excited states, one requires the force on the excited-state PES. This requirement limits the spectrum of electronic structure methods which are usable, with the field being limited to TD-DFT,[182–190] configuration–interaction singles (CIS),[163,186,191–196] CASSCF and CASPT2[197–203] (note that

analytic gradients for CASPT2 were not implemented until 2003![204]), equation-of-motion coupled-cluster singles and doubles (EOM-CCSD)[205–209] and its approximate form (EOM-CC2),[186,210–216] and sometimes multi-reference configuration interaction (MRCI).[217–220] Excited state gradients do not appear to be available for coupled-cluster triples methods, which would in any case be too computationally expensive to apply to systems with more than a handful of atoms, like their ground-state counterparts. Even for EOM-CCSD, CASPT2, and MRCI, the computational expense will vary with the implementation and application, and such calculations become impractical for systems with more than 10 or 15 atoms. DFT methods gain a significant advantage of practicality as the system size increases.

TD-DFT has seen broad use for electronic excited states in general, and excited-state dynamics and geometry optimization are no exception.[185–190] However, it still suffers from the deficiencies in describing multiple excitations and charge-transfer excitations which render it a less-than-general solution for vertical excitation energy calculations,[75,148,150,153,221–226] though recent developments show the state of affairs may be improving.[227–230] Constrained DFT (CDFT) is designed to directly construct charge- and spin-constrained states and as such can find charge-transfer states directly, using self-consistent ground state techniques.[5,6,8] The self-consistent nature of the solution means that nonlinear response of the density is included, and as such in principle permits the treatment of multiple excitations from the ground state. However, CDFT has limitations of its own; it is still a single-reference method (and as such suffers from the limitations of DFT in the face of strong static correlation), and it has no effective prescription for describing valence excitations. CDFT-configuration interaction (CDFT-CI) explicitly introduces multiple configurations to the electronic structure treatment, which greatly improves results for situations where static correlation is strongly present, such as dissociation curves and reaction transition states.[9,74] Additionally, it can treat the ground state and excited states on the same footing, as for conical intersections.[75] However, applications of CDFT-CI have heretofore remained somewhat limited due to the unavailability of gradients of the electronic energy; in this work, we present the theory and implementation of analytic energy gradients for CDFT-CI. These forces are used to optimize

the excited-state geometry of several systems, and also to optimize the transition-state geometries for a standard set of reaction barriers. In most cases, the energy does not change noticeably from the reference transition-state geometry to the optimized geometry, indicating that CDFT-CI-optimized geometries for transition states are of comparable quality to the reference geometries. Geometry optimization on the excited state also converges to geometries of acceptable quality.

4.2 Methods

We briefly summarize the equations of CDFT and CDFT-CI before proceeding to the derivation of expressions for the gradient of the energy. CDFT takes as input a density functional giving the energy $E[\rho]$ and adds a constraint Lagrange multiplier term to yield a new functional

$$E[\rho, V_k] = E[\rho] + \sum_k V_k \left(\int \hat{w}_k \rho(\mathbf{r}) d\mathbf{r} - N_k \right) \tag{4.1}$$

where $\hat{w}_k$ is a "weight" operator that probes the number of electrons in some particular region of space and N_k is a target value for that operator. Minimizing E with respect to ρ and maximizing with respect to V_k yields a state with the desired constrained charge and spin properties. The effect of the constraint terms in the energy expression are equivalent to adding an additional "constraint" potential $\sum_k V_k \hat{w}_k(\mathbf{r}) = \sum_k \hat{V}_k(\mathbf{r}) = \hat{V}$ acting on the electrons in Kohn-Sham theory.

CDFT-CI requires the user to specify a collection of different constrained states which are used as a basis/active space for constructing a configuration-interaction matrix; the basis states are specified just by charge and spin constraints on particular fragments of the system in question. This CI matrix then gives the CDFT-CI eigenvalue equation,

$$\begin{pmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \mathcal{E} \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \end{pmatrix}$$

We show only the two-state case, but the generalization to N states is easily made. The

diagonal elements of $\mathbf{H}$ are just the energies of the constrained states that form the basis for the active space; the off-diagonal elements are constructed as:[9,74]

$$H_{IJ} = H_{JI} = \frac{F_I + F_J}{2} S_{IJ} - \left\langle \Phi_I \middle| \frac{\hat{V}_I + \hat{V}_J}{2} \middle| \Phi_J \right\rangle \tag{4.2}$$

where Φ_I is the Kohn-Sham determinant for the Ith CDFT state, F_I is the energy of the Ith CDFT state in the presence of the constraining potential $\hat{V}_I$ for state I, and S_{IJ} is just $\langle \Phi_I | \Phi_J \rangle$. For the rest of this work, we will assume the two-state form, using capital letters I and J to index the different states; extension to the N state case is straightforward. We also attempt to adhere to the convention that i and j index occupied orbitals, a and b index virtual orbitals, and p and q index all orbitals, occupied and unoccupied.

If we write $\mathbf{H}C = \mathcal{E}\mathbf{S}C$, then we can easily take the derivative with respect to a nuclear coordinate x and write

$$\mathbf{H}^x C + \mathbf{H}C^x = \mathcal{E}^x \mathbf{S}C + \mathcal{E}\mathbf{S}^x C + \mathcal{E}\mathbf{S}C^x \tag{4.3}$$

The quantity we are interested in is $\mathcal{E}^x$, which we must solve for. Bracketing on the left with $C^\dagger$ and rearranging yields

$$\mathcal{E}^x = C^\dagger \left(\mathbf{H}^x - \mathcal{E}\mathbf{S}^x \right) C \tag{4.4}$$

$\mathcal{E}$ and C are already known from the single-point energy evaluation, so the only new terms required for the gradient expression are $\mathbf{H}^x$ and $\mathbf{S}^x$.

4.2.1 Overview

In order to actually use equation (4.4) to obtain $\mathcal{E}^x$, we must consider both diagonal terms of the form $\mathbf{H}^x_{II}$ and off-diagonal terms $\mathbf{H}^x_{IJ}$ and $\mathbf{S}^x_{IJ}$. (There are no diagonal overlap terms, since normalized states will always have unit self-overlap.) The diagonal terms $\mathbf{H}^x_{II}$ are just the energy gradient of the constrained states, so we focus on the off-diagonal elements $\mathbf{H}^x_{IJ}$ and $\mathbf{S}^x_{IJ}$. However, since we only use these two quantities in combination, for reasons that

will be discussed later, we define an auxiliary quantity $\mathbf{W} = \mathbf{H} - \mathcal{E}\mathbf{S}$ where $\mathcal{E}$ is treated as a *constant* and does not vary with changes in any other parameters. We then seek to compute $\mathbf{W}^x = (\mathbf{H} - \mathcal{E}\mathbf{S})^x = \mathbf{H}^x - \mathcal{E}\mathbf{S}^x$.

At this point, the high-level view is no longer sufficient and we must expand the expression to include our Kohn-Sham determinants that define the states in our CI basis.

$$W_{II} = H_{II} \tag{4.5}$$

$$W_{II}^x = H_{II}^x \tag{4.6}$$

$$W_{IJ} = \left(\frac{F_I + F_J}{2} - \mathcal{E}\right)\langle\Phi_I|\Phi_J\rangle - \left\langle\Phi_I\middle|\frac{\hat{V}_I + \hat{V}_J}{2}\middle|\Phi_J\right\rangle \tag{4.7}$$

$$W_{IJ}^x = \frac{F_I^x + F_J^x}{2}S_{IJ} - \left\langle\Phi_I\middle|\frac{\hat{V}_I^x + \hat{V}_J^x}{2}\middle|\Phi_J\right\rangle + \left\langle\Phi_I^x\middle|\hat{O}\middle|\Phi_J\right\rangle + \left\langle\Phi_I\middle|\hat{O}\middle|\Phi_J^x\right\rangle \tag{4.8}$$

$$\hat{O} = \left(\frac{F_I + F_J}{2} - \mathcal{E}\right) - \frac{\hat{V}_I + \hat{V}_J}{2} \tag{4.9}$$

However, the terms in $\langle\Phi_I^x|$ and $|\Phi_J^x\rangle$ are unreasonable to compute, given that the wavefunction gradient requires $O(N^3)$ space to store and $O(N^5)$ time to compute. As such, we seek an alternate route to W_{IJ}^x which does not involve the gradient of the wavefunction; we will adopt the standard framework of making the matrix element variational.[204]

The procedure to make the matrix element variational is quite complex, and does not have a linear logical order of execution. To assist in understanding the processes involved in the computation, a flow chart of the relevant expressions is presented in Figure 4-1. The labeled boxes in the flowchart correspond roughly to the subsections that follow, though we start off with the general computation of the gradient of a matrix element, used for both the promolecule contribution and the coupling element derivative.

4.2.2 Assembling a matrix element/coupling derivative

Looking forward, we will need to evaluate several expressions of similar form, so we step back from W_{IJ}^x and consider the general case of the matrix element of a one-electron (or

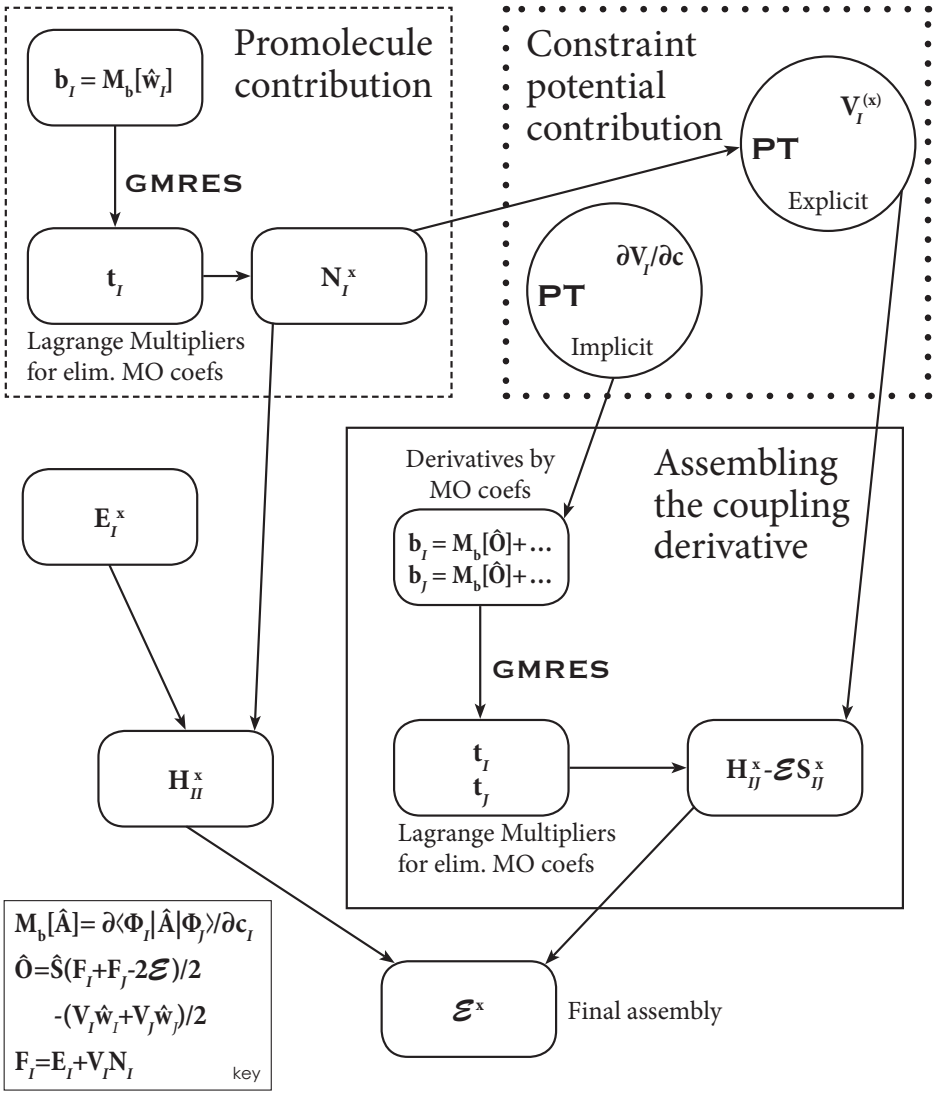

Figure 4-1: The flowchart for evaluating the gradient of the CDFT-CI energy. The gradient of the promolecule-adjusted constraint values is computed as a variational matrix element (dashed box), using the GMRES linear solver to obtain the Lagrange multipliers needed for variationality. These N_I^x are combined with the CDFT energy gradient to yield the diagonal elements of the Hamiltonian (bottom left), and also used to determine the explicit dependence of the constraint potential Lagrange multipliers on the nuclear coordinates, using a perturbation theory expression (upper right). A separate perturbation theory expression (also in the dotted box) gives the dependence of the constraint potential Lagrange multipliers on the MO coefficients (and thus the implicit dependence on nuclear position), which enters into the variationality of the coupling derivative (solid box), computed in a similar fashion to the promolecule contribution. With both diagonal and off-diagonal matrix elements available, the energy gradient is evaluated per equation (4.4) (bottom).

zero-electron) operator $\hat{O}$ between two states $|\Phi_I\rangle$ and $|\Phi_J\rangle$, which has been normalized for possible variations of the orbitals which underlie the states.

$$M[\hat{O}] = \frac{\left\langle \Phi_I \middle| \hat{O} \middle| \Phi_J \right\rangle}{\sqrt{\langle \Phi_I | \Phi_I \rangle \langle \Phi_J | \Phi_J \rangle}} \tag{4.10}$$

$$= \frac{\mathrm{Tr}\left[\mathbf{c}_I^\dagger \mathbf{O} \mathbf{c}_J \left(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_I \right)^{-1} \right] \det(\mathbf{c}_I^\dagger \mathbf{S} \mathbf{c}_J)}{\sqrt{\det(\mathbf{c}_I^\dagger \mathbf{S} \mathbf{c}_I) \det(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_J)}} \tag{4.11}$$

We will primarily be concerned with the case of $\hat{O} = (\hat{V}_I + \hat{V}_J)/2$ and also the case where there is no operator (so the matrix element becomes just an overlap), but will also need this machinery for $\hat{O} = \hat{w}$ in a special case, per section 4.2.3. To avoid the need for the expensive wavefunction gradient (or, equivalently, the gradient of the MO coefficients), we introduce an auxiliary function which has the same value as M at all points but is variational with respect to the MO coefficients; this allows the gradient of the MO coefficients to be ignored via the chain rule. We thus write

$$M^{\mathrm{var}} = M[\hat{O}] - L(\mathbf{c}_I, \mathbf{t}_I) - L(\mathbf{c}_J, \mathbf{t}_J) \tag{4.12}$$

Lagrange multipliers to eliminate dependence of M on the MO coefficients

This term L depends on Lagrange multipliers $\mathbf{t}$ which enforce variationality; as there are $N_{occ} \times N_{basis}$ relevant MO coefficients, so there are $N_{occ} \times N_{basis}$ Lagrange multipliers $\mathbf{t}$, coming in as

$$L(\mathbf{c}, \mathbf{t}) = \mathrm{Tr}\left[\mathbf{t}^\dagger \cdot \left(\mathbf{F}[\tilde{\mathbf{P}}]\mathbf{c} - \mathbf{S}\mathbf{c}\epsilon \right) \right] \tag{4.13}$$

where $\tilde{\mathbf{P}} = 3\mathbf{PSP} - 2\mathbf{PSPSP}$ is the density matrix after McWeeny's purification transformation, $\mathbf{P} = \mathbf{c}\mathbf{c}^\dagger$ is the density matrix, and $\mathbf{S}$ remains the atomic orbital overlap matrix. ϵ is a diagonal matrix of MO energies, which we define to be evaluated as $\epsilon_i = \frac{c_i^\dagger \mathbf{F} c_i}{c_i^\dagger \mathbf{S} c_i}$ to remain normalized when the orbitals themselves become unnormalized, per equation (4.11). We also introduce the Fock matrix, which has dependence on both the MO coefficients and the

nuclear position, but we do not need to enforce that $\partial M^{\text{var}}/\partial \mathbf{F} = 0$. Equation (4.13) is constructed such that the quantity in parentheses will always evaluate to zero when the system is at SCF convergence. Accordingly, L will also always be zero at convergence, so M^{var} will have the same value as M. Furthermore, $\partial M^{\text{var}}/\partial \mathbf{r}$ will also be zero by the self-consistency condition, which removes any need for gradients of $\mathbf{t}$ in evaluating chain-rule terms. (Note that the actual values of $\mathbf{t}$ are as-yet unspecified.) The McWeeny purified density matrix is required so that changes in the MO coefficients which do not preserve normalization do not affect the resulting Fock matrix; changes in normalization of the MO coefficients at first order will only affect the purified density matrix at the second order. No correction is needed for the MO coefficients that appear directly in $\mathbf{Fc} - \mathbf{Sc}\epsilon$, as that expression is merely enforcing that the orbitals remain eigenvectors; a change in normalization does not affect that condition.

