Developing a Model Chemistry for Multiconfiguration Pair-Density Functional Theory to Study Photochemistry and Molecular Interactions

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Jie Bao

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

PH.D. OF SCIENCE

Donald G. Truhlar

January, 2021

Acknowledgements

I would like to acknowledge my advisor, Dr. Donald Truhlar, for being a helpful and reliable advisor. I would like to acknowledge my collaborator, Dr. Laura Gagliardi, for her help in almost all my projects. I would like to acknowledge Dr. Junwei Lucas Bao, Dr. Shaohong Louis Li, Dr. Kamal Sharkas, Dr. Yinan Shu, Dr. Zoltán Varga, Dr. Chad Hoyer, and Dr. Pragya Verma for their help when I started work on theoretical chemistry. I would like to acknowledge Dr. Andrew Sand for his help in teaching me to code in OpenMolcas. I would like to thank Dr. Sijia Dong and Dr. Meagan Oakley for collaborations in two projects. I would like to thank Dr. Chen Zhou, Dr. Matthew Hermes and Thais Scott, for valuable and interesting discussions of some of my projects. At last, I would like to acknowledge Dr. Donald Truhlar, Dr. Laura Gagliardi, Dr. Jason Goodpaster, and Dr. Aaron Massari for serving on my committee for my thesis.

Dedication

To my family who give me support.

Abstract

Photochemical reaction, which starts by exciting a system into an electronically excited state, is ubiquitous, for example, in the atmosphere. This has made photochemical reactions a very interesting topic. Multiconfigurational pair-density functional theory (MC-PDFT) is a powerful and efficient method for studying photochemical processes. This method has proved very efficient compared with other wave function methods, such as multi-state complete active space second order perturbation theory (MS-CASPT2), especially for large systems. Successful as MC-PDFT is, there are some limitations that stop MC-PDFT from being applied to studying photochemistry problems. The first limitation is that, like other multireference methods, the performance of MC-PDFT depends on the quality of the reference wave function, which by convention is optimized by an active-space method, such as complete active space self-consistent field (CASSCF). The second limitation is that MC-PDFT is a single-state method that does not include state interaction between reference states. This means that MC-PDFT gives wrong topologies of potential energy surfaces, which are important in studying photochemical reactions. My work is focused on resolving these two limitations. We proposed the ABC scheme and the ABC2 scheme to automatically generate the active space that gives good-quality reference wave functions thus successfully reproducing vertical excitation energies obtained from experiments or high-level calculations. We proposed the extended multi-state PDFT (XMS-PDFT) and compressed-state multi-state PDFT (CMS-PDFT) as two options to introduce state-interaction in pair-density functional theory. Among two methods, XMS-PDFT is more efficient, while CMS-PDFT is more robust. Both methods proved successful in providing correct topologies of potential energy surfaces for a variety of systems.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Photochemistry plays an important role in the environment[1, 2, 3, 4, 5, 6, 7] and understanding the chemical processes in photochemistry is crucial. Kohn-Sham density functional theory[8] (KS-DFT) and time-dependent density functional theory (TD-DFT) with Tamm-Dancoff approximation and adiabatic approximation[9, 10, 11] are efficient methods to study ground states and excited states, respectively. However, these two density functional theories (DFT) with approximate density functionals are insufficient to study excited states in the following two cases. The first is that TD-DFT with adibatic approximation is insufficient to describe the doubly excited states,[12, 13, 14] and the second is the spin contamination that makes the excited states not eigenstates of the spin quantum number (S^2) in both KS-DFT and TD-DFT. Multireference (MR) wave function (WF) methods, such as complete active space second-order perturbation theory[15, 16, 17] (CASPT2), are free from these two problems. However, the application of the MR WF methods is limited to small systems such as azobenzene.[18]

Multiconfiguration pair-density functional theory[19, 20, 21] (MC-PDFT) is an MR method that combines WF methods and DFT. It is a more efficient method than CASPT2, especially for large systems.[22] Previous studies have shown that MC-PDFT is as accurate as CASPT2,[23] therefore MC-PDFT is a very promising method for studying photochemistry.

However, there are two limitations in applying MC-PDFT to photochemistry problems. The first is that the results of MC-PDFT depend on the active space of the reference wave function; thus one may spend a huge amount of time beforehand to find out the optimal active space. The second limitation is that MC-PDFT, as an MR method, does not include the state interaction as does multi-state (MS) CASPT2.

The themes of my work are to design automatic ways to generate active spaces and to introduce state interaction in MC-PDFT. The automatic active space selection procedures include ABC and ABC2 schemes, where the A, B and C are also three parameters in the scheme. The state interaction in MC-PDFT is introduced in two multi-state PDFT (MS-PDFT) methods. One is called extended MS-PDFT (XMS-PDFT) and the other is called compressed-state MS-PDFT (CMS-PDFT).

The thesis contains the following in Chapters 2 – 7.

- Chapter 2 studies the performance of MC-PDFT for predicting the adiabatic excitation energies and the equilibrium distances for cyano radical.
- Chapter 3 studies the performance of MC-PDFT for predicting the equilibrium distances and dissociation energies of weakly bonded alkaline earth metal dimers.
- Chapter 4 discusses an automatic scheme for selecting active spaces for MC-PDFT to compute vertical excitation energies.
- Chapter 5 proposes a new and more robust automatic scheme for selecting active space for MC-PDFT to compute vertical excitation energies.
- Chapter 6 introduces a method, called extended multi-state pair-density functional theory (XMS-PDFT), that efficiently introduces state interaction into pair-density functional theories.

- Chapter 7 talks about another multi-state PDFT that is as efficient as XMS-PDFT and is more robust than XMS-PDFT.

Chapter 2

Multiconfiguration pair-density functional theory for doublet excitation energies and excited state geometries: the excited states of CN

Reproduced from Ref. [24] with permission from the PCCP Owner Societies.

2.1 Introduction

The cyano radical, CN, is present in the spectra of the sun, planets, comets, and carbon stars,[25] and it participates in the formation of nitriles in the interstellar medium, as suggested by crossed molecular beam experiments.[26, 27] The first experimental study of the CN molecule dates back to a century ago, when a bright cyanogen spectrum was discovered.[28] Many experiments were subsequently carried out to study the spectrum of CN. For example, Ram *et al.*[29] studied the emission spectrum of the $B^2\Sigma^+$–$X^2\Sigma^+$ transition and improved some values of the spectroscopic constants. The spectroscopic constants for all the low-lying states of CN were collected by Herzberg *et al.*[30] Bradforth *et al.*[31] studied the photoelectron spectroscopy of several anions including CN^-, and their results showed a minor change in the ground-state geometry of CN. An interesting extension of the study on CN was carried out by Pieniazec *et al.*,[32] who modeled the spectroscopy of CN in

an aqueous environment.

Due to the availability of experimental data and the relative simplicity of CN as a diatomic molecule, CN can serve as a benchmark system for the prediction of doublet-to-doublet excitation energies. Since wave functions for open-shell systems tend to introduce spin contamination and are usually highly multiconfigurational[33], these transition energies are harder to predict than singlet-to-singlet transition energies by methods that remain affordable when one considers large systems where the most accurate methods for small systems become impractical.

Since experimental data are available for the equilibrium geometries of the excited states, we can study both vertical and adiabatic excitation energies, in contrast to polyatomic molecules, where benchmarks for adiabatic excitation energies are often unavailable. Guérin[34] carried out the first theoretical study of the $A^2\Pi_i$, $D^2\Pi_i$, $F^2\Delta$, and $J^2\Delta$ adiabatic excitations; he employed up to three configuration state functions (CSFs) per state. Schaefer and Heil[35] performed a full configuration interaction (CI) calculation using a small basis set, in which only one ζ is used for 2s and 2p orbitals of each atom; the errors in predicting adiabatic excitation energies with respect to experiment are about 0.4 eV. Das *et al.*[36] used multiconfiguration self-consistent field[37] (MCSCF) theory to study the spectrum of CN. Other workers[38, 39] used time-dependent linear-response Kohn-Sham density functional theory[9, 10, 11] (abbreviated TD-DFT) to calculate the vertical excitation energies. The most accurate computational results available so far are by Shi *et al.*[40] using multireference configuration interaction singles and doubles[41, 42] with the Davidson correction[43] (MR-CISD+Q) with the aug-cc-pV5Z[44] basis set and other basis sets, followed by extrapolation to the complete basis set (CBS) limit, including core correlation and scalar relativistic effects; this work has a very small mean unsigned error (MUE) of 0.01 eV when averaged over the first four adiabatic excitation energies (those for states $A^2\Pi_i$, $B^2\Sigma^+$, $D^2\Pi_i$, and $E^2\Sigma^+$). As shown by this example, for small systems, wave function methods are available that can

reproduce the experimental results very accurately. However, due to the cost of these wave function methods, such accurate calculations are usually too expensive for larger systems. Therefore we are interested in density functional methods because they can be applied to very large systems.

Methods based on a multiconfigurational reference wave function (such methods are called multireference (MR) methods) have advantages in treating open-shell systems since they can incorporate the correct spin and spatial symmetry in a natural way. Multiconfiguration pair-density functional theory[19, 20] (MC-PDFT) is a recently developed post-MCSCF method, which uses an MCSCF wave function to compute the kinetic energy, total density, and on-top density and then uses them to approximate the total electronic energy. The total electronic energy of MC-PDFT consists of the kinetic energy, electron-nuclear attraction, classical Coulomb repulsion, and the energy from a density functional called the on-top functional. Because of this computational prescription, MC-PDFT is much less expensive for large systems than other post-MCSCF methods such as MR-CISD and complete-active-space second-order perturbation theory[15](CASPT2). Studies carried out so far revealed that MC-PPDFT provides similar results to those from CASPT2[19, 45, 46, 47] for closed-shell and open-shell ground-state systems and excitation energies of closed-shell systems. MC-PDFT therefore provides a very promising method for a wide range of applications.