Derivatives by MO coefficients To enforce the variationality of M^{var}, we require

$$\frac{\partial M^{\text{var}}}{\partial \mathbf{c}} = 0 \tag{4.14}$$

$$\Rightarrow \quad \frac{\partial M}{\partial \mathbf{c}} = \frac{\partial L(\mathbf{c}, \mathbf{t})}{\partial \mathbf{c}} \tag{4.15}$$

for derivatives with respect to both $\mathbf{c}_I$ and $\mathbf{c}_J$. Both sides of this equation expand out to sizable expressions, which we treat singly.

$$\begin{aligned}\frac{\partial M}{\partial \mathbf{c}_I} &= \frac{1}{\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}}\left[\det\left(\mathbf{c}_I^\dagger\mathbf{S}\mathbf{c}_J\right)\mathbf{O}\mathbf{c}_J\left(\mathbf{c}_J^\dagger\mathbf{S}\mathbf{c}_I\right)^{-1}\right. \\ &\quad \left.-\mathbf{S}\mathbf{c}_J\left(\mathbf{c}_J^\dagger\mathbf{S}\mathbf{c}_I\right)^{-1}\mathbf{c}_I^\dagger\mathbf{O}\mathbf{c}_J\left(\mathbf{c}_J^\dagger\mathbf{S}\mathbf{c}_I\right)^{-1}\right] \\ &\quad + M[\hat{O}]\mathbf{S}\mathbf{c}_J\left(\mathbf{c}_J^\dagger\mathbf{S}\mathbf{c}_I\right)^{-1} - M[\hat{O}]\frac{\det\left(\mathbf{c}_I^\dagger\mathbf{S}\mathbf{c}_I\right)}{\langle\Phi_I|\Phi_I\rangle}\mathbf{S}\mathbf{c}_I\left(\mathbf{c}_I^\dagger\mathbf{S}\mathbf{c}_I\right)^{-1} \qquad (4.16)\end{aligned}$$

$$\frac{\partial L(\mathbf{c},\mathbf{t})}{\partial \mathbf{c}} = \mathrm{Tr}\left[\mathbf{t}^\dagger\cdot\left(\frac{\partial \mathbf{F}[\tilde{\mathbf{P}}]}{\partial \mathbf{c}}\mathbf{c}\right)\right] + \left(\mathbf{t}^\dagger\cdot\mathbf{F}[\tilde{\mathbf{P}}]\right)^\dagger - \left(\epsilon\mathbf{t}^\dagger\cdot\mathbf{S}\right)^\dagger - \mathrm{Tr}\left[\mathbf{t}^\dagger\cdot\left(\frac{\partial\epsilon}{\partial\mathbf{c}}\mathbf{c}\right)\right] \qquad (4.17)$$

$$\begin{aligned}&= t_{\nu j}\frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}}c_{\lambda j} + t_{\nu i}F_{\nu\mu}[\tilde{\mathbf{P}}] - t_{\nu i}S_{\nu\mu}\epsilon_\mu - t_{\nu j}S_{\nu\delta}\frac{c_{\lambda j}\left(\partial F_{\lambda\sigma}[\tilde{\mathbf{P}}]/\partial c_{\mu i}\right)c_{\sigma j}}{c_{\alpha j}S_{\alpha\beta}c_{\beta j}}c_{\delta j} \\ &\quad - 2t_{\nu i}S_{\nu\delta}\frac{F_{\mu\sigma}[\tilde{\mathbf{P}}]c_{\sigma i}}{c_{\alpha i}S_{\alpha\beta}c_{\beta i}}c_{\delta j} + 2t_{\nu i}S_{\nu\delta}\frac{c_{\lambda i}F_{\lambda\sigma}[\tilde{\mathbf{P}}]c_{\sigma i}}{\left(c_{\alpha i}S_{\alpha\beta}c_{\beta i}\right)^2}S_{\mu\gamma}c_{\gamma i}c_{\delta i} \qquad (4.18)\end{aligned}$$

However, we can see that the overall structure of the expression is really

$$M_b[\hat{O}] \equiv \frac{\partial M}{\partial c_{\mu i}} = b_{\mu i} = A^{\nu j}_{\mu i}t_{\nu j} \qquad (4.19)$$

where we have adopted explicit indices and the Einstein summation convention to make clear the complicated structure of some of these terms, and defined a new quantity $M_b[\hat{O}]$ for future use.[a] The final line, however, makes it clear that this is essentially just a linear system which can be solved for $\mathbf{t}$ in a conceptually straightforward fashion, regardless of the complexity of some of the terms. The $\mathbf{A}$ matrix that is involved in this linear system, however, is of size $(N_{occ}\cdot N_{virt})\times(N_{occ}\cdot N_{virt})$ which requires $O(N^4)$ storage and $O(N^6)$ time for a direct inversion, a step backwards from equation (4.9). However, we can solve the linear system in equation (4.19) without constructing $\mathbf{A}$, by using an iterative linear solver. This allows us to leverage the fact that the product $A^{\nu j}_{\mu i}x_{\nu j}$ may be evaluated efficiently without computing $\mathbf{A}$.

[a] This definition does give preference to $\mathbf{c}_I$ over $\mathbf{c}_J$, but this is immaterial given that the structure of M is symmetric with respect to the two states.

We have conveniently left unexpanded the expression $\partial F_{\lambda\sigma}[\tilde{\mathbf{P}}]/\partial c_{\mu i}$, a quantity whose determination is complicated by the use of the purified density $\tilde{\mathbf{P}}$. Performing this computation requires a breakdown of the different contributions to $\mathbf{F}$, with the Coulomb integrals and Hartree-Fock exchange being treated differently from DFT exchange and correlation functionals. (The one-electron Hamiltonian of course has no dependence on the MO coefficients.) Bearing in mind our need to compute only the product $\mathbf{A}x$ and not $\mathbf{A}$ itself, we examine the contraction of some pseudo-density-matrix $W_{\lambda\sigma}$ against $\partial F_{\lambda\sigma}[\tilde{\mathbf{P}}]/\partial c_{\delta i}$, looking at each of these terms in turn.

$$\frac{\partial J_{\lambda\sigma}[\tilde{\mathbf{P}}]}{\partial c_{\delta i}} W_{\lambda\sigma} = \frac{\partial}{\partial c_{\delta i}} \left((\mu\nu|\lambda\sigma)(3P_{\mu\alpha}S_{\alpha\beta}P_{\beta\nu} - 2P_{\mu\alpha}S_{\alpha\beta}P_{\beta\gamma}S_{\gamma\eta}P_{\eta\nu}) \right) W_{\lambda\sigma} \tag{4.20}$$

$$\begin{aligned} &= 6J_{\mu\delta}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} + 6S_{\delta\beta}P_{\beta\nu}J_{\mu\nu}[\mathbf{W}]c_{\nu i} \\ &- 4J_{\mu\delta}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}P_{\beta\gamma}S_{\gamma\eta}c_{\eta i} - 4S_{\delta\beta}P_{\beta\gamma}S_{\gamma\eta}P_{\eta\nu}J_{\mu\nu}[\mathbf{W}]c_{\mu i} \\ &- 4S_{\delta\eta}P_{\eta\nu}J_{\mu\nu}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} \end{aligned} \tag{4.21}$$

$$\begin{aligned} &= 2J_{\mu\delta}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} + 2S_{\delta\beta}P_{\beta\nu}J_{\mu\nu}[\mathbf{W}]c_{\nu i} \\ &- 4S_{\delta\eta}P_{\eta\nu}J_{\mu\nu}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} \end{aligned} \tag{4.22}$$

where we use the fact that $\mathbf{PSP} = \mathbf{P}$ at convergence. This requires only a single Coulomb build from the pseudo-density $\mathbf{W}$, and matrix multiplications with $\mathbf{S}$ and $\mathbf{P}$.

In a similar fashion, the expression for the exchange derivative becomes

$$\begin{aligned} \frac{\partial K_{\lambda\sigma}[\tilde{\mathbf{P}}]}{\partial c_{\delta i}} W_{\lambda\sigma} &= 2K_{\mu\delta}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} + 2S_{\delta\beta}P_{\beta\nu}K_{\mu\nu}[\mathbf{W}]c_{\mu i} \\ &- 4S_{\delta\eta}P_{\eta\nu}K_{\mu\nu}[\mathbf{W}]P_{\mu\alpha}S_{\alpha\beta}c_{\beta i} \end{aligned} \tag{4.23}$$

The DFT contributions are not quite as straightforward, in that a standard calculation of the XC matrix from the pseudodensity matrix is not the correct approach to this derivative. In fact, the XC matrix (for a pure functional) is more properly written like

$$v_{xc} = v_{xc}[\rho(\tilde{\mathbf{P}})]$$

From the chain rule,

$$\frac{\partial v_{xc}}{\partial c_{\delta i}} = \frac{\partial v_{xc}}{\partial \rho}\frac{\partial \rho}{\partial \tilde{\mathbf{P}}}\frac{\partial \tilde{\mathbf{P}}}{\partial c_{\delta i}} \tag{4.24}$$

$$= \left(\frac{\partial v_{xc}}{\partial \rho}\frac{\partial \rho}{\partial \tilde{\mathbf{P}}}\right)\frac{\partial \tilde{\mathbf{P}}}{\partial c_{\delta i}} \tag{4.25}$$

The quantity in parentheses is the implicit first derivative of the XC matrix, which is generally only used by being contracted against a "trial density", as we are doing here as we contract against $\mathbf{W}$. Thus,

$$\frac{\partial (v_{xc})_{\lambda\sigma}}{\partial c_{\delta i}} W_{\lambda\sigma} = \left(W_{\lambda\sigma}\frac{\partial (v_{xc})_{\lambda\sigma}}{\partial \rho}\frac{\partial \rho}{\partial \tilde{\mathbf{P}}_{\alpha\beta}}\right)\frac{\partial \tilde{\mathbf{P}}_{\alpha\beta}}{\partial c_{\delta i}} \tag{4.26}$$

$$= X_{\alpha\beta}\frac{\partial \tilde{\mathbf{P}}_{\alpha\beta}}{\partial c_{\delta i}} \tag{4.27}$$

$$= 2\mathbf{XPSc} + 2\mathbf{SPXc} - 4\mathbf{SPXPSc} \tag{4.28}$$

in an analogous fashion to the coulomb and exchange terms. We implicitly define the quantity $X_{\alpha\beta}$ as the contraction of $\mathbf{W}$ against the implicit first derivative of the XC matrix. All of these terms (Coulomb, exchange, and DFT) may be efficiently evaluated in $O(N^3)$ time or less and $O(N^2)$ space.

Formulation of the Lagrange multiplier linear system for an iterative solver The combination of the above three terms yields the overall contraction of $\partial\mathbf{F}/\partial\mathbf{c}$ against a trial density matrix $\mathbf{W}$. This, in turn, lets us return to $\partial L/\partial\mathbf{c}$ and the linear system $b_{\mu i} = A^{\nu j}_{\mu i} t_{\nu j}$ of equation (4.19). Having already determined that the system will be solved iteratively, we

now need a procedure for computing the matrix-vector product

$$A_{\mu i}^{\nu j} x_{\nu j} = x_{\nu j} \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} c_{\lambda j} + x_{\nu j} F_{\nu\mu}[\tilde{\mathbf{P}}] \delta_{ij} - x_{\nu j} S_{\nu\mu} \epsilon_\mu \delta_{ij} - x_{\nu j} S_{\nu\delta} \frac{c_{\lambda j} \left(\partial F_{\lambda\sigma}[\tilde{\mathbf{P}}] / \partial c_{\mu i} \right) c_{\sigma j}}{c_{\alpha j} S_{\alpha\beta} c_{\beta j}} c_{\delta j}$$
$$- 2 x_{\nu j} S_{\nu\delta} \frac{F_{\mu\sigma}[\tilde{\mathbf{P}}] c_{\sigma i}}{c_{\alpha i} S_{\alpha\beta} c_{\beta i}} c_{\delta i} \delta_{ij} + 2 x_{\nu j} S_{\nu\delta} \frac{c_{\lambda i} F_{\lambda\sigma}[\tilde{\mathbf{P}}] c_{\sigma i}}{\left(c_{\alpha i} S_{\alpha\beta} c_{\beta i} \right)^2} S_{\mu\gamma} c_{\gamma i} c_{\delta i} \delta_{ij} \tag{4.29}$$

$$= A_{0\,\mu i}^{\nu j} x_{\nu j} + \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} c_{\lambda j} x_{\nu j} - \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} c_{\lambda j} \frac{x_{\alpha j} S_{\alpha\beta} c_{\beta j}}{c_{\gamma j} S_{\gamma\delta} c_{\delta j}} c_{\nu j} \tag{4.30}$$

$$= A_{0\,\mu i}^{\nu j} x_{\nu j} + \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} c_{\lambda j} x_{\nu j} - \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} Y_{\lambda\nu} \tag{4.31}$$

$$= A_{0\,\mu i}^{\nu j} x_{\nu j} + \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} \left(c_{\lambda j} x_{\nu j} - Y_{\lambda\nu} \right) \tag{4.32}$$

$$= A_{0\,\mu i}^{\nu j} x_{\nu j} + \frac{\partial F_{\nu\lambda}[\tilde{\mathbf{P}}]}{\partial c_{\mu i}} W_{\lambda\nu} \tag{4.33}$$

$$A_{0\,\mu i}^{\nu j} x_{\nu j} = x_{\nu i} F_{\nu\mu}[\tilde{\mathbf{P}}] \delta_{ij} - x_{\nu i} S_{\nu\mu} \epsilon_\mu \delta_{ij}$$
$$- 2 x_{\nu i} S_{\nu\delta} \frac{F_{\mu\sigma}[\tilde{\mathbf{P}}] c_{\sigma i}}{c_{\alpha i} S_{\alpha\beta} c_{\beta i}} c_{\delta j} \delta_{ij} + 2 x_{\nu i} S_{\nu\delta} \frac{c_{\lambda i} F_{\lambda\sigma}[\tilde{\mathbf{P}}] c_{\sigma i}}{\left(c_{\alpha i} S_{\alpha\beta} c_{\beta i} \right)^2} S_{\mu\gamma} c_{\gamma i} c_{\delta i} \delta_{ij} \tag{4.34}$$

where we have explicitly constructed the pseudo-density matrix $\mathbf{W}$ to be contracted against $\partial F / \partial c$, and $\mathbf{Y}$ takes the form of an energy-weighted density matrix with orbital "energies"

$$\epsilon_j = \frac{x_{\alpha j} S_{\alpha\beta} c_{\beta j}}{c_{\gamma j} S_{\gamma\delta} c_{\delta j}} \tag{4.35}$$

$$Y_{\lambda\nu} = c_{\lambda j} \epsilon_j c_{\nu j} \tag{4.36}$$

It is worth noting that, if considered as a matrix, $\mathbf{A}_0$ is block-diagonal — each orbital only interacts with the corresponding "orbital" from x. In other words, all of the terms in A_0 include a Kronecker delta δ_{ij}, so there is only $O(N^3)$ work to be done in the overall multiplication.

Iterative linear solver (GMRES)

Now that the matrix-vector product is available, we can proceed to the iterative linear solver. We have implemented the GMRES (Generalized Minimum Residual) algorithm in Q-Chem; the algorithm is covered extensively elsewhere,[231,232] but we give a brief summary here and expound a bit more in Appendix A. The goal is to construct an approximate solution to the linear system

$$\mathbf{A} \cdot x = b$$

without explicitly operating on the matrix $\mathbf{A}$, instead only evaluating matrix-vector products $\mathbf{A} \cdot x_i$. GMRES requires the matrix $\mathbf{A}$ to be square. The outline of the solution is to find an approximate solution in a projected subspace at each step; by preconditioning the system appropriately, an approximate solution with sufficiently small residual can be obtained in a constant number of iterations, essentially independent of the dimension of $\mathbf{A}$. For the systems we consider here, that constant is around twenty iterations.

Preconditioning

The GMRES method is most effective for diagonal-dominant matrices (or rather, those whose eigenvalues are clustered near each other in the complex plane, which is not quite the same), frequently exhibiting quadratic convergence for such systems.[231] It is therefore quite common to introduce a (left) preconditioner, an approximation $\mathbf{A}_0 \approx \mathbf{A}$ which may be easily inverted, and then applying the GMRES algorithm to the preconditioned system:

$$\left(\mathbf{A}_0^{-1}\mathbf{A}\right) x = \mathbf{A}_0^{-1} b$$

Our linear system of equation (4.19) is not diagonal dominant, so we will need some non-trivial preconditioning strategy for GMRES to converge efficiently. However, there is a block-diagonal component to our $\mathbf{A}$, which in general is much larger in magnitude than the contribution coming from contractions against the derivative of the Fock matrix. This block-diagonal term, already referred to as $\mathbf{A}_0$ above, could be explicitly constructed with

$O(N^3)$ effort and blockwise inverted for $O(N^4)$ effort, but it proves more convenient to only consider the first two terms of $\mathbf{A}_0$,

$$A'^{\,\nu j}_{0\mu i} = \delta_{ij}\left(F_{\mu\nu} - S_{\mu\nu}\epsilon_j\right) \tag{4.37}$$

since in this formulation computing $\mathbf{A}_0'^{-1}x$ will only require $O(N^3)$ work.

This inversion is equivalent to solving the systems

$$\left(\mathbf{F} - \mathbf{S}\epsilon_i\right)\mathbf{x}_i = \mathbf{b}_i \tag{4.38}$$

which does not necessarily involve explicitly constructing $\mathbf{A}_0'^{-1}$ in matrix form. Conceptually, this is effected by transforming to the MO basis, where $\mathbf{F}$ and $\mathbf{S}$ are diagonal, so the inversion is trivial. Transforming back to the AO basis yields the desired solution.

$$b_{\mu i} = \left(F_{\mu\nu} - S_{\mu\nu}\epsilon_i\right)x_{\nu i} \tag{4.39}$$

$$c_{\mu p}b_{\mu i} = c_{\mu p}\left(F_{\mu\nu} - S_{\mu\nu}\epsilon_i\right)x_{\nu i} \tag{4.40}$$

$$= c_{\mu p}\left(F_{\mu\nu} - S_{\mu\nu}\epsilon_i\right)c_{\nu q}c_{\lambda q}S_{\lambda\sigma}x_{\sigma i} \tag{4.41}$$

$$= \delta_{pq}\left(\epsilon_p - \epsilon_i\right)c_{\lambda q}S_{\lambda\sigma}x_{\sigma i} \tag{4.42}$$

$$\frac{1}{\epsilon_p - \epsilon_i}c_{\mu p}b_{\mu i} = c_{\lambda p}S_{\lambda\sigma}x_{\sigma i} \qquad p \neq i \tag{4.43}$$

$$\frac{1}{\epsilon_p - \epsilon_i}c_{\nu p}c_{\mu p}b_{\mu i} = c_{\nu p}c_{\lambda p}S_{\lambda\sigma}x_{\sigma i} \tag{4.44}$$

$$c_{\nu p}D_{pi} = x_{\nu i} \tag{4.45}$$

$$D_{pi} = \frac{1}{\epsilon_p - \epsilon_i}c_{\mu p}b_{\mu i} \tag{4.46}$$

The actual determination of $\mathbf{x}$ involves just matrix-matrix products, yielding the desired $O(N^3)$ time. Since the energy denominator is only nonzero when $p \neq i$, this preconditioner will not treat components of $b_{\mu i}$ which are proportional to actual orbitals $c_{\mu i}$; however, these components are zero by construction, being eliminated by the normalization denominator in equation (4.11).

Assembling the coupling derivative

Having amassed a great deal of machinery to determine the Lagrange multipliers $\mathbf{t}$ which make the function M^{var} of equation (4.12) actually variational with respect to the MO coefficients, we now return to the chain rule and actually make use of them. The terms from $\partial M^{\text{var}}/\partial \mathbf{t}$ have already been shown to be zero, but it turns out that our definition of L also introduced a dependence on $\mathbf{F}$ to the full M^{var} which must be included in the chain rule terms so that

$$\frac{dM^{\text{var}}}{dx} = \frac{\partial M^{\text{var}}}{\partial \mathbf{O}}\frac{d\mathbf{O}}{dx} + \frac{\partial M^{\text{var}}}{\partial \mathbf{S}}\frac{d\mathbf{S}}{dx} + \frac{\partial M^{\text{var}}}{\partial \mathbf{F}}\frac{d\mathbf{F}}{dx} \tag{4.47}$$

In evaluating these chain-rule terms, it proves convenient to repartition this expression as

$$\frac{dM}{dx} = \frac{dM^{\text{var}}}{dx} = M_x[\hat{O}] - L_x(\mathbf{c}_I, \mathbf{t}_I) - L_x(\mathbf{c}_J, \mathbf{t}_J) \tag{4.48}$$

$M[\hat{O}]$ depends on $\mathbf{O}$ and $\mathbf{S}$ but not $\mathbf{F}$, whereas L does not depend on $\mathbf{O}$ but only on $\mathbf{S}$ and $\mathbf{F}$, so this repartitioning simplifies the formal parameters somewhat. Putting $n = 1/\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}$ for concision, we can then write

$$\begin{aligned} M_x[\hat{O}] = n \cdot \mathrm{Tr}\left[\mathbf{c}_I^\dagger \mathbf{O}^x \mathbf{c}_J \left(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_I\right)^{-1}\right] \det(\mathbf{c}_I^\dagger \mathbf{S} \mathbf{c}_J) \\ + n \cdot \mathrm{Tr}\left[\mathbf{c}_I^\dagger \mathbf{O} \mathbf{c}_J \left(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_I\right)^{-1} \left(\mathbf{c}_J^\dagger \mathbf{S}^x \mathbf{c}_I\right) \left(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_I\right)^{-1}\right] \det(\mathbf{c}_I^\dagger \mathbf{S} \mathbf{c}_J) \\ + M[\hat{O}] \cdot \mathrm{Tr}\left[\left(\mathbf{c}_J^\dagger \mathbf{S} \mathbf{c}_I\right)^{-1} \left(\mathbf{c}_J^\dagger \mathbf{S}^x \mathbf{c}_I\right)\right] \\ - \frac{1}{2} M[\hat{O}] \frac{\mathbf{c}_I^\dagger \mathbf{S}^x \mathbf{c}_I}{\det(\mathbf{c}_I^\dagger \mathbf{S} \mathbf{c}_I)} \end{aligned} \tag{4.49}$$

The L contribution does not need the explicit normalization term n;

$$L(\mathbf{c}, \mathbf{t}) = \mathrm{Tr}\left[\mathbf{t}^\dagger \cdot \left(\mathbf{F}[\tilde{\mathbf{P}}]\mathbf{c} - \mathbf{S}\mathbf{c}\epsilon\right)\right] \tag{4.50}$$

$$L_x(\mathbf{c}, \mathbf{t}) = \mathrm{Tr}\left[\mathbf{t}^\dagger \cdot \left(\mathbf{F}^{(x)}[\tilde{\mathbf{P}}]\mathbf{c} - \mathbf{S}^x\mathbf{c}\epsilon - \mathbf{S}\mathbf{c}\epsilon^{(x)}\right)\right] \tag{4.51}$$

where

$$\epsilon_p^{(x)} = \frac{c_p^\dagger \mathbf{F}^{(x)}[\tilde{\mathbf{P}}] c_p}{c_p^\dagger \mathbf{S} c_p} - \frac{c_p^\dagger \mathbf{F}[\tilde{\mathbf{P}}] c_p}{\left(c_p^\dagger \mathbf{S} c_p\right)^2} c_p^\dagger \mathbf{S}^x c_p \tag{4.52}$$

and $\mathbf{F}^{(x)}[\tilde{\mathbf{P}}]$ is the partial derivative of the Fock matrix including the McWeeny purification of the density matrix (but excluding the position dependence of the MO coefficients). That is,

$$\mathbf{F}^{(x)}[\tilde{\mathbf{P}}] = \mathbf{F}^{(x)}[\mathbf{P}] + \mathbf{J}[\tilde{\mathbf{P}}^{(x)}] + c_K \mathbf{K}[\tilde{\mathbf{P}}^{(x)}] + \left(\frac{\partial(v_{xc})}{\partial\rho}\frac{\partial\rho}{\partial\tilde{\mathbf{P}}}\right)\tilde{\mathbf{P}}^{(x)} \tag{4.53}$$

$$\tilde{\mathbf{P}}^{(x)} = 3\mathbf{P}\mathbf{S}^x\mathbf{P} - 2\mathbf{P}\mathbf{S}^x\mathbf{P}\mathbf{S}\mathbf{P} - 2\mathbf{P}\mathbf{S}\mathbf{P}\mathbf{S}^x\mathbf{P} \tag{4.54}$$

$$= -\mathbf{P}\mathbf{S}^x\mathbf{P} \tag{4.55}$$

The DFT contribution is again an instance of the implicit first derivative of the XC matrix used in equation (4.28) $\mathbf{F}^{(x)}[\mathbf{P}]$ is just the standard partial derivative of the Fock matrix with respect to the nuclear position, and c_K represents the coefficient of exact exchange in the density functional.

The procedure to obtain $M^x = dM^{\text{var}}/dx$ then is to determine $\mathbf{b}_I$ and $\mathbf{b}_J$ using the $M_b[\hat{O}]$ formula, and perform two GMRES calculations (using the appropriate $\mathbf{A}$ matrix for state I or J) to determine the Lagrange multipliers $\mathbf{t}_I$ and $\mathbf{t}_J$. Then, take $M_x[\hat{O}]$ from equation (4.49) and $L_x(\mathbf{c}, \mathbf{t})$ from equation (4.51) and substitute into equation (4.48) to obtain the final gradient.

It bears reiterating that $L_x \neq dL/dx$, since it omits the $\partial L/\partial\mathbf{c}$ terms which only cancel when the full quantity dM/dx is being evaluated. This is why $\mathbf{t}$ must be redetermined for each operator and for each state.

4.2.3 Promolecule contribution

Having established the general form for evaluating the gradient of a matrix element of a one-electron operator between two distinct states, we now step back and note a particular

issue with the formulation of CDFT-CI which makes the actual computation of $\mathcal{E}^x$ (equation (4.4)) more complicated. Recall that the CDFT equations involve minimizing the value of $E[\rho] + \sum_k V_k \left(\int \hat{w}_k \rho d^3r - N_k \right)$ with respect to ρ and maximizing with respect to V_k for fixed $\hat{w}_k$ and N_k (section 1.2.4). However, when adapting CDFT for use in CDFT-CI, the concept of a "promolecule density" was introduced which produced modified values of N_k for a given system (section 2.2.3).[9] This used the converged density from independent calculations on noninteracting fragments, in conjunction with $\hat{w}_k$, to produce new values of N_k. However, the converged density of even non-interacting fragments has a dependence on nuclear position (when there is more than one atom in a fragment), which in turn introduces a position dependence on the constraint values N_k used as input for the CDFT calculations to obtain $|\Phi_I\rangle$ and $|\Phi_J\rangle$. This dependence will in turn trickle through to affect the other properties of the system, such as the free energy of state I

$$F_I = E_I + \sum_k V_{I,k} N_{I,k} \tag{4.56}$$

in the presence of constraints. A regular energy gradient will yield E_I^x, and when the $N_{I,k}$ are fixed (i.e., the promolecule correction is not in use), that will be equal to F_I^x, since the constraint parameters would not change to first order. When the promolecule correction is being applied, though, there is a nonzero N^x contribution (but V is still at a stationary point), so the free energy gradient becomes

$$F_I^x = E_I^x + \sum_k V_{I,k} N_{I,k}^x \tag{4.57}$$

Thus, in equation (4.6) when we said that H_{II}^x is the gradient of the CDFT state free energy, it is the gradient of that energy provided that the constraint values are also changing according to the promolecule formalism, i.e., it is the gradient including this correction of equation (4.57). The form of this small correction is quite simple once the N_k^x are known. The determination of the N_k^x, however, is not quite so simple.

Recalling the definition of N_k:

$$N_k = \langle \Phi_I | \hat{w}_k | \Phi_I \rangle \tag{4.58}$$

we note that the expression is precisely the matrix element of a one-electron operator, so we can reuse the algebraic machinery developed for $M[\hat{O}]$ wholesale to obtain N_k^x, with $\mathbf{b}_k = 2M_b[\hat{w}_k]$ yielding Lagrange multipliers $\mathbf{t}_k$ via GMRES, and $N_k^x = M_x[\hat{w}_k] - L_x(\mathbf{c}, \mathbf{t}_k)$. That N_k is a matrix element between two identical states serves only to simplify the algebra in that only one set of Lagrange multipliers is needed and $\partial N_k / \partial \mathbf{c} = 2M_b[\hat{w_k}]$ due to the symmetry. However, evaluation of $\partial M[\hat{w}_k]/\partial \mathbf{c}$ is slightly complicated by the need to retain isolation between the independent fragments. The gradient of the weight operator (for the full system, not the isolated fragments) is needed to evaluate $M_x[\hat{w}_k]$; for the Becke weights used in this implementation of CDFT-CI, such gradient terms have been computed in reference 72.

4.2.4 Constraint potential contribution

We pause once more to note that, while the single-state values of N_k^x were sufficient to correct W_{II}^x as in equation (4.57), the off-diagonal elements W_{IJ}^x include the matrix element of the constraint potential $\hat{V}_k$ between two different states, and those quantities cannot be related to single-state values of N_k^x.

Given that the overall gradient of the constraint potential is

$$\hat{V}_k^x = V_k \hat{w}_k^x + V_k^x \hat{w}_k \tag{4.59}$$

and the gradients of the weight matrices, $\mathbf{w}_k^x$, were used in evaluating $M_x[\hat{w}_k]$, it remains only to determine the gradient of the constraint potential Lagrange multipliers, V_k^x. These are intrinsically linked to the constraint value gradients N_k^x, changing in lockstep while maintaining the SCF solution of the inner SCF procedure. As such, we require a mechanism to go from the promolecule-derived constraint value gradients N_k^x to a gradient of the con-

straint potential Lagrange multipliers, V_k^x. The nested double-SCF structure of the method of solution of the CDFT equations presents an interesting challenge to the theoretician! To briefly resummarize the double-SCF procedure, a given set of orbitals allows a core Fock matrix to be constructed as for a normal SCF electronic calculation, and this Fock matrix is used as a fixed base upon which constraint potentials are added; the total Fock matrix is diagonalized to yield new orbitals and updated values for the constraint potential Lagrange multipliers in an iterative loop. Once Lagrange multipliers are determined which cause the constraints to be satisfied, the inner SCF loop terminates and the current orbitals are used to construct a new, updated core Fock matrix. This new core Fock matrix is used as a base for a new inner SCF procedure to determine the constraint Lagrange multipliers. It becomes clear that over the course of the entire double SCF calculation, the Lagrange multipliers V_k depend on nuclear coordinates both "directly", through the explicit dependence of the core Fock matrix and AO overlap on the nuclear coordinates, and also indirectly, through the dependence of the core Fock matrix on the MO coefficients (which in turn depends on the nuclear coordinates). It proves convenient to separate these dependencies, pushing the implicit contribution back into the **b** vector of equation (4.19) for the quantity of interest, and only treating the explicit dependence at this junction.

Explicit contribution

Treating just the explicit contribution requires holding the MO coefficients used to build the core Fock matrix fixed (while varying the nuclear geometry to $x + \delta x$), essentially just limiting the calculation to a single cycle of the outer SCF loop. The resulting change in the constraint potential Lagrange multipliers δV_k are the quantities we need for the explicit contribution. Because we consider only explicit changes in the core Fock matrix (ignoring its nonlinear dependence on the MO coefficients), the resulting changes to the orbitals can be determined solely via perturbation theory. We start off with knowledge that the constraints must be satisfied initially

$$N_k = \langle \Phi | \hat{w}_k | \Phi \rangle \tag{4.60}$$

A change in the constraint potential will manifest as a change in the Fock matrix, δF, which will induce a change in the wavefunction, $|\delta\Phi\rangle$. This change in the wavefunction will then contribute to the change in N_k, δN_k. Therefore,

$$\delta N_k = \langle\Phi|\delta\hat{w}_k|\Phi\rangle + 2\langle\Phi|\hat{w}_k|\delta\Phi\rangle \tag{4.61}$$

$$= \langle\Phi|\delta\hat{w}_k|\Phi\rangle + 2\sum_i \langle\phi_i|\hat{w}_k|\delta\phi_i\rangle \tag{4.62}$$

$$= \langle\Phi|\delta\hat{w}_k|\Phi\rangle + 2\sum_{ip} \langle\phi_i|\hat{w}_k|\phi_p\rangle\langle\phi_p|\delta\phi_i\rangle \tag{4.63}$$

The orbital overlaps $\langle\phi_p|\delta\phi_i\rangle$ are computed primarily from perturbation theory, taking care to include corrections from the nonorthogonal basis;

$$\langle\phi_p|\delta\phi_i\rangle = \frac{\langle\phi_p|\delta\hat{F} - \epsilon_i\delta\hat{S}|\phi_i\rangle}{\epsilon_i - \epsilon_p} \tag{4.64}$$

$$= \frac{\langle\phi_p|\delta\hat{F}_{\text{core}} - \epsilon_i\delta\hat{S} + \sum_l \delta\hat{V}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \tag{4.65}$$

$$= \frac{\langle\phi_p|\delta\hat{F}_{\text{core}} - \epsilon_i\delta\hat{S} + \sum_l (\delta V_l)\hat{w}_l + V_l\delta\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \tag{4.66}$$

This expression only holds for $i \neq p$; the contribution from $p = i$ can be determined from the normalization constraint on the orbitals;

$$1 = \langle\phi_i|\hat{S}|\phi_i\rangle \tag{4.67}$$

$$0 = 2\langle\delta\phi_i|\hat{S}|\phi_i\rangle + \langle\phi_i|\delta\hat{S}|\phi_i\rangle \tag{4.68}$$

$$\langle\phi_i|\delta\phi_i\rangle = -\frac{1}{2}\langle\phi_i|\delta\hat{S}|\phi_i\rangle \tag{4.69}$$

There is no $\delta\epsilon_i$ contribution in equations (4.64) through (4.69) because only variations of the orbitals which preserve normalization are produced.

Substituting these expressions into equation (4.63) yields:

$$\begin{aligned}\delta N_k = \langle\Phi|\delta\hat{w}_k|\Phi\rangle \\ &+ 2\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\delta\hat{F}_{\text{core}} - \epsilon_i\delta\hat{S} + \sum_l(\delta V_l)\hat{w}_l + V_l\delta\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \\ &+ 2\sum_i\langle\phi_i|\hat{w}_k|\phi_i\rangle\langle\phi_i|\delta\phi_i\rangle \end{aligned} \tag{4.70}$$

$$\begin{aligned}\delta N_k - \langle\Phi|\delta\hat{w}_k|\Phi\rangle = 2\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\sum_l(\delta V_l)\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \\ &+ 2\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\delta\hat{F}_{\text{core}} - \epsilon_i\delta\hat{S} + \sum_l V_l\delta\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \\ &- \sum_i\langle\phi_i|\hat{w}_k|\phi_i\rangle\langle\phi_i|\delta\hat{S}|\phi_i\rangle \end{aligned} \tag{4.71}$$

$$\begin{aligned}\delta N_k - \langle\Phi|\delta\hat{w}_k|\Phi\rangle + \sum_i\langle\phi_i|\hat{w}_k|\phi_i\rangle\langle\phi_i|\delta\hat{S}|\phi_i\rangle \\ -2\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\delta\hat{F}_{\text{core}} - \epsilon_i\delta\hat{S} + \sum_l V_l\delta\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} = 2\sum_l\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p}(\delta V_l)\end{aligned} \tag{4.72}$$

$$b_k = A_k^l(\delta V_l) \tag{4.73}$$

which is a system of linear equations corresponding to the various constraints being applied. Here, we have $\delta N_k = dN_k/dx$ (obtained per section 4.2.3, $\delta\hat{w}_l = \hat{w}_l^x$, $\delta\hat{S} = \hat{S}^x$, $\delta\hat{F}_{\text{core}} = \hat{F}_{\text{core}}^{(x)}$, and $\delta V_l = \partial V_l/\partial x$, the latter of which are the desired quantities. Since the promolecule specification requires both charge and spin constraints on a given fragment, there will in general be at least two constraints, and thus a linear system to be solved. Fortunately, the **A** matrix does not have any position dependence and can be precomputed and inverted just once, with b (and thus δV_l) being computed for a single nuclear coordinate at a time. The b vector contains only quantities which are already known; the gradient of the weight matrix is easy to compute, and the gradient of the overlap and core Fock matrices are already present

as they were needed for equation (4.51).