In order to evaluate its performance for excitation energies of open-shell systems, we here report MC-PDFT calculations for adiabatic excitation energies of CN. In particular, we calculate the adiabatic excitation energies for the first four excited states and equilibrium distances for the first five states of CN using MC-PDFT and several other methods. The methods are divided into two categories, namely single-reference (SR) methods and the multireference (MR) methods. We compare the performance of the SR methods and the MR methods for reproducing the experimental results of ref. 6. We also evaluate the advantage of MC-PDFT over the SR methods, and point out some drawbacks of this MR density

functional method (at least in its current state of development) as compared to MR wave function methods.

2.2 Computational Details

We performed SR excited-state calculations by three kinds of methods: TD-DFT, equation-of-motion coupled-cluster theory with singles and doubles for electron excitation[48, 49] (EOM-EE-CCSD) and the completely renormalized EOM-CCSD with noniterative triples[50, 51] [CR-EOM-CCSD(T)], and equation-of-motion coupled-cluster theory for ionization potentials with single and double excitations (EOM-IP-CCSD).[52, 53]

The TD-DFT calculations use unrestricted Kohn-Sham DFT for the ground states whose response is calculated to obtain the excitation energies; the coupled cluster calculations use unrestricted Hartree-Fock[54] (UHF) wave functions for the reference states, and CCSD or CR-CCSD(T)[55, 56, 57] for the ground state. The basis set for SR calculations is maug-cc-pVTZ (to be abbreviated mVTZ).[58]

The TD-DFT calculations are performed in Gaussian 09[59], and the functionals include two generalized gradient approximation (GGA) exchange-correlation (xc) functionals, BLYP[60, 61] and PBE[62], three meta-gradient-approximation functionals, M06-L[63, 64], MN11-L[65] and MN15-L[66], and three hybrid meta-gradient-approximation functionals, M06[64, 67], MN11[68] and MN15.[69]

The EOM-EE-CCSD and the CR-EOM-CCSD(T) calculations are performed using the NWChem program.[70] The EOM-IP-CCSD calculations are performed using the GAMESS program.[71, 72]

We also performed MR calculations. The reference wave functions for these are state-averaged complete active space self-consistent-field theory[73] (CASSCF) wave functions with 7 active electrons in 7 active orbitals. The basis set is mVTZ.[58] In addition, the

cc-pVTZ basis set (to be abbreviated VTZ)[44] is used to understand the effect of omitting diffuse functions on the performance of some of the MR methods. The occupancies of the active orbitals in the dominant configurations of each of the states are shown in Table 2.1. The CASSCF calculations were performed with Molcas 8.1.[74]

Table 2.1: The Dominant Configuration for the Ground State and the Orbitals with a Changed Number of Electrons in the Dominant Configurations for Excited States at r_e

State	Configuration
$X^2\Sigma^+$	$(1\sigma)^2(2\sigma)^2(3\sigma)^2(4\sigma)^2(1\pi)^4(5\sigma)^1$
$A^2\Pi_i$	same except $(1\pi)^3(5\sigma)^2$
$B^2\Sigma^+$	same except $(4\sigma)^1(5\sigma)^2$
$D^2\Pi_i$	same except $(5\sigma)^0(2\pi)^1$
$E^2\Sigma^+$	same except $(1\pi)^3(5\sigma)^1$

We performed four kinds of MR calculations: MC-PDFT; multi-state complete active space second-order perturbation theory[75] (MS-CASPT2); MR-CISD; and MR-CISD+Q. The MS-CASPT2 and MC-PDFT calculations were performed using Molcas 8.1.[74] The MR-CISD and MR-CISD+Q calculations were performed using Molpro.[76, 77]

We performed three variations on the CASSCF and post-CASSCF calculations, and these are indicated by suffixes -7m, -7, and -10m, which are explained in Table 2.2. In the text, when no suffix is used, we are referring to the method in general.

Table 2.2: Notations for Three Levels of CASSCF and MR Calculations

Suffix	States averaged[a]	Basis set
-7m	7	maug-cc-pVTZ
-7	7	cc-pVTZ
-10m	10	maug-cc-pVTZ

[a]The number of states in the effective Hamiltonian of the multi-state CASPT2 calculations is the same as the number of states average in the CASSCF calculations

We used six functionals for the MC-PDFT calculations, namely translated24 (prefix t)[19] and fully translated[78] (prefix ft) versions of BLYP, PBE, and revPBE.[79] For labeling the MC-PDFT calculations, we only specify which functional we use without saying that is an MC-PDFT calculation. For example, ftPBE is an MC-PDFT calculation using fully translated PBE functional. All MC-PDFT calculations in the present article are based on CASSCF reference functions without restrictions.

To calculate the adiabatic excitation energy of CN, we first calculated the equilibrium value (r_e) of the internuclear distance (r) for each excited state. This was done by scanning r and fitting the three lowest-energy points to a quadratic function. Subtracting the energy minimum of the ground state from that of the excited state gives the adiabatic excitation energy (T_e) of that state.

2.3 Results and Discussion

2.3.1 An Illustration of the Dominant Electron Configuration for Each State of CN

At the equilibrium geometry (r_e =1.1718 $\mathring{A}$) of the ground state, the dominant configuration state of each of the CASSCF of the states is shown in Table 2.1. The ground state is the X state, with a singly occupied σ orbital. The A, B, D, and E states are obtained by single excitations, as indicated in Table 2.1.

2.3.2 Multireference Diagnostic for CN

The M diagnostic[80] is used to evaluate the multiconfigurational character of a state. The M diagnostic is defined as:

$$M_i = \frac{1}{2}[2 - n_i(\mathrm{MCDONO}) + \sum_{j}^{n_{j\,\mathrm{SOMO}}} |n_i(j) - 1| + n_i(\mathrm{MCUNO})\,] \qquad (2.1)$$

where n_i(MCDONO) is the natural occupation number (NON) of the most correlated doubly-occupied orbital in the dominant configuration for the state *i*; n_i(MCUNO) is the NON of the most correlated unoccupied orbital in the dominant configuration; and $n_i(j)$ is the NON of singly-occupied orbital (SOMO) *j*. The multiconfigurational character of a system is classified as small ($M_i < 0.05$), modest ($0.05 \leqslant M_i \leqslant 0.10$) or large ($M_i > 0.10$), depending on M_i. For example, for the CN ground state at the equilibrium distance, the NONs for orbitals from 6 to 8 are 1.92, 1.03 and 0.09, respectively. The occupations of these orbitals in the dominant configuration are $(6)^2(7)^1(8)^0$; therefore the most correlated doubly-occupied orbital is orbital 6, with n_1(MCDONO) being 1.92; correspondingly, n_1(MCUNO) is 0.09, and $n_1(1)$ is 1.03.

The results for the M_i for the five states calculated by CASSCF-7m at the equilibrium distances of the corresponding state are listed in Table 2.3. The results show that except for the first excited state, the A state, which has a modest multiconfigurational character, the other four states investigated here have a large multiconfigurational character, and the higher-energy states have more *m* character.

Table 2.3: M Diagnostic, Dominant Configurations and the Corresponding Weight of the Configuration in the Wave Function for Each State at the Experimental Equilibrium

State	Distance[a]	M_i	Dominant configuration(s)[b]
$X^2\Sigma^+$	1.1718	0.10	222u000 (85%)
$A^2\Pi_i$	1.2333	0.07	22u2000 (93%)/2u22000 (93%)
$B^2\Sigma^+$	1.150	0.18	u222000 (78%)
$D^2\Pi_i$	1.498	0.63	2220u00 (76%)/22200u0 (76%)
$E^2\Sigma^+$	1.3245	0.91	2u2d0u0 (28%) + 22udu000 (28%)

[a] Distances are in $\mathring{\mathrm{A}}$, taken from ref. [32].
[b] A slash sign separates the dominant configurations of two degenerate states in Π states, and a plus sign means the wave function for a state has configurations with similar weights. u and d means alpha electron and beta electron, respectively.

2.3.3 Adiabatic Excitation Energy for Each of the States

The adiabatic excitation energies (T_e) by the methods studied in this chapter and the experimental values (dashed horizontal lines) are shown in Fig. 2.1. The signed errors and MUEs are shown in Table 2.4. For the first excited state and the second excited state, the SR and the MR methods perform similarly in reproducing the experimental values. For the third and the fourth excited states, SR methods do not perform as well as MR methods. The SR methods systematically underestimate the adiabatic excitation energy for the E state, and they give values that are much lower than the experimental value. This systematic error is also observed in the previous literature in that SR methods underestimate the energy for the E state and MR methods have a correct order of energies; for example, in the TD-DFT calculation of vertical excitations by Guan *et al.*,[38] the third $^2\Sigma^+$ state is lower than the second $^2\Pi_i$ state (see Fig. 6 in ref. [38]); and in the MCSCF calculation by Das *et al.*[36], the E state lies above the D state (see Fig. 1 in ref.[36]).