Implicit dependence

Having constructed an expression for the explicit dependence of the constraint potential on the nuclear position, in the form of an expression $(\delta V)\hat{w} + V(\delta\hat{w})$ which is added to the gradient of the Fock matrix, it remains to treat the implicit dependence, through the MO coefficients used to build the core Fock matrix. As previously indicated, this will be included through the Lagrange multipliers in $L(\mathbf{c}, \mathbf{t})$, by including an extra contribution from $\partial V/\partial \mathbf{c}$ to the $\mathbf{b}$ vector in equation (4.19). Obtaining this $\partial V/\partial \mathbf{c}$ contribution actually requires a very similar perturbation-theory structure to that of the $\partial V/\partial x$ contribution above, including the need for a linear system in the various constraints. In this case the perturbation is now $\delta\hat{F} = \partial\hat{F}_{\text{core}}/\partial\mathbf{c}$ and $\delta V_l = \partial V_l/\partial\mathbf{c}$. The form of the equations is identical to equation (4.72) with simplifications that δN_k is zero (the target constraint value does not depend on the MO coefficients passed to the core Fock matrix), and $\delta\hat{S}$ and $\delta\hat{w}$ are also zero. The orbital overlap then becomes just

$$\langle\phi_p|\delta\phi_i\rangle = \frac{\langle\phi_p|\delta F_{\text{core}} + \sum_l(\delta V_l)\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p} \tag{4.74}$$

And the linear system to be solved:

$$-\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\delta\hat{F}_{\text{core}}|\phi_i\rangle}{\epsilon_i - \epsilon_p} = \sum_l\sum_{i\neq p}\langle\phi_i|\hat{w}_k|\phi_p\rangle\frac{\langle\phi_p|\hat{w}_l|\phi_i\rangle}{\epsilon_i - \epsilon_p}(\delta V_l) \tag{4.75}$$

$$b_k = A_k^l(\delta V_l) \tag{4.76}$$

An astute reader will note that the $\mathbf{A}$ matrix is identical to the one in equation (4.72). All of the b vectors may be generated at once as contractions against $\partial F_{\text{core}}/\partial\mathbf{c}$, repeated for the number of constraints applied to the system. Such contractions against $\partial F_{\text{core}}/\partial\mathbf{c}$ were described in equations (4.20) through (4.28).

4.2.5 Final assembly

At this point, all the pieces are in place to compute $(\mathbf{H} - \mathcal{E}\mathbf{S})^x_{IJ} = W^x_{IJ}$, the last remaining piece before equation (4.4) may be applied to obtain $\mathcal{E}^x$. In the now-familiar procedure, we construct

$$W^{\text{var}}_{IJ} = \frac{W_{IJ}}{\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}} - L(\mathbf{c}_I, \mathbf{t}_I) - L(\mathbf{c}_J, \mathbf{t}_J) \tag{4.77}$$

and solve for the Lagrange multipliers $\mathbf{t}_I$ and $\mathbf{t}_J$ which make W^{var}_{IJ} variational with respect to the MO coefficients. To do so, we need vectors $\mathbf{b}_I$ and $\mathbf{b}_J$ as input for GMRES calculations to determine $\mathbf{t}_I$ and $\mathbf{t}_J$; in our formalism the contribution from $\mathbf{H}$ is split out into terms arising from the constrained states, so this really looks like

$$\frac{\partial\left(\mathbf{H} - \mathcal{E}\mathbf{S}\right)_{IJ}}{\partial \mathbf{c}_I} \tag{4.78}$$

$$= \frac{\partial}{\partial \mathbf{c}_I} \frac{\left(\frac{F_I+F_J}{2} S_{IJ} - \left\langle \Phi_I \middle| \frac{\hat{V}_I+\hat{V}_J}{2} \middle| \Phi_J \right\rangle - \mathcal{E} S_{IJ}\right)}{\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}} \tag{4.79}$$

$$= -\frac{\partial}{\partial \mathbf{c}_I} \frac{\left\langle \Phi_I \middle| \frac{\hat{V}_I+\hat{V}_J}{2} \middle| \Phi_J \right\rangle}{\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}} + \left(\frac{F_I+F_J}{2} - \mathcal{E}\right) \frac{\partial}{\partial \mathbf{c}_I} \frac{S_{IJ}}{\sqrt{\langle\Phi_I|\Phi_I\rangle\langle\Phi_J|\Phi_J\rangle}} \tag{4.80}$$

In the notation we have developed, we can then write

$$\mathbf{b}_I = -M_b\left[\left(\hat{V}_I + \hat{V}_J\right)/2\right] + \left(\frac{F_I+F_J}{2} - \mathcal{E}\right) M_b[\hat{S}] + \frac{1}{2}\sum_l M[\hat{w}_{I,l}] \tag{4.81}$$

where $\partial V_{I,l}/\partial \mathbf{c}_I$ are determined from equation (4.76); $\mathbf{b}_J$ is determined similarly. A pair of GMRES calculations then yields $\mathbf{t}_I$ and $\mathbf{t}_J$, which we can then assemble into the final

$$\frac{dW_{IJ}}{dx} = -M_x\left[\left(\hat{V}_I + \hat{V}_J\right)/2\right] + \left(\frac{F_I+F_J}{2} - \mathcal{E}\right) M_x[\hat{S}] - L_x(\mathbf{c}_I, \mathbf{t}_I) - L_x(\mathbf{c}_J, \mathbf{t}_J) \tag{4.82}$$

noting that the $\mathbf{F}^{(x)}$ matrices which are used in computing the L_x must include the contributions from $\hat{V}^{(x)} = \frac{\partial V}{\partial x}\hat{w} + V\hat{w}^x$ (where $\partial V/\partial x$ come from equation (4.72)).

This completes the construction of W_{IJ}^x and returns the gradient calculation of $\mathcal{E}^x$ to equation (4.4), recalling that W_{II}^x comes from equation (4.57). We note that the CI vector C used in equation (4.4) is an eigenvector of the generalized eigenvector problem, in contrast to the coefficient vector produced by many CDFT-CI calculations, which is in the orthogonalized diabatic basis; a factor of $\mathbf{S}^{-1/2}$ allows interconversion.

4.3 Results

We have implemented CDFT-CI gradients in a development version of Q-Chem; since we do not have the benefit of the Hellmann-Feynman theorem for gradients, we expect that the CPU time for a CDFT-CI gradient evaluation should be comparable to that needed for a Hessian evaluation using regular DFT. As some indication of the qualitative similarity, we note that for the $OH + C_2H_6 \leftrightarrow H_2O + C_2H_5$ system presented below (with rather large basis and integration grid), some timings are presented in Table 4.1. The gradient evaluation

	HF	BLYP	B3LYP
CDFT-CI gradient	3053	3582	4570
DFT Hessian	112	1977	2012

Table 4.1: Total execution time, in seconds of CPU time. The non-HF results suffer from the exceptionally large integration grid used for this calculation; this penalty even affects the CDFT-CI HF gradient evaluation due to an implementation detail that CDFT is implemented as an exchange functional, forcing the DFT routines to be executed even if they are not needed.

is within roughly a factor of two of a DFT hessian evaluation, which is reasonable for our comparatively unoptimized code. We have endeavored to retain the $O(N^3)$ scaling behavior of DFT, albeit with a rather large pre-factor. (Each GMRES iteration requires some number of $O(N^3)$ matrix manipulations, and it is not atypical for 20 GMRES iterations to be required for convergence.) In the present implementation, $\mathbf{F}^{(x)}$ and $\mathbf{S}^x$ and $\mathbf{W}^x$ are constructed explicitly, which is the rate-limiting step. For the B3LYP case listed in Table 4.1, building $\mathbf{W}^x$ took 408 CPU-seconds, and building $\mathbf{F}^{(x)}[\mathbf{P}]$ took 2130 CPU-seconds, with the extra

steps for the transformation to $\mathbf{F}^{(x)}[\tilde{\mathbf{P}}]$ requiring another 620 CPU-seconds. However, these matrix gradients are only used in expressions which involve tracing over the matrix indices, so it should be possible as a future improvement to eliminate their explicit construction in favor of implicit routines which contract over those matrix indices. A reduction from the current 70% of execution time would be a very useful development.

4.3.1 Transition State Optimization

The present work also found it opportune to return to the set of reactions previously used to evaluate CDFT-CI,[74] taken from the HTBH38/04 and NHTBH38 databases of Zhao et al.[131,233] The newly implemented analytic gradients allow us to locate optimized transition-state geometries at the CDFT-CI level of theory, to compare against the reference geometries which were optimized at a QCISD/MG3 level of theory.[131] Since the CI vector should be strongly spread over both configurations at the transition-state, these geometry optimizations represent a stringent test of the CDFT-CI coupling gradient computation — any inaccuracies would be highlighted by the delocalized CI vector, given the availability of accurate reference data for comparison. Furthermore, the change between the CDFT-CI energy calculated at the reference geometry and at the CDFT-CI optimized geometry presents a measure of the "goodness" of the CDFT-CI geometry; systems with small energy change are expected to have a converged CDFT-CI geometry close to the reference geometry. It also presents an opportunity to once again examine the overall quality of CDFT-CI for barrier heights. The present calculations are performed using a 6-311++G** basis set, as opposed to the 6-311+G(3df,2p) basis set used in reference 74. For performing gradient evaluations, we also found it necessary to increase the quality of the DFT integration grid (though this is less critical when a smaller basis set is used) — SG-1 is insufficient to yield sufficiently converged SCF calculations. The present work uses a Lebedev grid with 100 radial points and 302 angular points (as implemented and truncated by Q-Chem). We also deem it sufficient to present results using a single functional, given the overall robust performance with multiple functionals in the previous work.[74] Table 4.2 shows the results for the 32 reactions considered,

with forward (and backward, when distinct) reaction barrier heights for CDFT-CI at the reference geometry, CDFT-CI at the optimized geometry, and the reference barrier heights.

Reaction	reference	CDFT-CI at ref. geom.	at optimized geom.
$H + HCl \leftrightarrow Cl + H_2$	5.7	4.59	6.18
	8.7	10.98	12.57
$OH + H_2 \leftrightarrow H + H_2O$	5.1	7.86	6.36
	21.2	19.66	18.16
$CH_3 + H_2 \leftrightarrow H + CH_4$	12.1	12.60	12.78
	15.3	13.51	13.69
$OH + CH_4 \leftrightarrow CH_3 + H_2O$	6.7	14.54	NA
	19.6	25.43	NA
$H + H_2 \leftrightarrow H_2 + H$	9.6	7.64	7.64
$OH + NH_3 \leftrightarrow H_2O + NH_2$	3.2	2.80	3
	12.7	11.47	12
$HCl + CH_3 \leftrightarrow Cl + CH_4$	1.7	6.27	5.82
	7.9	13.57	13.12
$OH + C_2H_6 \leftrightarrow H_2O + C_2H_5$	3.4	6.83	6.40
	19.9	22.28	21.84
$F + H_2 \leftrightarrow H + HF$	1.8	1.73	1.42
	33.4	29.64	29.33
$O + CH_4 \leftrightarrow OH + CH_3$	13.7	15.43	15.67
	8.1	11.52	11.76
$H + PH_3 \leftrightarrow PH_2 + H_2$	3.1	3.22	3.36
	23.2	27.62	27.76
$H + HO \leftrightarrow H_2 + O$	10.7	9.24	9.27
	13.1	12.24	12.27
$H + H_2S \leftrightarrow H_2 + HS$	3.5	5.20	5.04
	17.3	22.13	21.98

$O + HCl \leftrightarrow OH + Cl$	9.8	5.10	5.92
	10.4	8.49	9.30
$NH_2 + CH_3 \leftrightarrow CH_4 + NH$	8.0	9.70	9.66
	22.4	21.92	21.88
$NH_2 + C_2H_5 \leftrightarrow C_2H_6 + NH$	7.5	11.68	11.21
	18.3	19.34	18.87
$C_2H_6 + NH_2 \leftrightarrow NH_3 + C_2H_5$	10.4	12.98	12.88
	17.4	19.75	19.66
$NH_2 + CH_4 \leftrightarrow CH_3 + NH_3$	14.5	15.54	15.37
	17.8	17.76	17.58
$H + N_2O \leftrightarrow OH + N_2$	18.14	15.63	16.22
	83.22	79.40	79.99
$H + FH \leftrightarrow FH + H$	42.18	37.29	37.52
$H + ClH \leftrightarrow HCl + H$	18.0	22.40	22.18
$H + FCH_3 \leftrightarrow HF + CH_3$	30.38	26.04	25.94
	57.02	53.61	53.52
$H + F_2 \leftrightarrow HF + F$	2.27	0.97	2.62
	106.18	107.00	108.64
$CH_3 + FCl \leftrightarrow CH_3F + Cl$	7.43	2.86	1.32
	60.17	60.42	58.87
$F^- + CH_3F \leftrightarrow FCH_3 + F^-$	-0.34	0.55	0.25
$F^- \cdots CH_3F \leftrightarrow FCH_3 \cdots F^-$	13.38	14.76	14.61
$Cl^- + CH_3Cl \leftrightarrow ClCH_3 + Cl^-$	3.10	1.56	1.43
$Cl^- \cdots CH_3Cl \leftrightarrow ClCH_3 \cdots Cl^-$	13.61	11.28	11.49
$F^- + CH_3Cl \leftrightarrow FCH_3 + Cl^-$	-12.54	-13.28	-13.45
	20.11	20.94	20.77
$F^- \cdots CH_3Cl \leftrightarrow FCH_3 \cdots Cl^-$	2.89	2.53	2.56

	29.62	29.75	29.75
$OH^- + CH_3F \leftrightarrow HOCH_3 + F^-$	-2.78	-3.23	-3.24
	17.33	18.65	18.64
$OH^- \cdots CH_3F \leftrightarrow HOCH_3 \cdots F^-$	10.96	11.26	11.32
	47.20	46.91	47.15

Table 4.2: Reaction barrier heights from various methods; forward and backward reaction barrier heights are shown when distinct. Reference data from references 131 and 233; the CDFT-CI energy at the reference geometry, and the CDFT-CI energy at the optimized CDFT-CI geometry are shown. All energies in kcal/mol. CDFT-CI data in the 6-311++G** basis set.

Given that CDFT-CI with the reactant/product constrained states as its basis degenerates into ordinary DFT calculations on the reactant or product fragments at infinite separation, in some cases there is conflict between accurate forward and backward barrier heights, when DFT does not treat the reactant and product states equally well. Over the entire set of reactions (modulo those with no CDFT-CI transition state yet converged), CDFT-CI at the reference geometry has a mean error of 0.17 kcal/mol, and a mean absolute error of 2.05 kcal/mol; using energies from the optimized geometries performs essentially the same, with mean error of 0.17 kcal/mol and mean absolute error 2.01 kcal/mol. Systems with oxygen atoms or hydroxy radicals seem to be more likely to have worse energies after geometry optimization, suggesting that B3LYP may cause the oxygen atom to be unphysically electronegative. Breaking the reactions down by class, average results are shown in Table 4.3. There appear to be no substantial differences between the reference geometry results and the optimized geometry results at a per-category level, with the optimized geometries consistently performing slightly better. Overall, CDFT-CI seems to produce very good reaction transition states, being essentially statistically indistinguishable from the reference geometries.

	ME, initial	ME, optimized	MAE, initial	MAE, optimized
Hydrogen transfer	0.89	0.87	2.18	2.08
Heavy atom transfer	-1.94	-1.82	3.03	3.22
Nucleophilic substitution	-0.07	-0.10	0.88	0.82
All reactions	0.17	0.17	2.05	2.01

Table 4.3: Deviation of CDFT-CI barrier heights from the reference values. Mean errors and mean absolute errors are given, broken down by the type of reaction, at the initial (reference) geometry, and at the final optimized geometry.

4.3.2 Excited-state optimizations

Gradients on the ground-state allow for geometry optimization of critical points, both minima and saddle points (transition states). It is less common to have gradients of the excited state energy available, which enable optimization of minima on excited-state electronic PESs. Our CDFT-CI gradient implementation produces forces on both the ground and excited states, treating them on an equal footing. As a simple example, we optimize the first singlet excited state of H_2. This molecule has been exhaustively studied, and it is known that the lowest $^1\Sigma_g^+$ state has two minima, with the lower-energy minimum at a separation of 1.0 Å and a second minimum at 2.3 Å.[234] The two minima correspond, qualitatively, to a $1s \rightarrow 2s$ valence excitation and an linear combination of ionic states, respectively. Since CDFT is unable to describe valence excitations, it would be surprising if CDFT-CI could reproduce this double-minimum in the excited state. In fact, with the cc-pVDZ basis set and the B3LYP functional, using the standard four-state CDFT-CI active space for diatomics (H^+H^-, H^-H^+, $H^\uparrow H^\downarrow$, and $H^\downarrow H^\uparrow$), we find only a single minimum at a separation of 2.235 Å (from an optimization starting at the ground-state equilibrium geometry). Throughout the optimization, the state in question is dominated by contributions from the ionic configurations, as expected given the unavailability of valence-excitation states. Nonetheless, the ionic minimum seems to be treated correctly, given that our modest basis set is not intended to yield quantitatively accurate results. It remains telling, though, that only one minimum is found — a reminder that we must always be conscious of the composition of the active space. If the active space does not include the proper states to describe a portion of configuration space, then the

CDFT-CI energy will be unreliable; this active space dependence must be kept in mind when applying CDFT-CI to new systems.

H_2 is well studied in part because it is a very small test system, and it functions as a first test for new theoretical methods. However, the double-minimum in the excited state makes it less useful for assessing the validity of CDFT-CI given that we know it will fail for the valence excitations which comprise half of the double well. It is therefore useful to consider another simple molecule with well-known structure, but which has only a single minimum in the excited state. The diatomic Li_2 meets these criteria; it also has more than two electrons, presenting a somewhat more stringent test on the applicability of theoretical methods. A similar four-state geometry optimization on the first singlet excited state of Li_2 (also using cc-pVDZ and B3LYP) locates a geometry minimum (ungerade) at a separation of 3.169 Å. Furche and coworkers[184,185] have compiled a benchmark suite of reference adiabatic excitation energies (including relaxation in the excited state), including excited-state geometries for for more than twenty molecules.[184] They give the dilithium $^1\Sigma_u^+$ minimum to be at 3.11 Å (experiment[235]), with all the tabulated TD-DFT methods underpredicting the minimum except for TDHF. Given the proper active space, CDFT-CI successfully optimizes geometries for HOMO $\rightarrow$ LUMO excitations in these simple systems.

However, diatomic molecules do not make a particularly compelling case for the utility of analytic gradients, since with a single degree of freedom the geometry may be easily optimized even with just single-point calculations. Moving away from diatomics, then, we consider the ethylene cation. This system presents theoretical interest even in the ground state, in particular with the non-planar nature of the equilibrium state.[236–238] Consensus has been reached that the dihedral angle is around 25°,[239–242] but it is difficult to confidently state a more precise value. The doublet nature of this molecule allows for only a four-state CDFT-CI active space to be used once again, splitting the molecule in half through the carbon-carbon bond, so that the two fragments are both CH_2 units. Between those fragments, there are two possibilities for each of localizing the charge and the spin on the neutral (radical) fragment; the product space gives four possible constraints: $CH_2^+CH_2^\uparrow$, $CH_2^+CH_2^\downarrow$, $CH_2^\uparrow CH_2^+$,

and $CH_2^{\downarrow}CH_2^{+}$. A CDFT-CI ground-state optimization with that four-CDFT-state basis and again cc-pVDZ/B3LYP finds the dihedral angle at convergence to be 35.55°, even larger than the stock DFT state at 28.33°. However, this angle is known to be sensitive to the quality of the basis set,[242] so we do not necessarily seek quantitative accuracy. Of more interest to us at present is the behavior in the excited state; the minimum on the excited state is known to be at a conical intersection with the ground state, at a perpendicular geometry.[237] A CDFT-CI geometry optimization in the first excited state (starting from the equilibrium ground state structure) proceeds to a perpendicular geometry which is degenerate with the ground state. Unfortunately, Q-Chem's geometry optimizer does not treat conical intersections, so there is little more that may be said about this system at present.

The general pattern of a four-state CDFT-CI with simple charge/spin-constrained states has been successful in these previous applications, so it is natural to apply it to another polyatomic molecule (which does not have its excited-state minimum at a conical intersection). Ethane (C_2H_6) can be partitioned similarly to a diatomic (into two methyl groups) but has additional nuclear degrees of freedom, giving a more clear advantage to analytical gradients for geometry optimization. The four CDFT-CI states are now $CH_3^{+}CH_3^{-}$, $CH_3^{-}CH_3^{+}$, $CH_3^{\uparrow}CH_3^{\downarrow}$, and $CH_3^{\downarrow}CH_3^{\uparrow}$. At the ground-state equilibrium geometry, the $C-C$ distance is 1.54 Å with B3LYP/cc-pVDZ; the excited-state minimum for the ionic-like configuration has the carbons some 3.3 Å apart from each other. This is similar to the diatomics previously studied, a little more than twice the ground-state separation, indicating a commonality amongst the ionic-like minima. Here, the methyl groups have both become essentially planar and are parallel to each other, though they retain the staggered rotational conformation. The substantial geometry change is consistent with the mostly-continuous optical spectrum of ethane, given the minimal overlap with the ground state.[243] Again, CDFT-CI successfully locates the excited-state geometry of HOMO $\rightarrow$ LUMO excitations, given a sufficient active space.

4.4 Conclusions

We have derived and implemented the equations necessary to obtain analytical gradients of the CDFT-CI energy. The resulting implementation has been used to validate previously investigated reaction barrier heights at self-consistently optimized transition-state geometries, which have good accuracy as compared against reference values computed by high-level theory. Gradients are available equally for the ground state and electronic excited states, allowing for optimization of excited state geometries. As a density-functional method, CDFT-CI has potential application to large systems, with gradients allowing for excited-state dynamics on organic photoelectronic systems at the donor/acceptor interface, even with QM/MM embedding, provided an appropriate active space is available. There is also great potential in CDFT-CI as an economical method for tracking the decay of optically excited systems, including decay to conical intersections. However, that potential comes with a caveat, namely that the user must choose the active space for the calculation. Finding active spaces which remain valid over the entire area of the PES in question may prove to be challenging.

The availability of diabatic couplings and coupling gradients make possible another investigation of key interest to chemists: studying the Condon approximation that the electronic coupling is relatively invariant to changes in nuclear position. Now that we have implemented the gradient of the coupling element between states, we can proceed to throw it away (set $S_{12}^x = H_{12}^x = 0$ so that $\mathcal{E}^x \approx C_1^2 H_{11}^x + C_2^2 H_{22}^x$) and see how the omission changes the resulting nuclear dynamics. If the changes are small, then the Condon approximation can be safely applied for substantial computational speedup. Having the coupling derivative available allows the validity of the Condon approximation to be assessed on a system-by-system basis, giving greater confidence in the ensuing results.

It remains something of an open question what constitutes a "good" or "sufficient" active space for CDFT-CI calculations. Diatomics of low bond order seem well-understood, and the reactant/product split for the set of reaction transition states examined in this work produced good results, but no study has been made of whether increasing the active space would produce further improvement in transition states or elsewhere. Perhaps including

configurations with charge-transfer character would shift the location of reaction transition states; the ability to optimize transition-state geometries allows any such effects to be studied, and the results used to give guidance for the selection of active spaces in general.

Additionally, electronic excited states remain ever-tantalizing; to further assess CDFT-CI's usability in this space, it would be fruitful to study simple photoisomerization systems. With only one bond changing, the difficulty of selecting a CDFT-CI active space is reduced, making isomerization studies feasible. Such studies would help bring insight into how to choose CDFT-CI active spaces for effective description of electronic excited states, helping to bring the DFT toolbox into scope for studying the photochemistry of more generic large molecules.

Chapter 5

Conclusion

In this thesis, we have forayed deep into the depths of a particular flavor of electronic structure theory. To start, there was a quick pass through the basics of electronic structure, then greater detail about CDFT and the electronic couplings used to construct CDFT-CI therefrom. All of this, leading to a destination that intertwines its way away from the equilibrium ground state, amongst chemical reactions and photochemical transitions.

We have focused on Constrained Density Functional Theory–Configuration Interaction, which strives to bring the accuracy and efficiency of DFT methods to realms where that accuracy previously did not exist. In reaction transition states, near conical intersections, and on the excited state, we have used CDFT-CI to broaden the scope of DFT. This is done by introducing multiple configurations to the system, and a plausible approximate form for the couplings between them, ensuring a treatment of static correlation to complement DFT's inherent treatment of dynamic correlation through the exchange-correlation functional.

This static correlation is key at reaction transition states, combining attributes of the reactant and product configurations to yield a DFT energy that is accurate and unburdened by excessive self-interaction error. Improvement in reaction barrier heights of a factor of 2–3 was seen for a broad spread of functionals, with less correction needed for functionals which had better initial performance.

The static correlation brought in by CDFT-CI is also critical for correct description

of conical intersections, those diabolical degeneracies between electronic states. Here the correlation is not between reactant and product, but rather between ground and excited state. Even when intersections are constrained to occur by symmetry, standard TD-DFT methods for excited states fail to produce *conical* intersections; CDFT-CI brings in configurations from spin- and charge-localized states to yield the necessary static correlation and the correct topology of energetic degeneracy.

Static correlation ebbs and flows as reactions progress from reactants through transition state to products; to properly trace out the course of the reaction requires the forces on the nuclei, the gradient of the energy. At the transition state, this gradient must be zero, yet a single energy evaluation cannot attest to whether this condition is met. With the gradient of the CDFT-CI energy computationally available, the transition-state geometry can be located and confirmed to be a stationary point. At such saddle points, the CDFT-CI energy remains accurate, yielding reaction barrier heights which are still of excellent quality, being statistically indistinguishable from barrier heights evaluated using a reference geometry.

Forces from the CDFT-CI energy are not limited to the ground state; they are equally available on electronic excited states. Energy minima were located for diatomic and polyatomic molecules, but only for states dominated by ionic contributions. A second minimum of molecular hydrogen, corresponding to valence excitation, is completely absent; this highlights the need to select an active space appropriate for the task at hand. With an incomplete active space, incorrect results may be obtained.

Nonetheless, the real strength and weakness of CDFT-CI lies with us, the users. In a valence-bond-like sense, the CDFT-CI basis states that are *chosen* for the active space have a real physical interpretation and derive strength from that for interpreting the CDFT-CI description of the system. Yet before a calculation may be performed, the relevant basis states must be determined and verified, preventing CDFT-CI from being a "black box" electronic structure method. It thus requires proper intuition and insight, but with those both, CDFT-CI and dynamics upon it can take our molecules wherever in configuration space they need to go.

Appendix A

GMRES

The general class of iterative linear solvers endeavor to obtain an approximate solution to the linear system

$$\mathbf{A} \cdot x = b$$

without explicitly operating on the matrix $\mathbf{A}$, instead only evaluating the matrix-vector product $\mathbf{A} \cdot x_n$. In the case of GMRES, $\mathbf{A}$ must be square, though there exist iterative solvers which can act on non-square systems. The iterative solver gets its name by successively applying the matrix $\mathbf{A}$ to an initial vector and then iterating to generate a series of vectors which is used to produce an active space. After each iteration, the algorithm produces a trial solution vector for that cycle of the algorithm and determines whether this approximate solution x_n is a "good enough" solution to the system in question, that is, whether the residual $|\mathbf{A}x_n - b|$ is smaller than some convergence threshold. In the absence of other information about the system, the only vector available to use as a seed for the iterations is b, making the iterative series $S_b = \{b, \mathbf{A}b, \mathbf{A}^2b, \cdots \mathbf{A}^nb, \cdots\}$. The span of the truncated series at each step is called the nth Krylov space, $K_n = \text{span}\{b, \mathbf{A}b, \mathbf{A}^2b, \cdots \mathbf{A}^nb\}$ The approximate solution vector x_n is chosen as the best solution vector within the Krylov space K_n, as detailed below. However, in order to work with the Krylov space, a useful basis is needed, given that S_b will quickly become numerically unstable as the largest eigenvalue rapidly becomes dominant with successive iterations. Instead, Arnoldi iterations[231] are used

to construct an orthogonal basis Q_n for the Krylov space K_n; the nature of the construction of K_n allows these basis sets to be a nested collection $Q_n = \{q_0, q_1 \cdots q_n\}$ with $Q_n \subset Q_{n+1}$. Thus, at step n it is only necessary to determine q_n, and the rest of the set is reused from the previous iteration. Accordingly, q_{n-1} contains all the information gained from what would formally be written as $\mathbf{A}^{n-1}b$, so likewise $y_n = \mathbf{A}q_{n-1}$ contains all the new information for Q_n. Decomposing it into components $y_n = y_n^{\parallel} + y_n^{\perp}$ with $y_n^{\parallel} \in K_{n-1}$ and $y_n^{\perp} \notin K_{n-1}$, it is clear then that $q_n = y_n^{\perp}$, which may be evaluated using a Gram-Schmidt orthogonalization. ($q_0 = \frac{b}{|b|}$.) Having constructed a useful basis Q_n for the Krylov space K_n, the next step is find the approximate solution vector $x_n \in K_n = \text{span}\,[Q_n]$ which is the best approximate solution $\mathbf{A}x_n \approx b$ in the least-squares sense.

In order to do so, we use the projection of $\mathbf{A}$ into the basis Q_n. That is, we form a matrix $\mathbf{H}$ of dimension $(n+1) \times n$ where the (i,j)-th element is the overlap of $\mathbf{A}q_j$ with q_i; alternately, the jth column is the result of $\mathbf{A}q_j$ in the basis Q_n. In summary, this means that

$$\mathbf{A}\mathbf{Q}_n = \mathbf{Q}_{n+1}\mathbf{H} \tag{A.1}$$

and

$$\mathbf{Q}_{n+1}^{\dagger}\mathbf{A} = \mathbf{H}\mathbf{Q}_n^{\dagger} \tag{A.2}$$

which we will use later on. Not only is the product $\mathbf{A}q_j$ guaranteed to be in Q_n (by the sequential construction of Q_n and the extra row in $\mathbf{H}$), but we also have the property that $H_{ij} = 0$ when $i > j+1$. That is, $\mathbf{H}$ has Hessenberg structure (its lower-1 triangle is all zeros). Since $q_0 = \frac{b}{|b|}$, the target vector b is just $|b|e_0$ when projected into the basis Q_n. In this basis, then, our original equation $\mathbf{A}x = b$ becomes

$$\mathbf{H}y = |b|e_0 \tag{A.3}$$

This new equation has a structure which lends itself to direct solution, and is of much smaller dimensionality than the original equation, making it feasible to solve directly. Unfortunately, the linear system in equation (A.3) is overdetermined, and thus can only be solved in an

approximate sense; here it proves convenient to use the least-squares residual as the metric for solution quality (hence "minimal residual" in GMRES).

However, it is not immediately clear that the vector y which produces the minimum residual $\left|\mathbf{H}y - |b|e_0\right|$ corresponds to a solution which will minimize the residual of interest, namely $|\mathbf{A}x - b|$. Fortunately, each of the basis sets Q_n is by construction orthonormal, so Q_n (or $Q_n^{-1} = Q_n^\dagger$) acting on a vector in the appropriate space will not change its norm. As such, we note that

$$|\mathbf{A}x - b| = \left|\mathbf{Q}_{n+1}^\dagger \left(\mathbf{A}x - b\right)\right| \tag{A.4}$$

$$= \left|\mathbf{H}\mathbf{Q}_n^\dagger x - \mathbf{Q}_{n+1}^\dagger b\right| \tag{A.5}$$

$$= \left|\mathbf{H}y - |b|e_0\right| \tag{A.6}$$

so a minimum-residual solution to the linear equation in the small basis will have the same residual when transformed back to the full basis. (This is of course not a minimum residual solution in the full basis, since the system has an exact solution in the full basis; it is, however, the best residual from a solution vector in the current Krylov space.)

In fact, another (distinct) orthogonal matrix is needed to actually determine the vector y which instantiates the minimal residual vector; this matrix $\mathbf{Q}$ (again, distinct from any of the $\mathbf{Q}_n$) stems from the QR factorization of $\mathbf{H}$, partitioning as $\mathbf{H} = \mathbf{Q}\mathbf{R}$ with $\mathbf{Q}$ orthogonal and $\mathbf{R}$ upper-triangular.[231] It then becomes

$$|\mathbf{A}x - b| = \left|\mathbf{Q}\mathbf{R}y - |b|e_0\right| \tag{A.7}$$

$$= \left|\mathbf{R}y - |b|\mathbf{Q}^\dagger e_0\right| \tag{A.8}$$

At this point we note that although $\mathbf{R}$ is upper-triangular, it is still of dimension $(n+1) \times n$, and as such this system remains overdetermined. However, it is clear that since the last row of $\mathbf{R}$ is entirely zero, then the product $\mathbf{R}y$ will always have zero as its last element, and thus the norm of the whole expression will be bounded below by the last element of the vector $|b|\mathbf{Q}^\dagger e_0$. In the general case, the actual upper triangle of $\mathbf{R}$ is nonsingular, so the remaining

rows may be solved exactly, making the lower bound sharp.

Determination of the vector y producing the minimum residual then proceeds by normal solution of the triangular system

$$\mathbf{R}y = |b|\mathbf{Q}^{\dagger}e_0 \tag{A.9}$$

after projection to the $n \times n$ subspace.

However, the special structure of the GMRES problem means that we can effect a more efficient solution than using a standard QR partitioning routine and subsequent direct solution and substitution on the upper triangular system. This is because the original matrix $\mathbf{H}$ has Hessenberg structure — we can progress down the lower subdiagonal eliminating element by element with a series of 2×2 rotation blocks, bringing the factorization step down to $O(n^2)$ from the general-case $O(n^3)$. However, since we expect n to never be more than twenty or perhaps thirty in our application, and the dimensions of $\mathbf{A}$ scale as the square of the number of electrons in the system, this speedup probably does not contribute very much to the overall runtime of a CDFT-CI calculation.