For the first excited state, the best result is given by tPBE-7. Following that are MR-CISD-7 and tPBE-7m, which have absolute errors of 0.01 eV. In general, the most accurate results for the first adiabatic excitation energy are given by MR-CISD calculations and tPBE. Interestingly, the tPBE and ftPBE calculations are much more accurate than MS-CASPT2. The two best SR methods for the first excited state are TD-BLYP and CR-EOM-CCSD(T), with absolute errors of 0.06 and 0.07 eV, respectively.

For the second excited state, the best result comes from TD-BLYP-7m, followed by TD-M06-L and then tPBE-10m; roughly speaking, both SR and MR methods perform similarly for this state.

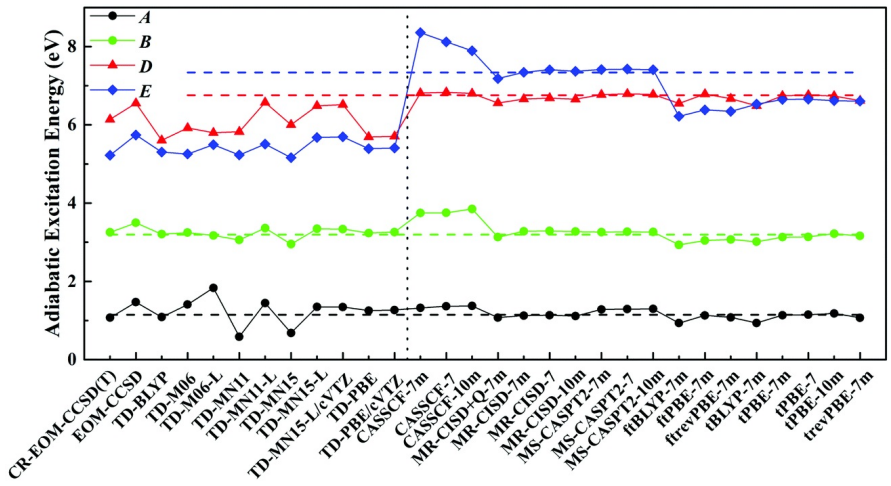

Figure 2.1: Adiabatic excitation energies (eV) by different methods. The horizontal dashes are the experimental adiabatic excitation energies. On the left-hand side of vertical dash line are the SR methods and on the right-hand side are the MR methods. J. J. Bao, L. Gagliardi, and D. G. Truhlar. Multiconfiguration pair-density functionaltheory for doublet excitation energies and excited state geometries: the excited statesof CN. Phys. Chem. Chem. Phys., 19:30089–30096, 2017. Reproduced by permission of the PCCP Owner Societies

Table 2.4: Signed Errors of Adiabatic Excitation Energies (in eV), the Mean Unsigned Error (MUE) Over First *n* Excitation Energies (MUE_*n*), and the Rank of MUE_*n* (rank_*n*)

	$A^2\Pi_i$	$B^2\Sigma^+$	$D^2\Pi_i$	$E^2\Sigma^+$	MUE_4	Rank_4	MUE_3	Rank_3
Single reference methods								
CR-EOM-CCSD(T)	-0.07	0.06	-0.62	-2.11	0.71	22	0.25	18
EOM-EE-CCSD	0.32	0.30	-0.20	-1.60	0.60	20	0.27	20
EOM-IP-CCSD	0.02	0.03	na[a]	na	na	na	na	na
TD-BLYP	-0.06	0.01	-1.15	-2.03	0.81	26	0.41	25
TD-M06	0.26	0.06	-0.83	-2.08	0.81	25	0.38	23
TD-M06-L	0.69	-0.02	-0.96	-1.84	0.88	27	0.55	29
TD-M11	-0.56	-0.14	-0.93	-2.11	0.93	29	0.54	28
TD-M11-L	0.30	0.17	-0.19	-1.82	0.62	21	0.22	15
TD-MN15	-0.46	-0.24	-0.76	-2.17	0.91	28	0.49	27
TD-MN15-L	0.20	0.15	-0.27	-1.66	0.57	19	0.21	14
TD-MN15-L/cVTZ	0.20	0.14	-0.24	-1.64	0.56	18	0.19	13
TD-PBE	0.11	0.04	-1.07	-1.95	0.79	23	0.41	24
TD-PBE/cVTZ	0.12	0.07	-1.05	-1.93	0.79	24	0.41	26
Multireference methods								
CASSCF-7m	0.18	0.56	0.06	1.02	0.45	17	0.26	19
CASSCF-7	0.22	0.56	0.08	0.79	0.41	15	0.28	21
CASSCF-10m	0.23	0.66	0.05	0.56	0.37	14	0.31	22
MRCISD+Q-7m	-0.07	-0.06	-0.20	-0.15	0.12	7	0.11	12
MRCISD-7m	-0.02	0.09	-0.10	0.01	0.05	1	0.07	6
MRCISD-7	-0.01	0.10	-0.07	0.07	0.06	3	0.06	4
MRCISD-10m	-0.03	0.08	-0.10	0.03	0.06	2	0.07	7
MS-CASPT2-7m	0.14	0.07	0.01	0.08	0.07	4	0.07	8
MS-CASPT2-7	0.15	0.07	0.04	0.09	0.09	6	0.09	10
MS-CASPT2-10m	0.15	0.07	0.02	0.07	0.08	5	0.08	9
ftBLYP-7m	-0.21	-0.26	-0.21	-1.12	0.45	16	0.23	17
ftPBE-7m	-0.02	-0.15	0.03	-0.95	0.29	11	0.07	5
ftrevPBE-7m	-0.07	-0.12	-0.09	-0.99	0.32	12	0.09	11
tBLYP-7m	-0.21	-0.18	-0.27	-0.81	0.37	13	0.22	16
tPBE-7m	-0.01	-0.06	-0.02	-0.69	0.19	9	0.03	3
tPBE-7	0.00	-0.05	0.00	-0.68	0.18	8	0.02	1
tPBE-10m	0.03	0.03	-0.02	-0.72	0.20	10	0.03	2
Experimental values								
	1.146	3.193	6.756	7.334				

[a] na denotes not available.

The third excited state (the D state) is where the SR methods start to become much worse than the MR methods. Among the SR methods, only the two local meta-GGAs, MN11-L and MN15-L, and the EOM-EE-CCSD calculations have an error whose magnitude is 0.24 eV or less. The magnitude of the error is in the range 0.62–1.15 eV for the rest of the SR methods, with the largest unsigned errors (1.05–1.15 eV) for the GGAs. It is very striking that the poor performance of TD-DFT using GGAs can be improved by using MC-PDFT with the translated versions of the same GGAs. For example, the error by TD-PBE is -1.07 eV. In the tPBE-7m and the ftPBE-7m calculations, the error for this state is -0.02 eV and 0.03 eV, respectively.

For the fourth excited state (the E state), MR-CISD-7m has an error of 0.01 eV, i.e., it almost perfectly reproduces the experimental result. Following this in accuracy are the MR-CISD calculations using the other two basis sets and then the MS-CASPT2 calculations. The MR-CISD+Q-7m calculation, compared to that from the MR-CISD-7m, overcorrects the adiabatic excitation energy and ranks after the MS-CASPT2 calculations. The tPBE-7m calculation has an error around -0.7 eV, which is not as good as that by the expensive wave function theories, but is a significant improvement over the SR methods. A possible reason that the SR methods to perform similarly to the MR methods on the first two excitation energies, but not the higher-energy states is that the multiconfigurational character for the first three states, although indicated as large by the M diagnostic, can still apparently be properly described by the SR methods. For the higher states, though, the multiconfigurational character is so strong (huge M diagnostics of 0.63–0.91) that the SR methods fail to give a correct description of those states.

When the MUE averaged over the first four excited adiabatic excitation energies [MUE(4) as tabulated in Table 2.4] is considered, the MR-CISD calculations are the best among all the calculations. Following in accuracy are the MS-CASPT2 calculations, and then MR-CISD+Q. The tPBE and the trevPBE calculations rank after these MR wave function calculations, but

are still better than the SA-CASSCF calculations. The SR methods are overall worse than the MR methods. When MUE(3) is considered, i.e., only the three lowest excited states are considered, tPBE-7 is the best among all the calculations, with an mean unsigned error of only 0.02 eV, which is even better than the two MR wave function methods. The second and the third best results come from the other two tPBE calculations, with an error of 0.03 eV.

2.3.4 Equilibrium Distances

The equilibrium distances for the ground and excited states are shown graphically in Fig. 2.2, and the signed errors are recorded in Table 2.5 and the MUEs are recorded in Table 2.6. For the ground state (pink in the figure), the A state (black), and the B state (green), some SR methods are able to give results as good as those by the MR methods. However, for the D state (red) and the E state (blue), SR methods are significantly worse than MR methods.

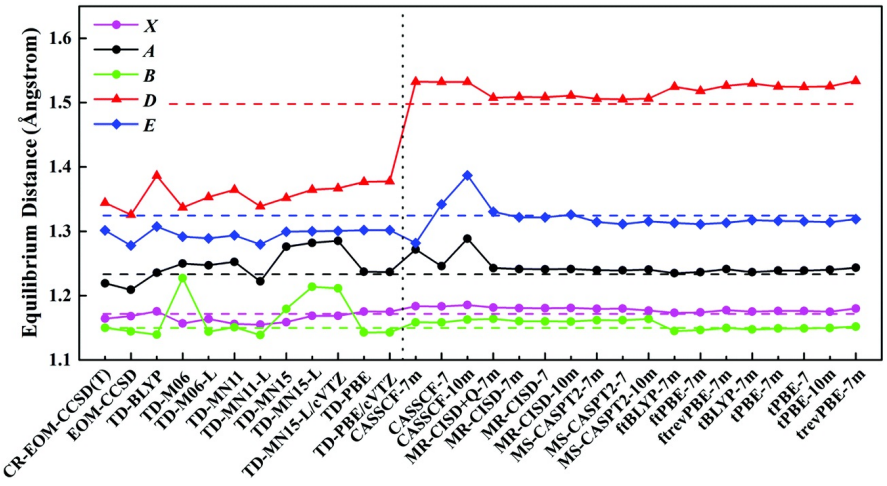

Figure 2.2: The equilibrium distance ($\mathring{\mathrm{A}}$) for each state by various methods. Single-reference methods are on the left of the vertical dash line, and multireference methods are on the right. Note that the order of equilibrium distances is not in the order of energies for these states. J. J. Bao, L. Gagliardi, and D. G. Truhlar. Multiconfiguration pair-density functionaltheory for doublet excitation energies and excited state geometries: the excited statesof CN. Phys. Chem. Chem. Phys., 19:30089–30096, 2017. Reproduced by permission of the PCCP Owner Societies

For the ground state and the first excited state, BLYP, PBE, and the MC-PDFT methods – both translated and fully translated – give only small errors in geometries. Taking the ground state as an example, the best C-N distance is given by ftBLYP-7m, which has an unsigned error of 0.001 $\mathring{\mathrm{A}}$. The MC-PDFT calculations in general have unsigned errors ranging from 0.001 to 0.008 $\mathring{\mathrm{A}}$. The calculations using PBE, BLYP, and MN15-L and the CCSD calculation have unsigned errors that are less than 0.005 $\mathring{\mathrm{A}}$. For the B state, the BLYP, PBE, M11, and M06-L functionals and CR-EOM-CCSD(T) yield results that are comparable to and even better than most MR methods, for example, MR-CISD-7m. Even though PBE and BLYP have small errors, MC-PDFT calculations further improve the results. For example, the magnitudes of the errors of tPBE-7m, tPBE-7 and tPBE-10m are all less than or equal to 0.001 and 0.000 $\mathring{\mathrm{A}}$, respectively, compared to that of TD-PBE, which is -0.007 $\mathring{\mathrm{A}}$.

Due to relatively small M value for the first three states, the SR methods perform similarly to the MR methods in both adiabatic excitation energies and equilibrium distances.

For the fourth state (the *D* state), the results from BLYP and PBE have errors that are about 10 times larger than for the lower states. The errors are -0.112 and -0.121 $\mathring{\mathrm{A}}$, respectively. Despite the fact that the MC-PDFT errors are also larger for this state, the error is still small compared to the errors of those SR methods. For example, the errors are only 0.032 $\mathring{\mathrm{A}}$ for tBLYP-7m and 0.027 $\mathring{\mathrm{A}}$ for tPBE-7m.

Although the SR methods systematically misestimate the adiabatic excitation energy for the E state, their results for the equilibrium distance are better, although eleven of the 12 unsigned errors for the SR methods for the E state are greater than or equal to 0.023 $\mathring{\mathrm{A}}$, with the one smaller error being -0.017 $\mathring{\mathrm{A}}$. The MC-PDFT calculations also have systematic errors for the geometry of this state, but the errors are much smaller than for the SR methods and are comparable to those from the MR wave-function methods. For example, the errors of MR-CISD-7m, MS-CASPT2-7m and tPBE-7m are -0.003, -0.010 and -0.008 $\mathring{\mathrm{A}}$, respectively.

Table 2.5: Signed Errors of Equilibrium Distances (r_e in $\mathring{A}$)

	$X^2\Sigma^+$	$A^2\Pi_i$	$B^2\Sigma^+$	$D^2\Pi_i$	$E^2\Sigma^+$
CR-EOM-CCSD(T)	-0.007[a]	-0.014	0.000	-0.154	-0.023
EOM-EE-CCSD	-0.004[a]	-0.024	-0.006	-0.170	-0.046
EOM-IP-CCSD	-0.004	-0.006	-0.006	na[b]	na
TD-BLYP	0.004	0.002	-0.011	-0.112	-0.017
TD-M06	-0.015	0.017	0.077	-0.161	-0.033
TD-M06-L	-0.008	0.014	-0.006	-0.145	-0.036
TD-M11	-0.016	0.019	0.001	-0.133	-0.031
TD-M11-L	-0.017	-0.011	-0.011	-0.159	-0.045
TD-MN15	-0.013	0.043	0.029	-0.146	-0.025
TD-MN15-L	-0.003	0.049	0.064	-0.133	-0.024
TD-MN15-L/cVTZ	-0.003	0.052	-0.061	0.131	-0.024
TD-PBE	0.003	0.004	-0.007	-0.121	-0.023
TD-PBE/cVTZ	0.003	0.003	-0.007	-0.120	-0.023
Single reference methods					
CASSCF-7m	0.012	0.038	0.009	0.034	-0.043
CASSCF-7	0.012	0.012	0.008	0.034	0.017
CASSCF-10m	0.014	0.055	0.013	0.034	0.063
MR-CISD+Q-7m	0.010	0.010	0.014	0.010	0.006
MRCISD-7m	0.009	0.008	0.010	0.011	-0.003
MRCISD-7	0.009	0.008	0.010	0.010	-0.003
MRCISD-10m	0.009	0.008	0.010	0.013	0.001
MS-CASPT2-7m	0.008	0.006	0.012	0.008	-0.010
MS-CASPT2-7	0.008	0.006	0.012	0.007	-0.013
MS-CASPT2-10m	0.005	0.007	0.014	0.008	-0.009
ftBLYP-7m	0.001	0.002	-0.005	0.027	-0.012
ftPBE-7m	0.002	0.003	-0.003	0.020	-0.013
ftrevPBE-7m	0.006	0.008	0.000	0.028	-0.011
tBLYP-7m	0.003	0.003	-0.003	0.032	-0.007
tPBE-7m	0.004	0.006	-0.001	0.027	-0.008
tPBE-7	0.004	0.006	-0.001	0.026	-0.009
tPBE-10m	0.003	0.007	0.000	0.027	-0.010
Experimental values					
	1.172	1.233	1.150	1.498	1.325

[a]Note that the ground state in EOM-EE-CCSD and CR-EOM-CCSD(T) calculations are CCSD and CR-CCSD(T), respectively.
[b]na denotes not available.

Table 2.6: MUE of r_{e} over First *n* States (MUE_*n*) and the Rank of MUE_*n* (Rank_*n*)

	MUE_5	Rank_5	MUE_4	Rank_4	MUE_3	Rank_3
Single reference methods						
CR-EOM-CCSD(T)	0.040	21	0.044	23	0.007	12
EOM-EE-CCSD	0.050	25	0.051	25	0.011	21
EOM-IP-CCSD	na[a]	na	na	na	0.005	10
TD-BLYP	0.029	17	0.032	18	0.006	11
TD-M06	0.061	29	0.067	29	0.036	28
TD-M06-L	0.042	23	0.043	22	0.009	19
TD-M11	0.040	22	0.042	21	0.012	23
TD-M11-L	0.049	24	0.050	24	0.013	24
TD-MN15	0.051	26	0.058	26	0.028	27
TD-MN15-L	0.055	28	0.062	28	0.039	29
TD-MN15-L/cVTZ	0.054	27	0.062	27	0.039	30
TD-PBE	0.032	19	0.034	20	0.005	9
TD-PBE/cVTZ	0.031	18	0.034	19	0.005	8
Multireference methods						
CASSCF-7m	0.027	16	0.023	16	0.020	25
CASSCF-7	0.017	15	0.017	15	0.011	20
CASSCF-10m	0.036	20	0.029	17	0.027	26
MR-CISD+Q-7m	0.010	13	0.011	14	0.011	22
MRCISD-7m	0.008	2	0.009	9	0.009	18
MRCISD-7	0.008	1	0.009	6	0.009	16
MRCISD-10m	0.008	3	0.010	11	0.009	17
MS-CASPT2-7m	0.009	6	0.009	3	0.009	14
MS-CASPT2-7	0.009	8	0.008	2	0.009	13
MS-CASPT2-10m	0.009	5	0.009	4	0.009	15
ftBLYP-7m	0.009	7	0.009	5	0.003	1
ftPBE-7m	0.008	4	0.007	1	0.003	2
ftrevPBE-7m	0.011	14	0.010	13	0.005	7
tBLYP-7m	0.010	11	0.010	12	0.003	3
tPBE-7m	0.009	10	0.010	10	0.004	6
tPBE-7	0.009	9	0.009	7	0.004	5
tPBE-10m	0.010	12	0.009	8	0.003	4

[a]na denotes not available.

Averaging the errors over the first three states gives MUE(3) in Table 2.6, and the ranks based on MUE(3) shows that ftPBLYP-7m has the very best performance of all methods tested over the first three states; the second, third and fourth ranked are the three tPBE calculations. The SR methods also show a strong ability of describing the first three excited states with a MUE of 0.005 or 0.006 $\mathring{\mathrm{A}}$, which is even better than the MR-CISD and the MS-CASPT2 calculations. The MUE(4) shows that the SR methods are not performing as well as the MR methods, and also that the errors for MC-PDFT calculations become bigger and have similar MUEs to those from the MR-CISD and the MS-CASPT2. When MUE(5) is considered, the MR-CISD and the MS-CASPT2 calculations are the best methods to describe the bond length. However, the ftPBE-7m still gives an error that is the same to the MR-CISD calculations to within 0.001 $\mathring{\mathrm{A}}$.

Our results show that the effect of adding or not including diffuse basis functions is generally small for CN, and so basis set effects will not be discussed in detail.