In determining the factorization, the matrix $\mathbf{Q}$ need not be maintained explicitly and only $|b|\mathbf{Q}^{-1}e_0$ need be stored; this can be updated in-place if the factorization is effected via pairwise rotations as described above. The solution of the upper-triangular system for y is then straightforward; given this y_n, at cycle n of the GMRES algorithm, the approximate solution vector in the full basis is given by $x_n = \sum_{i=0}^{n}[y_n]_i q_i = \mathbf{Q}_n y_n$ through a standard change of basis.

Amusingly, GMRES is a special case of the Direct Inversion in the Iterative Subspace (DIIS) method which is almost universally used as a convergence accelerator for electronic structure SCF calculations.

Bibliography

[1] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover Publications, New York, 1996.

[2] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, 1964.

[3] R. G. Parr and W. Yang. *Density-Functional Theory of Atoms and Molecules*. Oxford University Press, New York, 1989.

[4] P. H. Dederichs, S. Blugel, R. Zeller, and H. Akai. Ground-states of constrained systems — application to cerium impurities. *Phys. Rev. Lett.*, 53:2512–2515, 1984.

[5] Q. Wu and T. Van Voorhis. Direct optimization method to study constrained systems within density-functional theory. *Phys. Rev. A*, 72:024502, 2005.

[6] Q. Wu and T. Van Voorhis. Constrained density functional theory and its application in long-range electron transfer. *J. Chem. Theory Comput.*, 2:765–774, 2006.

[7] Q. Wu and T. Van Voorhis. Extracting electron transfer coupling elements from constrained density functional theory. *J. Chem. Phys.*, 125:164105, 2006.

[8] Q. Wu and T. Van Voorhis. Direct calculation of electron transfer parameters through constrained density functional theory. *J. Phys. Chem. A*, 110:9212–9218, 2006.

[9] Q. Wu, C.-L. Cheng, and T. Van Voorhis. Configuration interaction based on constrained density functional theory: A multireference method. *J. Chem. Phys.*, 127(16):164119, 2007.

[10] Ze Zhang and Sashi Satpathy. Electron states, magnetism, and the Verwey transition in magnetite. *Phys. Rev. B*, 44(24):13319–13331, 1991.

[11] M. M. Steiner, R. C. Albers, and L. J. Sham. Quasiparticle properties of Fe, Co, and Ni. *Phys. Rev. B*, 45(23):13272–13284, 1992.

[12] I. V. Solovyev and P. H. Dederichs. *Ab initio* calculations of Coulomb U parameters for transition-metal impurities. *Phys. Rev. B*, 49(10):6736–6740, 1994.

[13] G. Vielsack and W. Weber. Search for negative U in the $Ba_{1-x}K_xBi_{1-y}Pb_yO_3$ system using constrained density-functional theory. *Phys. Rev. B*, 54(9):6614–6623, 1996.

[14] W. E. Pickett, S. C. Erwin, and E. C. Ethridge. Reformulation of the LDA$+U$ method for a local-orbital basis. *Phys. Rev. B*, 58(3):1201–1209, 1998.

[15] I. Schnell, G. Czycholl, and R. C. Albers. Hubbard-U calculations for Cu from first-principle Wannier functions. *Phys. Rev. B*, 65(7):075103, 2002.

[16] A. N. Yaresko, V. N. Antonov, and P. Fulde. Localized U $5f$ electrons in UPd_3 from LDA $+ U$ calculations. *Phys. Rev. B*, 67(15):155103, 2003.

[17] Małgorzata Wierzbowska, Anna Delin, and Erio Tosatti. Effect of electron correlations in Pd, Ni, and Co monowires. *Phys. Rev. B*, 72(3):035439, 2005.

[18] Matteo Cococcioni and Stefano de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the LDA $+ U$ method. *Phys. Rev. B*, 71(3):035105, 2005.

[19] I. V. Solovyev and M. Imada. Screening of Coulomb interactions in transition metals. *Phys. Rev. B*, 71(4):045103, 2005.

[20] Kazuma Nakamura, Ryotaro Arita, Yoshihide Yoshimoto, and Shinji Tsuneyuki. First-principles calculation of effective onsite Coulomb interactions of $3d$ transition metals: Constrained local density functional approach with maximally localized Wannier functions. *Phys. Rev. B*, 74(23):235113, 2006.

[21] I. V. Solovyev. First-principles Wannier functions and effective lattice fermion models for narrow-band compounds. *Phys. Rev. B*, 73(15):155117, 2006.

[22] L. Cano-Cortés, A. Dolfen, J. Merino, J. Behler, B. Delley, K. Reuter, and E. Koch. Spectral broadening due to long-range Coulomb interactions in the molecular metal TTF-TCNQ. *Eur. Phys. J. B*, 56:173–176, 2007.

[23] C Tablero. Correlation effects for highly Cr doped ZnSe. *J. Phys.: Condens. Matter*, 19(46):466209, 2007.

[24] Yoshiki Imai, Yuichi Otsuka, and Masatoshi Imada. Applications of path-integral renormalization group method combined with density functional theory. *J. Phys.: Condens. Matter*, 19(36):365230, 2007.

[25] C. Tablero. Correlation and nuclear distortion effects of Cr-substituted ZnSe. *J. Chem. Phys.*, 126(16):164703, 2007.

[26] Vladimir I. Anisimov, Jan Zaanen, and Ole K. Andersen. Band theory and Mott insulators: Hubbard U instead of Stoner I. *Phys. Rev. B*, 44(3):943–954, 1991.

[27] N Stefanou. Electronic structure of $4d$ impurities in Rb: a local-spin-density approximation +U density-functional study. *J. Phys.: Condens. Matter*, 6(50):11221, 1994.

[28] I. V. Solovyev, P. H. Dederichs, and V. I. Anisimov. Corrected atomic limit in the local-density approximation and the electronic structure of d impurities in Rb. *Phys. Rev. B*, 50(23):16861–16871, 1994.

[29] S. Satpathy, Zoran S. Popović, and Filip R. Vukajlović. Electronic structure of the perovskite oxides: $La_{1-x}Ca_xMnO_3$. *Phys. Rev. Lett.*, 76(6):960–963, 1996.

[30] M. Pajda, J. Kudrnovský, I. Turek, V. Drchal, and P. Bruno. *Ab initio* calculations of exchange interactions, spin-wave stiffness constants, and Curie temperatures of Fe, Co, and Ni. *Phys. Rev. B*, 64(17):174402, 2001.

[31] Yoshiki Imai, Igor Solovyev, and Masatoshi Imada. Electronic structure of strongly correlated systems emerging from combining path-integral renormalization group with the density-functional approach. *Phys. Rev. Lett.*, 95(17):176405, 2005.

[32] A. O. Shorikov, A. V. Lukoyanov, M. A. Korotin, and V. I. Anisimov. Magnetic state and electronic structure of the δ and α phases of metallic Pu and its compounds. *Phys. Rev. B*, 72(2):024458, 2005.

[33] A. Yamasaki, L. Chioncel, A. I. Lichtenstein, and O. K. Andersen. Model Hamiltonian parameters for half-metallic ferromagnets NiMnSb and CrO_2. *Phys. Rev. B*, 74(2):024419, 2006.

[34] V. Anisimov, Dm. Korotin, S. Streltsov, A. Kozhevnikov, J. Kuneš, A. Shorikov, and M. Korotin. Density-functional calculation of the Coulomb repulsion and correlation strength in superconducting LaFeAsO. *Sov. Phys. JETP Lett.*, 88:729–733, 2008.

[35] V I Anisimov, Dm M Korotin, M A Korotin, A V Kozhevnikov, J Kuneš, A O Shorikov, S L Skornyakov, and S V Streltsov. Coulomb repulsion and correlation strength in LaFeAsO from density functional and dynamical mean-field theories. *J. Phys.: Condens. Matter*, 21(7):075602, 2009.

[36] Yong Wang, Manish K. Niranjan, J. D. Burton, Joonhee M. An, Kirill D. Belashchenko, and Evgeny Y. Tsymbal. Prediction of a spin-polarized two-dimensional electron gas at the $LaAlO_3$/EuO(001) interface. *Phys. Rev. B*, 79(21):212408, 2009.

[37] Siqi Shi, A. L. Wysocki, and K. D. Belashchenko. Magnetism of chromia from first-principles calculations. *Phys. Rev. B*, 79(10):104404, 2009.

[38] A. O. Shorikov, Z. V. Pchelkina, V. I. Anisimov, S. L. Skornyakov, and M. A. Korotin. Orbital-selective pressure-driven metal to insulator transition in FeO from dynamical mean-field theory. *Phys. Rev. B*, 82(19):195101, 2010.

[39] Hong Jiang, Ricardo I. Gomez-Abal, Patrick Rinke, and Matthias Scheffler. First-principles modeling of localized d states with the GW@LDA$+U$ approach. *Phys. Rev. B*, 82(4):045108, 2010.

[40] J. M. An, S. V. Barabash, V. Ozolins, M. van Schilfgaarde, and K. D. Belashchenko. First-principles study of phase stability of Gd-doped EuO and EuS. *Phys. Rev. B*, 83(6):064105, 2011.

[41] S K Ghosh and V A Singh. Energy theorems in constrained density-functional theory. *J. Phys.: Condens. Matter*, 1(11):1971, 1989.

[42] R. Astala and M. J. Stott. Kohn-Sham inversion and iterative energy minimization. *Phys. Rev. B*, 73(11):115127, 2006.

[43] H. Akai, S. Blügel, R. Zeller, and P. H. Dederichs. Isomer shifts and their relation to charge transfer in dilute Fe alloys. *Phys. Rev. Lett.*, 56(22):2407–2410, 1986.

[44] Oleg Pankratov and Matthias Scheffler. Localized excitons and breaking of chemical bonds at III-V (110) surfaces. *Phys. Rev. Lett.*, 75(4):701–704, 1995.

[45] M. Methfessel, Vincenzo Fiorentini, and Sabrina Oppo. Connection between charge transfer and alloying core-level shifts based on density-functional calculations. *Phys. Rev. B*, 61(8):5229–5236, 2000.

[46] L. R. C. Fonseca, J. L. Jimenez, J. P. Leburton, and Richard M. Martin. Self-consistent calculation of the electronic structure and electron-electron interaction in self-assembled InAs-GaAs quantum dot structures. *Phys. Rev. B*, 57(7):4017–4026, 1998.

[47] O. Eriksson, J. M. Wills, M. Colarieti-Tosti, S. Lebègue, and A. Grechnev. Many-body projector orbitals for electronic structure theory of strongly correlated electrons. *Int. J. Quantum Chem.*, 105(2):160–165, 2005.

[48] B Hourahine, B Aradi, and T Frauenheim. DFTB+ and lanthanides. *J. Phys. Conf. Ser.*, 242(1):012005, 2010.

[49] G. F. Voronoi. New applications of continuous parameters to the theory of quadratic forms: Part 2. research on primitive polyhedra. *J. Reine Angew. Math.*, 134:198–287, 1908.

[50] A. D. Becke. A multicenter numerical integration scheme for polyatomic molecules. *J. Chem. Phys.*, 88:2547–2553, 1988.

[51] F. L. Hirshfeld. Bonded-atom fragments for describing molecular charge densities. *Theor. Chim. Acta*, 44:129–138, 1977.

[52] Patrick Bultinck, Christian Van Alsenoy, Paul W. Ayers, and Ramon Carbó-Dorca. Critical analysis and extension of the Hirshfeld atoms in molecules. *J. Chem. Phys.*, 126(14):144111, 2007.

[53] Timothy C. Lillestolen and Richard J. Wheatley. Redefining the atom: atomic charge densities produced by an iterative Stockholder approach. *Chem. Commun.*, 45:5909–5911, 2008.

[54] Morrel H. Cohen and Adam Wasserman. Revisiting N-continuous density-functional theory: chemical reactivity and "atoms" in "molecules". *Israel J. Chem.*, 43:219–227, 2004.

[55] Morrel H. Cohen and Adam Wasserman. On hardness and electronegativity equalization in chemical reactivity theory. *J. Stat. Phys.*, 125:1121–1139, 2006.

[56] Morrel H. Cohen and Adam Wasserman. On the foundations of chemical reactivity theory. *J. Phys. Chem. A*, 111(11):2229–2242, 2007.

[57] R. F. W. Bader. *Atoms in molecules — a quantum theory.* Oxford University Press, Oxford, 1990.

[58] R. S. Mulliken. Electronic populations analysis on LCAO-MO molecular wave functions. *J. Chem. Phys.*, 23:1833–1840, 1955.

[59] P.-O. Löwdin. On the non-orthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *J. Chem. Phys.*, 18:365–375, 1950.

[60] J. P. Foster and F. Weinhold. Natural hybrid orbitals. *J. Am. Chem. Soc.*, 102(24):7211–7218, 1980.

[61] Alan E. Reed, Robert B. Weinstock, and Frank Weinhold. Natural population analysis. *J. Chem. Phys.*, 83(2):735–746, 1985.

[62] Hyunju Chang, J. F. Harrison, T. A. Kaplan, and S. D. Mahanti. Cluster study of the neutron-scattering form factor for antiferromagnetic $KNiF_3$ and NiO. *Phys. Rev. B*, 49(22):15753–15758, 1994.

[63] H. Oberhofer and J. Blumberger. Charge constrained density functional molecular dynamics for simulation of condensed phase electron transfer reactions. *J. Chem. Phys.*, 131:064101, 2009.

[64] I. Rudra, Q. Wu, and T. Van Voorhis. Accurate magnetic exchange couplings in transition-metal complexes from constrained density-functional theory. *J. Chem. Phys.*, 124:024103, 2006.

[65] Benjamin Kaduk, Tim Kowalczyk, and Troy Van Voorhis. Constrained density functional theory. *Chem. Rev.*, 112:321–370, 2012.

[66] Ernest R. Davidson and Larry E. Nitzsche. Vertical excitation energy to the lowest $^1\pi\pi^*$ state of acrolein. *J. Am. Chem. Soc.*, 101:6524–6526, 1979.

[67] P. Pulay. Convergence acceleration of iterative sequences. the case of SCF iteration. *Chem. Phys. Lett.*, 73:393–398, 1980.

[68] P. Pulay. Improved SCF convergence acceleration. *J. Comput. Chem.*, 3:556–560, 2004.

[69] P. H.-L. Sit, M. Cococcioni, and N. Marzari. Realistic quantitative descriptions of electron transfer reactions: Diabatic free-energy surfaces from first-principles molecular dynamics. *Phys. Rev. Lett.*, 97:028303, 2006.

[70] P. Ghosh and R. Gebauer. Computational approaches to charge transfer excitations in a zinc tetraphenylporphyrin and C_{70} complex. *J. Chem. Phys.*, 132:104102, 2010.

[71] Seth Difley, Lee-Ping Wang, Sina Yeganeh, Shane R. Yost, and Troy Van Voorhis. Electronic properties of disordered organic semiconductors via QM/MM simulations. *Acc. Chem. Res.*, 43(7):995–1004, 2010.

[72] Benny G. Johnson, Peter M. W. Gill, and John A. Pople. The performance of a family of density functional methods. *J. Chem. Phys.*, 98(7):5612–5626, 1993.

[73] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. *Molecular electronic-structure theory.* John Wiley and Sons, New York, 2000.

[74] Q. Wu, B. Kaduk, and T. Van Voorhis. Constrained density functional theory based configuration interaction improves the prediction of reaction barrier heights. *J. Chem. Phys.*, 130(3):034109, 2009.

[75] B. Kaduk and T. Van Voorhis. Conical intersections using constrained density functional theory–configuration interaction. *J. Chem. Phys.*, 133:061102, 2010.

[76] Harald Oberhofer and Jochen Blumberger. Electronic coupling matrix elements from charge constrained density functional theory calculations using a plane wave basis set. *J. Chem. Phys.*, 133:244105, 2010.

[77] J. E. Subotnik, S. Yeganeh, R. J. Cave, and M. A. Ratner. Constructing diabatic states from adiabatic states: Extending generalized Mulliken Hush to multiple charge centers with Boys localization. *J. Chem. Phys.*, 129:244101, 2008.

[78] T. Pacher, H. Köppel, and L. S. Cederbaum. Quasidiabatic states from ab initio calculations by block diagonalization of the electronic Hamiltonian: Use of frozen orbitals. *J. Chem. Phys.*, 95:6668–6680, 1991.

[79] T. Pacher, L.S. Cederbaum, and H. Köppel. Adiabatic and quasidiabatic states in a gauge theoretical framework. *Adv. Chem. Phys.*, 84:293–391, 1993.

[80] Paola Cattaneo and Maurizio Persico. Ab initio determination of quasi-diabatic states for multiple reaction pathways. *Chem. Phys.*, 214(1):49–60, 1997.

[81] Klaus Ruedenberg and Gregory J. Atchity. A quantum chemical determination of diabatic states. *J. Chem. Phys.*, 99:3799–3803, 1993.

[82] A. Thiel and H. Köppel. Proposal and numerical test of a simple diabatization scheme. *J. Chem. Phys.*, 110:9371–9383, 1999.

[83] H. Köppel, J. Gronki, and S. Mahapatra. Construction scheme for regularized diabatic states. *J. Chem. Phys.*, 115:2377–2388, 2001.

[84] R. J. Cave and M. D. Newton. Generalization of the Mulliken-Hush treatment for the calculation of electron transfer matrix elements. *Chem. Phys. Lett.*, 249:15–19, 1996.