2.4 Conclusions

In this chapter, the adiabatic excitation energies (T_{e}) and the equilibrium distances (r_{e}) for the five states (ground doublet state and first four excited doublet states) are studied using methods that include

- SR methods

  the EOM-EE-CCSD wave function method

  the CR-EOM-CCSD(T) wave function method

  TD-DFT with eight local exchange-correlation density functionals

- MR wave function methods

CASSCF

MS-CASPT2

MR-CISD

MR-CISD+Q

- an MR density functional method

  MC-PDFT with five on-top density functionals.

The SR calculations give reasonably good adiabatic excitation energies (T_e) and equilibrium distances (r_e) for the first two excited states, and in some cases even give better results than MR-CISD and MS-CASPT2 calculations for these states. However, the SR methods fail to give reliable predictions of (T_e) and (r_e) for the third and fourth excited state. Thus it becomes necessary to use MR methods to study the more highly excited states.

The MR-CISD shows the most accurate results among all the calculations when the unsigned errors are averaged over all four Te values, followed by MS-CASPT2. The MC-PDFT calculations are not as accurate as the two MR wave function methods in this average over four excitation energies, but they still yield reasonably accurate re for all four states and reasonable accurate Te for the first three excited states. The excitation energies are significantly improved compared with the SR methods, although the error for the E state is still large in absolute terms, which we interpret as the difficulty of the calculating the correlation energy for a state with such a very large multireference diagnostic (M = 0.91).

The MC-PDFT method has some especially notable successes. For the first excited state, the most accurate excitation energy by any of the methods studied is given by tPBE. For the second excited state, tPBE is more accurate than MS-CASPT2, MR-CISD, CASSCF, EOM-EE-CCSD, and TD-PBE. For the third excited state, the unsigned errors for the tPBE and ftPBE are 0.02–0.03 eV, as compared to 1.07 eV for TD-PBE and 0.62 eV for CR-EOM-CCSD(T). For equilibrium distances averaged over the ground and first two excited states,

all five MC-PDFT methods give more accurate results than any of the other methods studied, whether SR or MR. For equilibrium distances averaged over the ground and first three excited states, ftPBE has the best accuracy of all methods studied. For equilibrium distances averaged over the ground and first four excited states, all five MC-PDFT methods give more accurate results than any of the SR methods, even CR-EOM-CCSD(T). These results are very encouraging for practical applications of MC-PDFT to the electronic spectroscopy of open-shell systems.

Chapter 3

Weak Interactions in Alkaline Earth Metal Dimers by Pair-Density Functional Theory

3.1 Introduction

Alkaline earth atoms and their dimers are widely used as trapped ultracold atoms and molecules in their singlet ground states.[81, 82, 83] The binding energy of Be_2 is 2.7 kcal/mol.[84] In comparison, Li, which is its neighbor in the periodic table, has a dimer binding energy of 20.4 kcal/mol,[85] which is 8 times larger, and Ne_2 (selected for comparison because Ne is the noble gas in the same row of the periodic table) has a binding energy of 0.08 kcal/mol (32 times smaller).[86] We see that the alkaline earth dimers occupy a middle ground between noble gas dimers and typical covalent bonds. The binding in alkaline earth dimers and clusters has been widely studied and is especially fascinating because of their closed-subshell electronic structure.[87, 88, 89, 90, 91, 92, 93, 94, 95]

An especially interesting property of Be_2 is the 2s–2p hybridization.[96] Many studies

of alkaline earth dimers have been carried out, especially for Be_2, which in the earliest experimental[97, 98] and theoretical[99] investigations was believed to be unstable. Even when later theoretical calculations showed a weak bond for Be_2, controversy remained about whether the equilibrium distance was about 2.5[100, 101, 102, 103, 104]or 4.5 $\mathring{\mathrm{A}}$.[105, 106] Experiments have now shown that the equilibrium distance is 2.44–2.45 $\mathring{\mathrm{A}}$, with a well depth of about 2.66 kcal/mol.[84, 107, 108] Theoretical studies with various wave function methods showed that very high levels such as full configuration interaction (FCI) calculations with large basis sets are required to calculate the binding energy correctly.[109, 96, 110, 111]

Of the many studies carried out for heavier alkaline earth dimers,[112, 113, 114, 115, 116, 117, 118, 119, 120, 121, 122, 123, 124] we emphasize the calculations done by coupled-cluster theory with single and double excitations and quasiperturbative connected triple excitations, [CCSD(T)],[125] with a perturbative treatment of relativistic effects[126, 127, 128] and the basis set extrapolated to the complete basis set (CBS) limit.[129] Reference [129] agrees with experimental values very well for Ca_2 and Sr_2, differing in experiments by no more than 0.06 kcal/mol, and therefore, these results are used as reference values for Ca_2, Sr_2, and Ba_2 in the present work. Radium is the heaviest alkaline earth metal, but there is a lack of experimental data, and we found only two theoretical papers that provide estimated thermodynamic data for Ra_2.[130, 131]

In the earliest work, Kohn–Sham density functional theory[8] (KS-DFT) with local exchange–correlation functionals was hard pressed to accurately describe these systems, in part because many local functionals (for example, generalized gradient approximations (GGAs)) do not usually describe van der Waals (vdW) interactions accurately.[132, 133, 134] Progress in obtaining better energies for weak interactions by KS-DFT has mainly involved four kinds of strategies: combining KS-DFT with damped dispersion by molecular mechanics,[135, 136] introducing nonlocal correlation into the exchange–correlation density functional,[137, 138, 139, 140, 141, 142] using parametrized functionals containing kinetic

energy density,[143] and using range separation to combine KS-DFT for short interelectronic separations with wave function theory for long interelectronic separations.[144]

Multiconfiguration pair-density functional theory[19] (MC-PDFT) uses on-top density functionals; in current practice these are translated from existing GGA exchange– correlation functionals, e.g., tPBE is a translation of the PBE[62] exchange–correlation functional, and ftPBE,[78] which is a full translation of the PBE functional. MC-PDFT calculations have usually been performed using the kinetic energy, density, and on-top density from complete active space self-consistent field[145] (CASSCF) calculations, in which case they may be called CAS-PDFT; reviews are available.[20, 21] Here we study whether CASSCF with translated GGAs can describe weak bonding well, even though the untranslated functional does not perform very well in KS-DFT and even though the density functional we use does not involve any of the four strategies mentioned in the previous paragraph.

3.2 Computational Details

Dissociation energies (D_e) and ground-state equilibrium bond distances (r_e) of Be_2, Mg_2, Ca_2, Sr_2, Ba_2, and Ra_2 were calculated by the two multireference (MR) methods: (i) CAS-PDFT with either the tPBE or the ftPBE on-top functional and (ii) complete active space second-order perturbation theory[15, 16, 17] (CASPT2).

We used three basis sets that are designed for calculations that include scalar relativistic effects, ANO-RCC, ANO-RCC-VQZP, and ANO-RCC-VTZP,[146] where ANO-RCC is a very large contracted basis to be denoted AR, ANO-RCC-VQZP is contracted down to valence quadruple-ζ plus polarization and will be denoted ARQ, and ANO-RCC-VTZP is contracted down to valence triple-ζ plus polarization and will be denoted ART. The actual number of primitive and contracted basis functions in each of these basis sets for each of the dimers is given in the SI in Tables S20–S23. The SI also considers nonelativistic calculations with

nonrelativistic basis sets.

The CAS-PDFT and CASPT2 calculations are based on the same CASSCF wave functions. We examined the use of two popular choices for the active space. The first is the widely used full valence (FV) active space, in which the active electrons are all of the valence electrons and the active orbitals are those formed from the valence orbitals of all of the atoms; for alkaline earth dimers, this yields four active electrons in eight active orbitals ($\sigma_{n\mathrm{s}}$, $\sigma^*_{n\mathrm{s}}$, σ_{pz}, σ^*_{pz}, π_{px}, π^*_{px},π_{py} and π^*_{py} formed primarily by the valence s and p orbitals of the two centers). The second active space, which was tested for only the largest AR basis set, is based on the correlated participating orbitals (CPO) scheme,[47, 147, 80] which for alkaline earth dimers involves all occupied bonding and antibonding orbitals plus a correlating orbital for each, and therefore, there are four active electrons in four active orbitals, which are the bonding and antibonding orbitals and a correlating orbital for each, namely, ($\sigma_{n\mathrm{s}}$, $\sigma^*_{n\mathrm{s}}$, σ_{pz} and σ^*_{pz} (in previous work, we distinguished three levels of the CPO scheme, but for alkaline earth dimers, they are all the same because there are no occupied nonbonding valence orbitals).

The CASPT2 calculations use a default ionization energy–electron affinity (IPEA) shift of 0.25 au[148] and an imaginary shift of 0.2 au.[149]

We also performed unrestricted KS-DFT with the PBE functional, which is the Kohn–Sham exchange–correlation functional corresponding to the tPBE and ftPBE on-top functionals of MC-PDFT.

A second-order Douglas–Kroll–Hess[150, 151] (DKH) relativistic Hamiltonian is used for all of the calculations to account for scalar relativistic effects. All of the calculations were done in Molcas 8.1.[74]

In some of the following discussion, we use a simplified notation: MC-PDFT with the ftPBE and tPBE on-top functionals are simply called ftPBE and tPBE, respectively. Additionally, we use a prefix to indicate an active space, for example, CPO-tPBE and FV-CASPT2.

The KS-DFT calculations with the PBE exchange–correlation functional are called KS-PBE.

3.3 Results and Discussion

3.3.1 Multiconfigurational Character

One of the challenges in these alkaline earth dimers is that the multiconfigurational nature of the system is important for describing the character of the bonding. The M diagnostic[80] indicates how strongly multiconfigurational a system is. Details of how to calculate M diagnostics can be found in ref [80], but a simplified description is that M measures the deviations of frontier natural orbital occupation numbers from integers. The M diagnostics for the alkaline earth metal dimers at their experimental equilibrium distances can be found in Table 3.1, and the results show that they are all strongly multiconfigurational systems ($M >$ 0.10). Although all six dimers are strongly correlated, Be_2 is the most strongly correlated, and Mg_2 and Ra_2 are least strongly correlated.

Table 3.1: M Diagnostics and Dominant Configurations at the Equilibrium Geometry As Calculated by FV-CASSCF Calculations[a]

	M	Sqaured coefficients			
		$\sigma^2_{ns}\sigma^{*2}_{ns}$	$\sigma^2_{ns}\sigma^{*2}_{npz}$	$\sigma^1_{ns}\sigma^{*1}_{ns}\pi^1_{npx}\pi^{*1}_{npy}$	$\sigma^1_{ns}\sigma^{*1}_{ns}\pi^1_{npy}\pi^{*1}_{npx}$
Be_2	0.27	0.80	0.069	0.024	0.024
Mg_2	0.12	0.86	0.007	0.023	0.023
Ca_2	0.13	0.83	0.022	0.026	0.026
Sr_2	0.12	0.84	0.018	0.025	0.025
Ba_2	0.14	0.83	0.019	0.020	0.020
Ra_2	0.10	0.87	0.010	0.021	0.021

[a]Geometries for Be_2 to Ba_2 are experimental geometries. For Ra_2, the CPO-tPBE, FV-ftPBE, and CPO-ftPBE calculations all give 5.28 $\mathring{\mathrm{A}}$, and therefore, that distance was used for Ra_2 in this table.

In order to show which configurations dominate the deviation from a single-configuration

wave function, Table 3.1 shows the dominant configurations in the FV-CASSCF wave function at the equilibrium geometry. The same four configurations have the largest squared coefficients for all six dimers. We observe that a determinant with configuration $\sigma_{n\mathrm{s}}{}^2\sigma_{\mathrm{pz}}{}^2$ plays a much more important role in the wave function for Be_2 than it does for the rest of the dimers, which is consistent with the important role of s–p hybridization in Be_2 that has previously been pointed out by Schmidt *et al.*[96]

In the CPO-CASSCF calculations, only the first two configurations shown in Table 3.1 have significant weights; the configuration coefficients are in the SI.

3.3.2 Binding Energies (D_e) for dimers from Be_2 to Ba_2

The reference values of the bond energies and equilibrium distances are given in Table 3.2. The MC-PDFT and CASPT2 results calculated with the largest basis set (the AR basis set) using the two active spaces are given in Table 3.3. We consider four mean unsigned errors (MUEs) for MC-PDFT; these MUEs correspond to the four possible combinations of two active spaces and two on-top density functionals. We also give two MUEs for CASPT2; they correspond to the two active spaces. For the bond energy, three of the four MUEs for MC-PDFT are smaller than the better of the two CASPT2 MUEs, and the fourth is the same. This is a key finding of the present study. Table 3.3 also shows mean signed errors (MSEs) and standard deviations. Table 3.3 also shows mean signed errors and standard deviations.

Note that the ftPBE calculations overestimate the bond energy for Be_2 and Ba_2, underestimate it for Mg_2, and either overestimate it or underestimate it, depending on the active space, for Ca_2 and Sr_2. The smallest errors of the ftPBE calculations are achieved for Mg_2 and Ba_2, and the largest error is for Be_2.

Table 3.2: Reference Values for Dissociation Energies ($D_{\rm e}$, in kcal/mol) and Equilibrium Distances ($r_{\rm e}$, in $\mathring{\rm A}$)

	$D_{\rm e}$	$r_{\rm e}$
Be_2	2.66[a]	2.454a
Mg_2	1.24[b]	3.890b
Ca_2	3.13[c]	4.287c
Sr_2	3.02[c]	4.663c
Ba_2	3.87[c]	4.972c

[a]These results are from the most recent experimental value (2.658 $\pm$ 0.006 kcal/mol) in ref [84].
[b]These results are from experimental values in ref [112].
[c]The results are from high-level theoretical calculations in ref [129].

3.3.3 Equilibrium Internuclear Distances ($r_{\rm e}$) for dimers from Be_2 to Ba_2

Table 3.1 also shows the accuracy that can be achieved for the bond distances. Again we consider four MUEs for MC-PDFT; these MUEs correspond to the four possible combinations of two active spaces and two on-top density functionals. We also give two MUEs for CASPT2; they correspond to the two active spaces. For the equilibrium distance, all four MUEs for MC-PDFT are smaller than either of the CASPT2 MUEs. This is another important finding of the present study.

3.3.4 Larger Active Spaces

We also explored bigger active spaces for MC-PDFT calculations on Be_2, and the results are in Table S2 in the SI. The performance is somewhat worse with the larger active spaces. This result is consistent with a previous detailed study on H_2 where we found that very large active spaces can degrade the result.[152] In general, we recommend that the active space should be large enough to include the dominant static correlation but not significantly larger.

Table 3.3: Dissociation Energy, Equilibrium Distance, Mean Signed Error (MSE), Mean Unsigned Error (MUE), and Standard Deviation (StdDev) for Calculations with the Large AR Basis Set[a]

		tPBE		ftPBE		CASPT2	
active space	molecule	D_e	r_e	D_e	r_e	D_e	r_e
FV	Be_2	3.58	2.404	1.42	2.436	0.00[b]	2.554
	Mg_2	0.49	3.800	–0.10[b]	4.242	0.60	4.313
	Ca_2	2.64	4.284	1.56	4.376	1.81	4.536
	Sr_2	2.76	4.657	1.77	4.752	1.94	4.855
	Ba_2	3.89	4.960	2.80	5.038	2.88	5.102
	Ra_2	2.34	5.275	1.48	5.379	1.82	5.521
	MSE	–0.11	–0.032	–1.3	0.12	–1.3	0.22
	MUE	0.49	0.032	1.3	0.12	1.3	0.22
	StdDev[c]	0.64	0.037	0.18	0.14	0.78	0.13
CPO	Be_2	4.09	2.440	4.44	2.375	4.50	2.423
	Mg_2	1.00	3.767	0.96	3.768	2.83	4.068
	Ca_2	3.46	4.304	3.12	4.314	4.39	4.439
	Sr_2	3.58	4.682	3.27	4.682	4.53	4.806
	Ba_2	4.80	4.983	4.52	4.975	5.73	5.056
	Ra_2	3.11	5.276	2.79	5.282	4.24	5.410
	MSE	0.60	–0.018	0.48	–0.030	1.6	0.11
	MUE	0.70	0.037	0.59	0.050	1.6	0.12
	StdDev[c]	0.63	0.060	0.80	0.066	0.25	0.084

[a]All energies in this table are in kcal/mol, and all distances are in $\mathring{\mathrm{A}}$. Error statistics are based on Be_2–Ba_2 because accurate results are not available for Ra_2.

[b]These nonbonding results correspond to local minima on potential energy curves. See Figure 3.1b as an example where CASPT2 fails. See also refs [105] and [106] for similar results, consistent with the difficulty in studying Be_2.

[c]The standard deviation in the tables is the standard deviation of the unsigned errors from their mean.

Therefore, it is encouraging that we get generally good results with the two standard active space choices in Table 3.3.

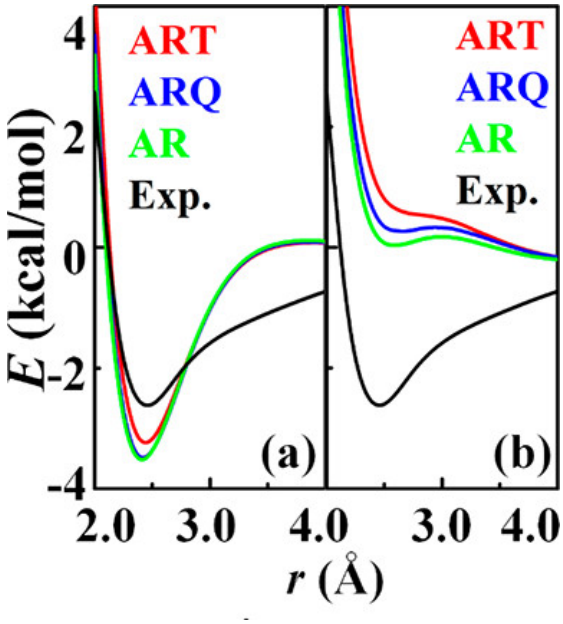

Figure 3.1: Potential energy curve (zero at 20 Å) of Be2 near equilibrium by (a) FV-ftPBE and (b) FV-CASPT2, using the ART, ARQ, and AR basis sets. Also shown is the experimental curve (abbreviated Exp.).

3.3.5 Basis Set Dependence

The MC-PDFT and CASPT2 results calculated with the contracted basis sets using the two active spaces are given in Table 3.4. With smaller valence basis sets, namely ARQ and ART, a general trend is found that the results of CASPT2 change more (as compared to the larger AR basis set) than do the ftPBE calculations. Comparison of Table 3.4 to Table 3.3 shows that the accuracy of CASPT2 degrades significantly with smaller basis sets, which is not surprising because of the slow convergence of perturbation theory with basis set size. For ftPBE we see a smaller degradation in bond energies and no degradation in bond distances. Thus MC-PDFT retains its accuracy comparatively better with basis sets that

would be more practical for larger systems.