[85] R. J. Cave and M. D. Newton. Calculation of electronic coupling matrix elements for ground and excited state electron transfer reactions: Comparison of the generalized Mulliken-Hush and block diagonalization methods. *J. Chem. Phys.*, 106(22):9213–9226, 1997.

[86] H. Sponer and E. Teller. Electronic spectra of polyatomic molecules. *Rev. Mod. Phys.*, 13:75, 1941.

[87] A. Warshel and R. M. Weiss. An empirical valence bond approach for comparing reactions in solutions and in enzymes. *J. Am. Chem. Soc.*, 102:6218–6226, 1980.

[88] Johan Åqvist and Arieh Warshel. Simulation of enzyme reactions using valence bond force fields and other hybrid quantum/classical approaches. *Chem. Rev.*, 93(7):2523–2544, 1993.

[89] T. Pacher, L. S. Cederbaum, and H. Köppel. Approximately diabatic states from block diagonalization of the electronic Hamiltonian. *J. Chem. Phys.*, 89(12):7367–7381, 1988.

[90] Hyung J. Kim and James T. Hynes. A theoretical model for S_N1 ionic dissociation in solution. 1. activation free energetics and transition-state structure. *J. Am. Chem. Soc.*, 114:10508–10528, 1992.

[91] Hisao Nakamura and Donald G. Truhlar. The direct calculation of diabatic states based on configurational uniformity. *J. Chem. Phys.*, 115(22):10353–10372, 2001.

[92] Hisao Nakamura and Donald G. Truhlar. Direct diabatization of electronic states by the fourfold way. II. dynamical correlation and rearrangement processes. *J. Chem. Phys.*, 117(12):5576–5593, 2002.

[93] Hisao Nakamura and Donald G. Truhlar. Extension of the fourfold way for calculation of global diabatic potential energy surfaces of complex, multiarrangement, non-Born–Oppenheimer systems: Application to HNCO(S_0,S_1). *J. Chem. Phys.*, 118(15):6816–6829, 2003.

[94] Gregory J. Atchity and Klaus Ruedenberg. Determination of diabatic states through enforcement of configurational uniformity. *Theor. Chim. Acta*, 97:47–58, 1997.

[95] T. Pacher, C. A. Mead, L. S. Cederbaum, and H. Köppel. Gauge theory and quasidiabatic states in molecular physics. *J. Chem. Phys.*, 91:7057–7062, 1989.

[96] Lingchun Song and Jiali Gao. On the construction of diabatic and adiabatic potential energy surfaces based on *ab initio* valence bond theory. *J. Phys. Chem. A*, 112:12925–12935, 2008.

[97] Y. Mo and J. Gao. An ab initio molecular orbital-valence bond (MOVB) method for simulating chemical reactions in solution. *J. Phys. Chem. A*, 104:3012–3020, 2000.

[98] George D. Purvis III and Rodney J. Bartlett. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.*, 76:1910–1918, 1982.

[99] C. David Sherrill, Anna I. Krylov, Edward F. C. Byrd, and Martin Head-Gordon. Energies and analytic gradients for a coupled-cluster doubles model using variational brueckner orbitals: Application to symmetry breaking in O_4^+. *J. Chem. Phys.*, 109(11):4171–4181, 1998.

[100] John P. Perdew and Mel Levy. Comment on "Significance of the highest occupied Kohn-Sham eigenvalue". *Phys. Rev. B*, 56(24):16021–16028, 1997.

[101] Y. Zhang and W. Yang. A challenge for density functionals: self-interaction error increases for systems with a noninteger number of electrons. *J. Chem. Phys.*, 109:2604–2608, 1998.

[102] Thomas Bally and G. Narahari Sastry. Incorrect dissociation behavior of radical ions in density functional calculations. *J. Phys. Chem. A*, 101:7923–7925, 1997.

[103] A. Ruzsinszky, J. P. Perdew, G. I. Csonka, O. A. Vydrov, and G. E. Scuseria. Density functionals that are one- and two- are not always many-electron self-interaction-free, as shown for H_2^+, He_2^+, LiH^+, and Ne_2^+. *J. Chem. Phys.*, 126:104102, 2007.

[104] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.*, 125(20):201102, 2006.

[105] J. L. Durant. Evaluation of transition state properties by density functional theory. *Chem. Phys. Lett.*, 256:595–602, 1996.

[106] A. Daniel Boese and Jan M. L. Martin. Development of density functionals for thermochemical kinetics. *J. Chem. Phys.*, 121:3405–3416, 2004.

[107] Y. Zhao, N. E. Schultz, and D. G. Truhlar. Design of density functionals by combining the method of constraint satisfaction with parametrization for thermochemistry, thermochemical kinetics, and noncovalent interactions. *J. Chem. Theory Comput.*, 2:364–382, 2006.

[108] O. A. Vydrov, J. Heyd, A. V. Krukau, and G. E. Scuseria. Importance of short-range versus long-range Hartree-Fock exchange on the performance of hybrid density functionals. *J. Chem. Phys.*, 125:074106, 2006.

[109] Paula Mori-Sánchez, Aron J. Cohen, and Weitao Yang. Self-interaction-free exchange-correlation functional for thermochemistry and kinetics. *J. Chem. Phys.*, 124:091102, 2006.

[110] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Development of exchange-correlation functionals with minimal many-electron self-interaction error. *J. Chem. Phys.*, 126:191109, 2007.

[111] O. A. Vydrov, G. E. Scuseria, and J. P. Perdew. Tests of functionals for systems with fractional electron number. *J. Chem. Phys.*, 126:154109, 2007.

[112] P. Mori-Sánchez, A. J. Cohen, and W. Yang. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Phys. Rev. Lett.*, 100:146401, 2008.

[113] Axel D. Becke and Erin R. Johnson. A unified density-functional treatment of dynamical, nondynamical, and dispersion correlations. *J. Chem. Phys.*, 127:124108, 2007.

[114] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Assessment and formal properties of exchange-correlation functionals constructed from the adiabatic connection. *J. Chem. Phys.*, 127:034101, 2007.

[115] Ross M. Dickson and Axel D. Becke. Reaction barrier heights from an exact-exchange-based density-functional correlation model. *J. Chem. Phys.*, 123:111101, 2005.

[116] A. J. Cohen, P. Mori-Sánchez, and W. Yang. Insights into current limitations of density functional theory. *Science*, 321:792–794, 2008.

[117] Wei Wu and Sason Shaik. VB-DFT: a nonempirical hybrid method combining valence bond theory and density functional energies. *Chem. Phys. Lett.*, 301:37–42, 1999.

[118] Rodney J. Bartlett, Victor F. Lotrich, and Igor V. Schweigert. Ab initio density functional theory: The best of both worlds? *J. Chem. Phys.*, 123:062205, 2005.

[119] J. Gräfenstein and D. Cremer. Development of a CAS-DFT method covering non-dynamical and dynamical electron correlation in a balanced way. *Mol. Phys.*, 103:279–308, 2005.

[120] T. Leininger, H. Stoll, H. J. Werner, and A. Savin. Combining long-range configuration interaction with short-range density functionals. *Chem. Phys. Lett.*, 275:151–160, 1997.

[121] Stefan Grimme and Mirko Waletzke. A combination of Kohn-Sham density functional theory and multi-reference configuration interaction methods. *J. Chem. Phys.*, 111:5645–5655, 1999.

[122] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, 1965.

[123] J. Behler, B. Delley, S. Lorenz, K. Reuter, and M. Scheffler. Dissociation of O_2 at Al(111): The role of spin selection rules. *Phys. Rev. Lett.*, 94:036104, 2005.

[124] J. Behler, B. Delley, K. Reuter, and M. Scheffler. Nonadiabatic potential-energy surfaces by constrained density functional theory. *Phys. Rev. B*, 75:115409, 2007.

[125] J. R. Schmidt, N. Shenvi, and J. C. Tully. Controlling spin contamination using constrained density functional theory. *J. Chem. Phys.*, 129:114110, 2008.

[126] Udo W. Schmitt and Gregory A. Voth. The computer simulation of proton transport in water. *J. Chem. Phys.*, 111:9361–9381, 1999.

[127] S. Difley and T. Van Voorhis. Exciton/charge-transfer electronic couplings in organic semiconductors. *J. Chem. Theory Comput.*, 7(3):594–601, 2011.

[128] X.-Y. Zhu, Q. Yang, and M. Muntwiler. Charge-transfer excitons at organic semiconductor surfaces and interfaces. *Acc. Chem. Res.*, 42(11):1779–1787, 2009.

[129] Veaceslav Coropceanu, Jérôme Cornil, Demetrio A. da Silva Filho, Yoann Olivier, Robert Silbey, and Jean-Luc Brédas. Charge transport in organic semiconductors. *Chem. Rev.*, 107(4):926–952, 2007.

[130] T. Kowalczyk, Z. Lin, and T. Van Voorhis. Fluorescence quenching by photoinduced electron transfer in the Zn^{2+} sensor zinpyr-1: A computational investigation. *J. Phys. Chem. A*, 114:10427–10434, 2010.

[131] Yan Zhao, Núria González-García, and Donald G. Truhlar. Benchmark database of barrier heights for heavy atom transfer, nucleophilic substitution, association, and unimolecular reactions and its use to test theoretical methods. *J. Phys. Chem. A*, 109(9):2012–2018, 2005.

[132] John Clarke Slater. *The Self-Consistent Field for Molecules and Solids. Quantum Theory of Molecules and Solids, Vol. 4.* McGraw-Hill, New York, 1974.

[133] S. H. Vosko, L. Wilk, and M. Nusair. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis. *Can. J. Phys.*, 58:1200–1211, 1980.

[134] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38:3098–3100, 1988.

[135] Chengteh Lee, Weitao Yang, and Robert G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37:785–789, 1988.

[136] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–3868, 1996.

[137] A. D. Becke. Density-functional thermochemistry. III. the role of exact exchange. *J. Chem. Phys.*, 98:5648–5652, 1993.

[138] Axel D. Becke. Density-functional thermochemistry. V. systematic optimization of exchange-correlation functionals. *J. Chem. Phys.*, 107:8554–8560, 1997.

[139] Philip J. Wilson, Thomas J. Bradley, and David J. Tozer. Hybrid exchange-correlation functional determined from thermochemical data and *ab initio* potentials. *J. Chem. Phys.*, 115:9233–9242, 2001.

[140] Peter M.W Gill, Benny G Johnson, and John A Pople. A standard grid for density functional calculations. *Chem. Phys. Lett.*, 209:506–512, 1993.

[141] Benjamin G. Janesko and Gustavo E. Scuseria. Hartree-Fock orbitals significantly improve the reaction barrier heights predicted by semilocal density functionals. *J. Chem. Phys.*, 128:244112, 2008.

[142] Sason Shaik, Devesh Kumar, and Sam P. de Visser. A valence bond modeling of trends in hydrogen abstraction barriers and transition states of hydroxylation reactions catalyzed by cytochrome P450 enzymes. *J. Am. Chem. Soc.*, 130:10128–10140, 2008.

[143] E. Runge and E. K. U. Gross. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.*, 52:997–1000, 1984.

[144] Mark E. Casida, Christine Jamorski, Kim C. Casida, and Dennis R. Salahub. Molecular excitation energies to high-lying bound states from time-dependent density-functional response theory: Characterization and correction of the time-dependent local density approximation ionization threshold. *J. Chem. Phys.*, 108:4439–4449, 1998.

[145] Chao-Ping Hsu, So Hirata, and Martin Head-Gordon. Excitation energies from time-dependent density functional theory for linear polyene oligomers: Butadiene to decapentaene. *J. Phys. Chem. A*, 105:451–458, 2001.

[146] Peter Elliott, Filipp Furche, and Kieron Burke. Excited states from time-dependent density functional theory. *Rev. Comp. Chem.*, 28:91, 2008.

[147] David R. Yarkony. Diabolical conical intersections. *Rev. Mod. Phys.*, 68(4):985, 1996.

[148] B. G. Levine, C. Ko, J. Quenneville, and T. J. Martínez. Conical intersections and double excitations in time-dependent density functional theory. *Mol. Phys.*, 104:1039–1051, 2006.

[149] Charles W. Bauschlicher and Stephen R. Langhoff. Full configuration-interaction study of the ionic-neutral curve crossing in LiF. *J. Chem. Phys.*, 89:4246–4254, 1988.

[150] A. Dreuw and M. Head-Gordon. Failure of time-dependent density functional theory for long-range charge-transfer excited states: The zincbacteriochlorin-bacterlochlorin and bacteriochlorophyll-spheroidene complexes. *J. Am. Chem. Soc.*, 126:4007–4016, 2004.

[151] Neepa T. Maitra, Fan Zhang, Robert J. Cave, and Kieron Burke. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.*, 120:5932–5937, 2004.

[152] Robert J. Cave, Fan Zhang, Neepa T. Maitra, and Kieron Burke. A dressed TDDFT treatment of the $2\,^1A_g$ states of butadiene and hexatriene. *Chem. Phys. Lett.*, 389:39–42, 2004.

[153] K.J.H. Giesbertz and E.J. Baerends. Failure of time-dependent density functional theory for excited state surfaces in case of homolytic bond dissociation. *Chem. Phys. Lett.*, 461:338–342, 2008.

[154] A. Zangwill and Paul Soven. Density-functional approach to local-field effects in finite systems: Photoabsorption in the rare gases. *Phys. Rev. A*, 21:1561–1572, 1980.

[155] E. K. U. Gross and Walter Kohn. Local density-functional theory of frequency-dependent linear response. *Phys. Rev. Lett.*, 55:2850–2852, 1985.

[156] A. J. C. Varandas. Energy switching approach to potential surfaces. II. two-valued function for the water molecule. *J. Chem. Phys.*, 107:867–878, 1997.

[157] Ioannis S. K. Kerkines, Ioannis D. Petsalakis, Panagiotis Argitis, and Giannoula Theodorakopoulos. Fluorescence properties of organic dyes: quantum chemical studies on the green/blue neutral and protonated DMA-DPH emitters in polymer matrices. *Phys. Chem. Chem. Phys.*, 13:21273–21281, 2011.

[158] Denis Jacquemin, Eric A. Perpète, Giovanni Scalmani, Michael J. Frisch, Xavier Assfeld, Ilaria Ciofini, and Carlo Adamo. Time-dependent density functional theory investigation of the absorption, fluorescence, and phosphorescence spectra of solvated coumarins. *J. Chem. Phys.*, 125:164324, 2006.

[159] Michael Hartmann, Jiří Pittner, and Vlasta Bonačić-Koutecký. *Ab initio* nonadiabatic dynamics involving conical intersection combined with Wigner distribution approach to ultrafast spectroscopy illustrated on Na_3F_2 cluster. *J. Chem. Phys.*, 114:2123, 2001.

[160] Alexey L. Kaledin and Keiji Morokuma. An *ab initio* direct-trajectory study of the photodissociation of ClOOCl. *J. Chem. Phys.*, 113:5750, 2000.

[161] Uwe Müller and Gerhard Stock. Surface-hopping modeling of photoinduced relaxtion dynamics on coupled potential-energy surfaces. *J. Chem. Phys.*, 107:6230, 1997.

[162] Dominik Schemmel and Martin Schütz. Molecular aniline cluster. II. the low-lying electronic excited states. *J. Chem. Phys.*, 133:134307, 2010.

[163] Andrzej L. Sobolewski and Wolfgang Domcke. Computational studies of the photophysics of hydrogen-bonded molecular systems. *J. Phys. Chem. A*, 111:11725–11735, 2007.

[164] Joshua D. Coe and Todd J. Martínez. *Ab initio* multiple spawning dynamics of excited state intramolecular proton transfer: the role of spectroscopically dark states. *Mol. Phys.*, 106:537–545, 2008.

[165] Mihajlo Etinski and Christel M. Marian. *Ab initio* investigation of the methylation and hydration effects on the electronic spectra of uracil and thymine. *Phys. Chem. Chem. Phys.*, 12:4915–4923, 2010.

[166] Adrian W. Lange and John M. Herbert. Both intra- and interstrand charge-transfer excited states in aqueous B-DNA are present at energies comparable to, or just above, the $^1\pi\pi^*$ excitonic bright states. *J. Am. Chem. Soc.*, 131:3913–3922, 2009.

[167] Hanneli R. Hudock and Todd J. Matrínez. Excited-state dynamics of cytosine reveal multiple intrinsic subpicosecond pathways. *ChemPhysChem*, 9:2486–2490, 2008.

[168] Gerrit Groenhof, Lars V. Schäfer, Martial Boggio-Pasqua, Maik Goette, Helmut Grubmüller, and Michael A Robb. Ultrafast deactivation of an excited cytosine-guanine base pair in DNA. *J. Am. Chem. Soc.*, 129:6812–6819, 2007.

[169] Wolfgang J. Schreier, Tobias E. Schrader, Florian O. Koller, Peter Gilch, Carlos E. Crespo-Hernández, Vijay N. Swaminathan, Thomas Carell, Wolfgang Zinth, and Bern Kohler. Thymine dimerization in DNA is an ultrafast photoreaction. *Science*, 315:625–629, 2007.

[170] Carlos E. Crespo-Hernández, Boiko Cohen, and Bern Kohler. Base stacking controls excited-state dynamics in A · T DNA. *Nature*, 436:1141–1144, 2005.

[171] Thomas Gustavsson, Ákos Bányász, Elodie Lazzarotto, Dimitra Markovitsi, Giovanni Scalmani, Michael J. Frisch, Vincenzo Barone, and Roberto Improta. Singlet excited-state behavior of uracil and thymine in aqueous solution: a combined experimental

and computational study of 11 uracil derivatives. *J. Am. Chem. Soc.*, 128:607–619, 2006.