Table 3.4: Dissociation Energy, Equilibrium Distance, Mean Signed Error (MSE), Mean Unsigned Error (MUE), and Standard Deviation (StdDev) for Calculations with the Large AR Basis Set[a]

		ART		ARQ	
method	molecule	D_{e}	r_{e}	D_{e}	r_{e}
ftPBE	Be_2	3.30	2.435	3.56	2.412
	Mg_2	0.40	3.855	0.47	3.812
	Ca_2	2.44	4.303	2.55	4.296
	Sr_2	2.53	4.676	2.70	4.669
	Ba_2	3.68	5.000	3.86	4.970
	MSE	–0.31	0.001	–0.16	–0.021
	MUE	0.57	0.022	0.52	0.027
	StdDev[b]	0.59	0.026	0.66	0.038
CASPT2	Be_2	–0.56	2.513	–0.24	2.630
	Mg_2	0.56	4.690	0.64	4.397
	Ca_2	1.13	4.790	1.41	4.653
	Sr_2	1.17	5.199	1.69	4.972
	Ba_2	3.07	5.154	3.88	5.016
	MSE	–1.71	0.42	–1.31	0.28
	MUE	1.71	0.42	1.31	0.28
	StdDev[b]	1.03	0.30	1.11	0.18

[a] All energies in this table are in kcal/mol, and all distances are in $\mathrm{\AA}$.
[b] The standard deviation in the tables is the standard deviation of the unsigned errors from their mean.

3.3.6 Comparison to KS-DFT

The KS-PBE results are given in Table 3.5. The KS-DFT bond energies with the parent GGA are, on average, much less accurate than MC-PDFT or CASPT2. We see that MC-PDFT with either active space gives the most accurate bond distances, followed by CPO-CASPT2, then KS-PBE, and at last FV-CASPT2.

Table 3.5: Dissociation Energy, Equilibrium Distance, Mean Signed Error (MSE), Mean Unsigned Error (MUE), and Standard Deviation (StdDev) for Calculations with KS-DFT Using the PBE Exchange–Correlation Functional[a]

	ART		ARQ		AR	
molecule	D_{e}	r_{e}	D_{e}	r_{e}	D_{e}	r_{e}
Be_2	9.7	2.458	9.8	2.438	9.8	2.431
Mg_2	3.2	3.518	3.2	3.512	3.2	3.510
Ca_2	5.2	4.197	5.3	4.188	5.4	4.171
Sr_2	4.7	4.556	4.8	4.555	4.9	4.544
Ba_2	6.4	4.863	6.6	4.845	6.6	4.845
Ra_2	3.6	5.197	3.5	5.205	3.8	5.172
MSE	3.1	–0.14	3.2	–0.15	3.2	–0.15
MUE	3.1	0.14	3.2	0.15	3.2	0.15
StdDev[b]	2.3	0.14	2.3	0.14	2.2	0.13

[a]All energies in this table are in kcal/mol, and all distances are in $\mathring{\mathrm{A}}$. Error statistics are based on Be_2–Ba_2 because accurate results are not available for Ra_2.
[b]The standard deviation in the tables is the standard deviation of the unsigned errors from their mean.

3.3.7 Potential Energy Curves

Figure 3.2 shows dissociation curves of Mg_2 and Ca_2 by FV-ftPBE, FV-CASPT2, and KS-PBE as compared to the experimental curve from ref. [153] and ref. [154] , respectively. (The potential energy curve for Be_2 is shown in Figure 3.2 discussed below, and the potential energy curves for all six dimers are given in tabular form Table S18-S23 in the Supporting Information. We did not find experimental potential energy curves for Sr_2, Ba_2, or Ra_2.) Figure 1 shows that the FV-ftPBE curve agrees with the experiment very well for both equilibrium distances and dissociation energies, but it is underestimates for the long-range tail of the potential, which may be attributed to underestimation of dispersion energy contributions.

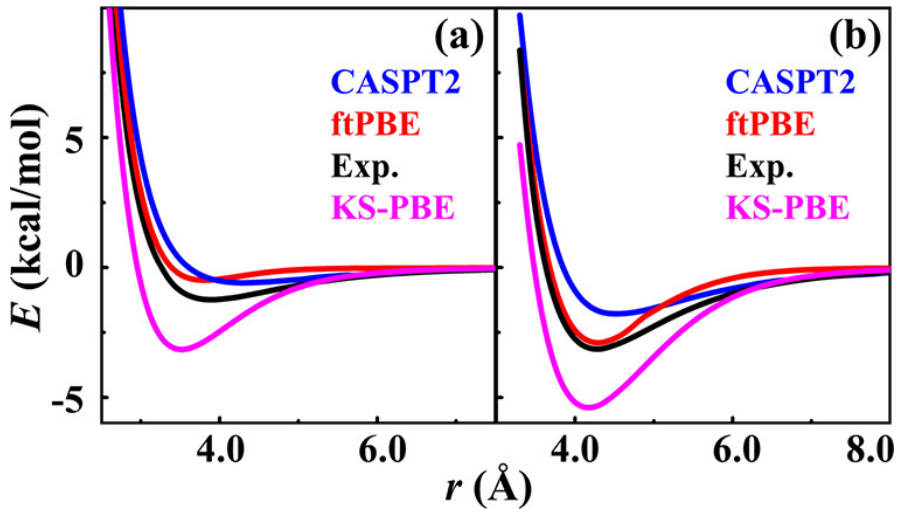

Figure 3.2: Dissociation curves for (a) Mg_2 and (b) Ca_2 by FV-CASPT2, FV-ftPBE, and KS-PBE using the AR basis set. Also shown is the experimental curve (labeled "Exp.").

3.3.8 Performance of Other On-top Functionals

In Tables S4-S6 we present results for other on-top translated and fully translated functionals[78] compared with each other and compared with their untranslated KS analogs, namely KS-revPBE[79] and KS-OPBE,[62, 155] together with their translated or fully translated versions. We found that fully translated functionals are usually more accurate than their corresponding translated ones, and the fully translated functionals have better performance than their corresponding KS parent functionals. One possible reason for the improved performance of the ft functionals is that they include the gradient of the on-top density.

3.3.9 MP2 vs. CASPT2

One may ask whether the improved performance of ftPBE or tPBE over KS-PBE originates from MC-PDFT being able to handle multiconfigurational systems better. The M diagnostics presented above show that these systems are inherently multiconfigurational (i.e., strongly correlated), but it is also useful to consider another aspect of multireference

character, namely the deviation of CASPT2 based on CASSCF from single-configuration Møller-Plesset perturbation theory,[156, 157] (MP2) based on a restricted Hartree-Fock calculation. The MP2 results are given in the SI. When using the AR basis set, we find – unexpectedly – that MP2 performs better than CASPT2. This demonstrates how difficult it is to identify sources of errors in approximate calculations. There is a subtle interplay between all of the approximations, and there is some cancellation of errors, and for weak interactions this can subvert any simple attempt to divide the errors into categories.

3.3.10 Core Correlation

All the CASPT2 calculations reported here and in the SI make the frozen-core approximation. The SI contains a comparison of full MP2, denoted as MP2(Full), which includes excitation of all core orbitals, and frozen-core MP2, denoted FC-MP2, which does not allow excitations of core orbitals. The latter is important because all the CASPT2 calculations reported here and in the SI make the frozen-core approximation. The comparison is in Tables S7 and S8, and it shows that sometimes core correlation improves the results and sometimes in makes them worse. This is perhaps not surprising because it is known that core polarization is particularly important in group 2 of the periodic table.[158, 159, 160] One advantage of density functional methods over wave function methods is that density functional methods always include core correlation, at least to the extent allowed by the basis set.

3.3.11 Diffuse Basis Functions and Core Polarization Basis Functions

Tables S9-S12 present additional tests for Be_2 and Mg_2 showing that the results are not sensitive to diffuse functions or to the core polarization functions in the basis sets.

3.3.12 D_e and r_e for Ra_2

Because there is no accurate experiment available for Ra_2, we also made ftPBE calculations for Ra_2. The results given by FV-ftPBE and CPO-tPBE are in a quite good agreement with some previous calculations. A CCSD(T) calculation using exact two-component[161, 162] (X2C) theory for the relativistic effect gave 2.6 kcal/mol for D_e and 5.324 $\mathring{A}$ for r_e.[130] A CASPT2 calculation using an active space with 4 electrons in 16 orbitals and the AR basis set gave 1.82 kcal/mol for D_e and 5.478 $\mathring{A}$ for r_e.[129] The data are in a good agreement with our FV-CASPT2 calculation. The ftPBE result lies between the CCSD(T) and previous CASPT2 results for the dissociation energy, but is closer to the CCSD(T) results for bond distance.

3.4 Conclusion

Alkaline-earth dimers embody two of the most challeging problem in quantum chemistry – weak interactions and strongly correlated systems. The present study shows that MC-PDFT provides improved accuracy as compared to KS-DFT and CASPT2 for weak bonds in metal dimers. Furthermore, MC-PDFT has the advantage of having less dependence on basis sets than one finds for wave function methods like MP2 or CASPT2. This reduced basis-set dependence is usually an advantage of density functional methods over methods that must explicitly build the electron-electron cusp into the wave function in order to accurately approximate the dynamic correlation energy, and MC-PDFT provides a way to take advantage of this in a setting that includes multi-configurational effects in the representation of the electron density.