[172] Jean-Marc L. Pecourt, Jorge Peon, and Bern Kohler. DNA excited-state dynamics: ultrafast internal conversion and vibrational cooling in a series of nucleosides. *J. Am. Chem. Soc.*, 123:10370–10378, 2001.

[173] Spiridoula Matsika. Radiationless decay of excited states of uracil through conical intersections. *J. Phys. Chem. A*, 108:7584–7590, 2004.

[174] Hanneli R. Hudock, Benjamin G. Levine, Alexis L. Thompson, and Todd J. Martinez. First principles dynamics of photoexcited DNA and RNA bases. *AIP Conf. Proc.*, 963:219–222, 2007.

[175] T. Kowalczyk, S. R. Yost, and T. Van Voorhis. Assessment of the ΔSCF density functional theory approach for electronic excitations in organic dyes. *J. Chem. Phys.*, 134:054128, 2011.

[176] Jérôme Cornil, David Beljonne, Jean-Philippe Calbert, and Jean-Luc Brédas. Interchain interactions in organic π-conjugated materials: impact on electronic structure, optical response, and charge transport. *Adv. Mater.*, 13:1053–1067, 2001.

[177] R. H. Friend, R. W. Gymer, A. B. Holmes, J. H. Burroughes, R. N. Marks, Cl. Taliani, D. D. C. Bradley, D. A. Dos Santos, J. L. Brédas, M. Lögdlund, and W. R. Salaneck. Electroluminescence in conjugated polymers. *Nature*, 397:121–128, 1999.

[178] Jean-Luc Brédas, Joseph E. Norton, Jérôme Cornil, and Veaceslav Coropceanu. molecular understanding of organic solar cells: the challenges. *Acc. Chem. Res.*, 42:1691–1699, 2009.

[179] Adélia J. A. Aquino, Mario Barbatti, and Hans Lischka. Excited-state properties and environmental effects for protonated Schiff basis: a theoretical study. *ChemPhysChem*, 7:2089–2096, 2006.

[180] Olli Lehtonen, Dage Sundholm, and Tommy Vänskä. Computational studies of semiconductor quantum dots. *Phys. Chem. Chem. Phys.*, 10:4535–4550, 2008.

[181] Leszek Lapinski, Maciej J. Nowak, Jacek Nowacki, Michat F. Rode, and Andrzej L. Sobolewski. A bistable molecular switch driven by photoinduced hydrogen-atom transfer. *ChemPhysChem*, 10:2290–2295, 2009.

[182] Carole Van Caillie and Roger D. Amos. Geometric derivatives of excitation energies using SCF and DFT. *Chem. Phys. Lett.*, 308:249–255, 1999.

[183] Carole Van Caillie and Roger D. Amos. Geometric derivatives of density functional theory excitation energies using gradient-corrected functionals. *Chem. Phys. Lett.*, 317:159–164, 2000.

[184] Filipp Furche and Reinhart Ahlrichs. Adiabatic time-dependent density functional methods for excited state properties. *J. Chem. Phys.*, 117:7433–7447, 2002.

[185] Robert Send, Michael Kuhn, and Filipp Furche. Assessing excited state methods by adiabatic excitation energies. *J. Chem. Theory Comput.*, 7:2376–2386, 2011.

[186] Mario Barbatti, Giovanni Granucci, Maurizio Persico, Matthias Ruckenbauer, Mario Vazdar, Mirjana Eckert-Maksić, and Hans Lischka. The on-the-fly surface-hopping program system NEWTON-X: application to ab initio simulation of the nonadiabatic photodynamics of benchmark systems. *J. Photochem. Photobiol. A*, 190:228–240, 2007.

[187] J. Plötner and A. Dreuw. Pigment Yellow 101: a showcase for photo-initiated processes in medium-sized molecules. *Chem. Phys.*, 347:472–482, 2007.

[188] Paul Wiggins, J. A. Gareth Williams, and David J. Tozer. Excited state surfaces in density functional theory: a new twist on an old problem. *J. Chem. Phys.*, 131:091101, 2009.

[189] A. Dreuw and M. Head-Gordon. Single-reference ab initio methods for the calculation of excited states of large molecules. *Chem. Rev.*, 105:4009–4037, 2005.

[190] Michael Wormit and Anndreas Dreuw. Carotenoid radical cation formation in LH2 of purple bacteria: a quantum chemical study. *J. Phys. Chem. B*, 110:24200–24206, 2006.

[191] James B. Foresman, Martin Head-Gordon, John A. Pople, and Michael J. Frisch. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.*, 96:135–149, 1992.

[192] Davir Maurice and Martin Head-Gordon. Analytical second derivatives for excited electronic states using the single excitation configuration interaction method: theory and application to benzo[a]pyrene and chalcone. *Mol. Phys.*, 96:1533–1541, 1999.

[193] Chang ik Song and Young Min Rhee. Dynamics on the electronically excited state surface of the bioluminescent firefly luciferase-oxyluciferin system. *J. Am. Chem. Soc.*, 133:12040–12049, 2011.

[194] R. Nithya, N. Santhanamoorthi, P. Kolandaivel, and K. Senthilkumar. Structural and spectral properties of 4-bromo-1-naphthyl chalcones: A quantum chemical study. *J. Phys. Chem. A*, 115:6594–6602, 2011.

[195] Wolfgang J. Schreier, Igor Pugliesi, Florian O. Koller, Tobias E. Schrader, Wolfgang Zinth, and Markus Braun. Vibrational spectra of the ground and the singlet excited $\pi\pi^*$ state of 6,7-dimethyl-8-ribityllumazine. *J. Phys. Chem. B*, 115:3689–3697, 2011.

[196] Gianfranco Bocchinfuso, Claudia Mazzuca, Antonio Palleschi, Roberto Pizzoferrato, and Pietro Tagliatesta. Photophysical properties of 1,3,5-tris(2-naphthyl)benzene and related less-arylated compounds: Experimental and theoretical investigations. *J. Phys. Chem. A*, 113:14887–14895, 2009.

[197] Kerstin Andersson and Björn O. Roos. Multiconfigurational second-order perturbation theory: a test of geometries and binding energies. *Int. J. Quantum Chem.*, 45:591–607, 1993.

[198] Manuela Merchán and Björn O. Roos. A theoretical determination of the electronic spectrum of formaldehyde. *Theor. Chim. Acta*, 92:227–239, 1995.

[199] Arvi Rauk, Dake Yu, Piotr Borowski, and Björn Roos. CASSCF, CASPT2, and MRCI investigations of formyloxyl radical ($HCOO^{\cdot}$). *Chem. Phys.*, 197:73–80, 1995.

[200] Christopher S. Page and Massimo Olivucci. Ground and excited state CASPT2 geometry optimizations of small organic molecules. *J. Comput. Chem.*, 24:289–309, 2003.

[201] S. Olsen, Kristina Lamonthe, and Todd J. Martinez. Protonic gating of excited-state twisting and charge localization in GFP chromophores: a mechanistic hypothesis for reversible photoswitching. *J. Am. Chem. Soc.*, 132:1192–1193, 2010.

[202] Xue fang Yu, Shohei Yamazaki, and Tetsuya Taketsugu. Concerted or stepwise mechanism? CASPT2 and LC-TDDFT study of the excited-state double proton transfer in the 7-azaindole dimer. *J. Chem. Theory Comput.*, 7:1006–1015, 2011.

[203] P. B. Coto, A. Strambi, and M. Olivucci. Effect of opsin on the shape of the potential energy surfaces at the conical intersection of the Rhodopsin chromophore. *Chem. Phys.*, 347:483–491, 2008.

[204] Paolo Celani and Hans-Joachim Werner. Analytical energy gradients for internally contracted second-order multireference perturbation theory. *J. Chem. Phys.*, 119:5044–5057, 2003.

[205] John F. Stanton. Many-body methods for excited state potential energy surfaces. I. general theory of energy gradients for the equation-of-motion coupled-cluster method. *J. Chem. Phys.*, 99:8840–8847, 1993.

[206] John F. Stanton and Jürgen Gauss. Analytic energy gradients for the equation-of-motion coupled-cluster method: Implementation and application to the HCN/HNC system. *J. Chem. Phys.*, 100:4695–4698, 1994.

[207] John F. Stanton and Jürgen Gauss. Analytic energy derivatives for the equation-of-motion coupled-cluster method: algebraic expressions, implementation and application to the S_1 state of HFCO. *Theor. Chim. Acta*, 91:267–289, 1995.

[208] Lori A. Burns, Daniel Murdock, and Patrick H. Vaccaro. An exploration of electronic structure and nuclear dynamics in tropolone: II. the $\tilde{A}\,^1B_2(\pi^*\pi)$ excited state. *J. Chem. Phys.*, 130:144304, 2009.

[209] Dennis Gerbig, Hans Peter Reisenauer, Chia-Hua Wu, David Ley, Wesley D. Allen, and Peter R. Schreiner. Phenylhydroxycarbene. *J. Am. Chem. Soc.*, 132:7273–7275, 2010.

[210] Ove Christiansen, John F. Stanton, and Jürgen Gauss. A coupled cluster study of the 1^1A_{1g} and 1^1B_{2u} states of benzene. *J. Chem. Phys.*, 108:3987–4001, 1998.

[211] Andreas Köhn and Christof Hättig. Analytic gradients for excited states in the coupled-cluster model CC2 employing the resolution-of-the-identity approximation. *J. Chem. Phys.*, 119:5021–5036, 2003.

[212] Robert Brause, Monika Santa, Michael Schmitt, and Karl Kleinermanns. Determination of the geometry change of the phenol dimer upon electronic excitation. *ChemPhysChem*, 8:1394–1401, 2007.

[213] Yunlong Zhang, Gotard Burdzinski, Jacek Kubickki, Shubham Vyas, Christopher M. Hadad, Michel Sliwa, Olivier Poizat, Guy Buntinx, and Matthew S. Platz. Study of the S_1 excited state of para-methoxy-3-phenyl-3-methyl diazirine by ultrafast time resolved UV-Vis and IR spectroscopies and theory. *J. Am. Chem. Soc.*, 131:13784–13790, 2009.

[214] Evgeniy V. Gromov, Irene Burghardt, Horst Köppel, and Lorenz S. Cederbaum. Photoinduced isomerization of the photoactive yellow protein (PYP) chromophore: Interplay of two torsions, a HOOP mode and hydrogen bonding. *J. Phys. Chem. A*, 115:9237–9248, 2011.

[215] Szymon Smolarek, Alexander Vdovin, Anouk Rijs, Cornelis A. van Walree, Marek Z. Zgierski, and Wybren J. Buma. High-resolution spectroscopy of jet-cooled 1,1'-diphenylethylene: Electronically excited and ionic states of a prototypical cross-conjugated system. *J. Phys. Chem. A*, 115:9399–9410, 2011.

[216] Michał F. Rode and Andrzej L. Sobolewski. Effect of chemical substituents on the energetical landscape of a molecular photoswitch: An ab initio study. *J. Phys. Chem. A*, 114:11879–11889, 2010.

[217] Hans-Joachim Werner and Ernst-Albrecht Reinsch. The self-consistent electron pairs method for multiconfiguration reference state functions. *J. Chem. Phys.*, 76:3144, 1982.

[218] Ron Shepard, Isaiah Shavitt, Russell M. Pitzer, Donald C. Comeau, Melanie Pepper, Hans Lischka, Peter G. Szalay, Reinhart Ahlrichs, Franklin B. Brown, and Jian-Guo Zhao. A progress report on the status of the COLUMBUS MRCI program system. *Int. J. Quantum Chem.*, 34:149–165, 1988.

[219] Hans Lischka, Ron Shepard, Russell M. Pitzer, Isaiah Shavitt, Michal Dallos, Thomas Müller, Péter G. Szalay, Michael Seth, Gary S. Kedziora, Satoshi Yabushita, and Zhiyong Zhang. High-level multireference methods in the quantum-chemistry program system COLUMBUS: analytic MR-CISD and MR-AQCC gradients and MR-AQCC-LRT for excited states, GUGA spin-orbit CI and parallel CI density. *Phys. Chem. Chem. Phys.*, 3:664–673, 2001.

[220] H. Lischka, M. Dallos, and R. Shepard. Analytic MRCI gradient for excited states: formalism and application to the $n - \pi^*$ valence- and $n - (3s, 3p)$ Rydberg states of formaldehyde. *Mol. Phys.*, 100:1647–1658, 2002.

[221] A. Dreuw, J. L. Weisman, and M. Head-Gordon. Long-range charge-transfer excited states in time-dependent density functional theory require non-local exchange. *J. Chem. Phys.*, 119:2943–2946, 2003.

[222] Michael J. G. Peach, Peter Benfield, Trygve Helgaker, and David J. Tozer. Excitation energies in density functional theory: an evaluation and a diagnostic test. *J. Chem. Phys.*, 128:044118, 2008.

[223] Sundaram Arulmozhiraja and Michelle L. Coote. $^1\mathrm{L}_a$ and $^1\mathrm{L}_b$ states of indole and azaindole: is density functional theory inadequate? *J. Chem. Theory Comput.*, 8:575–584, 2012.

[224] Andriy V. Kityk. Absorption and fluorescence spectra of heterocyclic isomers from long-range-corrected density functional theory in polarizable continuum approach. *J. Phys. Chem. A*, 116:3048–3055, 2012.

[225] Ryan M. Richard and John M. Herbert. Time-dependent density-functional description of the $^1\mathrm{L}_a$ state in polycyclic aromatic hydrocarbons: charge-transfer character in disguise? *J. Chem. Theory Comput.*, 7:1296–1306, 2011.

[226] Gjergji Sini, John S. Sears, and Jean-Luc Brédas. Evaluating the performance of DFT functionals in assessing the interaction energy and ground-state charge-transfer of donor/acceptor complexes: tetrathiafulvalene-tetracyanoquinodimethane (TTF-TCNQ) as a model case. *J. Chem. Theory Comput.*, 7:602–609, 2011.

[227] P. Romaniello, D. Sangalli, J. A. Berger, F. Sottile, L. G. Molinari, L. Reining, and G. Onida. Double excitations in finite systems. *J. Chem. Phys.*, 130:044108, 2009.

[228] D. Hofmann, T. Körzdörfer, and S. Kümmel. Kohn-Sham self-interaction correction in real time. *Phys. Rev. Lett.*, 108:146401, 2012.

[229] Natalia Kuritz, Tamar Stein, Roi Baer, and Leeor Kronik. Charge-transfer-like $\pi \to \pi^*$ excitations in time-dependent density functional theory: a conundrum and its solution. *J. Chem. Theory Comput.*, 7:2408–2415, 2011.

[230] Peter Elliott, Sharma Goldson, Chris Canahui, and Neepa T. Maitra. Perspectives on double-excitations in TDDFT. *Chem. Phys.*, 391:110–119, 2011.

[231] Lloyd N. Trefethen and David Bau III. *Numerical Linear Algebra*. Society for Industrial and Applied Mathematics, Philadelphia, 1997.

[232] Youcef Saad and Martin H. Schultz. GMRES: a generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM J. Sci and Stat. Comput.*, 7:856–869, 1986.

[233] Yan Zhao, Benjamin J. Lynch, and Donald G. Truhlar. Multi-coefficient extrapolated density functional theory for thermochemistry and thermochemical kinetics. *Phys. Chem. Chem. Phys.*, 7:43–52, 2005.

[234] Ernest R. Davidson. First excited $^1\Sigma_g^+$ state of the hydrogen molecule. *J. Chem. Phys.*, 35:1189, 1961.

[235] Klaus-Peter Huber and Gerhard Herzberg. *Molecular Spectra and Molecular Structure. IV. constants of diatomic molecules*. Van Nostrand Reinhold Co., New York, 1979.

[236] C. Sannen, G. Raşeev, C. Galloy, G. Fauville, and J. C. Lorquet. Unimolecular decay paths of electronically excited species. II. the $C_2H_4^+$ ion. *J. Chem. Phys.*, 74:2402–2411, 1981.

[237] H. Köppel, L. S. Cederbaum, and W. Domcke. Strong nonadiabatic effects and conical intersections in molecular spectroscopy and unimolecular decay: $C_2H_2^+$. *J. Chem. Phys.*, 77:2014–2022, 1982.

[238] Kouichi Takeshita. A theoretical study on ionization of ethylene with analysis of vibrational structure of the photoelectron spectra. *J. Chem. Phys.*, 95:1838–1846, 1991.

[239] Horst Köppel, W. Domcke, L. S. Cederbaum, and W. von Niessen. Vibronic coupling effects in the photoelectron spectrum of ethylene. *J. Chem. Phys.*, 69:4252–4263, 1978.

[240] P. M. Dehmer and J. L. Dehmer. Photoelectron spectroscopy using a supersonic molecular beam source. the $^2B_{3u} + e^- \leftarrow {}^1A_g$ transition in ethylene. *J. Chem. Phys.*, 70:4574–4580, 1979.

[241] J. E. Pollard, D. J. Trevor, J. E. Reutt, Yuan T. Lee, and D. A Shirley. Torsional potential and intramolecular dynamics in the $C_2C_4^+$ photoelectron spectra. *J. Chem. Phys.*, 81:5302–5309, 1984.

[242] Nicholas C. Handy, Ross H. Nobes, and Hans-Joachim Werner. The ethylene radical cation: twisted or planar? *Chem. Phys. Lett.*, 110:459–463, 1984.

[243] Gerhard Herzberg. *Molecular Spectra and Molecular Structure. III. Electronic spectra and electronic structure of polyatomic molecules.* Van Nostrand Co., Princeton, NJ, 1967.