Chapter 4

Automatic selection of an active space for calculating electronic excitation spectra by MS-CASPT2 or MC-PDFT

4.1 Introduction

An early issue in the history of electronic structure theory was the determination of the best orbitals and configurations for configuration interaction (CI) calculations; we give only a few representative references.[163, 164, 165, 166, 167, 168, 169, 170] As time progressed, configuration interaction calculations lost favor to size-extensive and more powerful coupled cluster calculations based on a single configuration, where the question of configuration selection does not arise, and to active space methods where internal correlation is calculated by an active-space multiconfiguration self-consistent field (MCSCF) calculation followed by a post-MCSCF calculation of external correlation. Methods employing the latter approach

are called multireference methods. The choice of orbitals and configurations for the active space is even thornier than for single-reference CI or coupled cluster calculations because it is a common occurrence that the desired active orbitals cannot be maintained in the active space during SCF iterations. This can be ameliorated by using better initial guesses, and this is an important part of current strategies. As active-space methods are being applied to larger and more complex systems, and as molecular dynamics considerations demand wider exploration of potential energy surfaces, there have been increasing efforts[171, 80, 172, 173, 174, 175, 176, 177, 47, 178, 179, 180, 181, 182, 183, 184] to develop procedures and algorithms, even automated ones, for the selection of active space orbitals, improved virtual orbitals, and/or configurations and even to create expert systems so the computer can make such decisions. The present article is one attempt to take a step in that direction.

Here we propose an automatic selection scheme of orbitals in a complete-active-space self-consistent-field (CASSCF)[145, 73, 185] calculation to be used as a reference for a post-CASSCF calculation. The method involves three parameters, which we will call A, B, and C, so the method will be called the ABC scheme. The goal is to set up a deterministic set of steps that qualifies as theoretical model chemistry by the criteria of Pople:[186] "The approximate mathematical procedure must be precisely formulated. This should be general and continuous as far as possible. Thus, particular procedures for particular molecules or particular symmetries should be avoided." We will define a three-step procedure with three parameters, and we will test it by applying precisely the same three steps with the same three parameters to a set of ten test molecules. We realize that the procedure will not be completely general (i.e., as one considers a great variety of systems, it will surely be necessary to improve it), but we hope to make a good start in the direction of giving some prescriptions on how to make these multireference calculations more automatized for different classes of systems. What we are presenting here is one particular prescription for a class of systems. Similar prescriptions can be envisioned for different classes of systems.

We tested the ABC scheme on ten doublet radicals in three groups:

Group 1: 9-electron systems, in particular, CH_3 and NH_2

Group 2: 13-electron systems, in particular, C_2H, CN, BO, N_2^+, and CO^+

Group 3: 15-electron systems, in particular, C_2H_3, CH_2N, and CH_2O^+

For each radical our goal is to calculate accurate excitation energies of the first five excited doublet states. To judge the results, we calculate accurate benchmarks for all 50 excitation energies by multireference configuration interaction with single and double excitations[41, 42, 187, 188, 189] with a Davidson correction[43, 190] (MR-CISD+Q). We test the ABC scheme for two different levels of post-CASSCF calculation: multi-state complete active space second-order perturbation theory[15, 16, 17, 75] (MS-CASPT2) and multiconfiguration pair-density functional theory[19, 20, 21] (MC-PDFT).

4.2 ABC Scheme

The state-averaged ground state necessarily has a higher energy than the single-state ground state, but the state-averaged method allows one, in principle, to obtain a balanced treatment of ground and excited states. Therefore, we use state-averaged calculations in the present chapter.

The method is designed to calculate the first N_{EE} electronic excitation energies. This means one must calculate the ground state and N_{EE} excited states for a total of $N_{\mathrm{S}} = N_{\mathrm{EE}} + 1$ states. A familiar difficulty in trying to calculate these excitation energies by a post-MCSCF method is making sure that the orbitals involved as holes or particles in the configurations that dominate these excitations are present and suitably optimized in the MCSCF reference function. The holes (orbitals from which excitations are made) should be present in the active space as occupied orbitals. However, if one includes too many occupied orbitals in the active spaces, the calculations will become unwieldy or even unaffordable. If one includes

too few, then the calculations will be inaccurate. In either case one might end up with an unbalanced choice of holes. Our first parameter (A) will be used to automate the choice of doubly occupied orbitals that may serve as holes.

An even more vexing problem is to make sure that the calculations include the orbitals needed for particles (orbitals into which excitations are made). The virtual orbitals of a Hartree–Fock calculation are well-known to be unphysical,[191, 192, 193] and our experience is that simply using the low-energy unoccupied orbitals of a Hartree–Fock calculation as initial guesses for the active space orbitals often leads to active spaces that are missing the key particle orbitals needed for an accurate treatment of excited states. One can obtain better results with a state-averaged (SA) restricted active space self-consistent field (RASSCF)[194] calculation in which the particle orbitals are optimized for the excitations of interest. But one has to have initial guess orbitals for the state-averaged calculation, and the first state-averaged calculation must necessarily have orbitals that are not optimized in a previous calculation. In our experience, when one does that, the needed particle orbitals for the low-energy excitations are not necessarily the lowest-energy orbitals. Thus, one must use a large virtual space to be sure the needed particle orbitals are optimized and present. A key step in our procedure (it will be called step 2, where step 1 is the calculation used to generate the initial guess orbitals for step 2) will be to find and optimize the needed particle orbitals in an automatic way without making case-specific decisions and without visualizing the orbitals. Our goal is to find out if one can do this systematically in an affordable manner, and the second parameter (B) will be used to automate the number of virtual orbitals used in this step.

If one's target is to calculate accurate energy differences between the ground state and N_{EE} excited states, then one should do state-averaged calculations over at least N_{EE} + 1 states. But in our experience the N_{EE}th excited state often interacts strongly with the next excited state (some configuration state functions contribute significantly to more than one

low-energy eigenstate), and in fact, to get a balanced treatment of N_{EE} + 1 excited states, it may be necessary to do a state average that includes additional excited states. Our third parameter (C) is used to automate the choice of the number of states.

With the above considerations as motivations, we arrived at the following scheme.

We specify the number of unpaired electrons in the ground electronic state as V. The ten radicals employed in the present test of the scheme all have ordinary doublet ground states for which V = 1, but the method is designed to also be applicable to other multiplicities, for example, to closed-shell singlet molecules with V = 0. We use the notation HDOMO for the highest doubly occupied molecular orbital in the dominant configuration of the ground state. The ABC scheme of automatically selecting active spaces has three steps.

In the first step, one puts 2A + V active electrons in A + V active orbitals and performs a ground-state CASSCF[73] calculation. (Note that when V is zero, this becomes simply a restricted Hartree–Fock (RHF) calculation, which is a special case of CASSCF, but for open-shell cases it is an MCSCF calculation.) We will also use A in subsequent steps. If the configurations that make significant contributions to the targeted excitations involve excitations from orbitals no deeper than HDOMO–(A–1), then one should include A doubly occupied orbitals in the active space of all steps, and A becomes our first parameter. The physical justification for the A parameter is that the lowest-lying excitations tend to come from the highest-energy occupied orbitals. The question we will address is can we obtain good results for a range of systems with a single value of A? In the second step, one runs an SA-RASSCF[194] calculation, with 2A electrons and A orbitals in the RAS1 space, with V electrons and V orbitals in the RAS2 space, and with B virtual orbitals. The initial guesses for all orbitals are taken from step 1. In particular, the B virtual orbitals are the virtual orbitals with B lowest orbital energies in step 1. The maximum number of holes in RAS1 is 1, and the maximum number of particles in RAS3 is 1. The choice of limiting the level of excitations to singles in this step is based on the consideration that this is still a preparatory step to

generate qualitatively good orbitals for the final step. We do not exclude the possibility that for some classes of problems one has to allow up to double excitations. Nevertheless, for all the results in the present article, we included only single excitations in step 2. The number of states averaged in this calculation is N_{S} + C. We calculate C extra roots for two physical reasons: (i) The N_{S} lowest states in reality need not be the same as the N_{S} lowest states of step 2; step 2 is designed to tell us what orbitals to use in step 3, not to give accurate energies. (ii) As mentioned above, in the motivation of the method, a balanced treatment of the NS lowest-energy states may require averaging over a greater number of states because the states are multiconfigurational and some configuration state functions contribute to more than one low-energy physical state.

To prepare for step 3, the user examines the dominant configurations of the excited states. If V is 0 or 1, the dominant excitations are all single excitations (because step 2 automatically only includes single excitations if V is 0 or 1), and one identifies the hole orbital and the particle orbital for each of the N_{S} + C dominant excitations (one dominant excitation for each state) by comparing its dominant configuration with that of the ground state. These holes and particles are called the dominant transition orbitals. This identification is done after the transformation to state-averaged natural orbitals, so these particle and hole orbitals are state-averaged natural orbitals.

The third step is an SA(N_{S}+C)-CASSCF calculation in which the initial active orbitals are the dominant transition orbitals of step 2. The initial guesses for the SA(N_{S}+C)-CASSCF process are the final orbitals of the RASSCF calculation in step 2. If one is using MS-CASPT2, one uses the wave function obtained from the CASSCF calculation in step 3 as the reference function for an MS-CASPT2 calculation with N_{S} + C states in the model space, and the N_{S} lowest-energy states of this MS-CASPT2 calculation are used to calculate NEE excitation energies.

If one is using MC-PDFT, one calculates the kinetic energy, density, and on-top pair

density for the $N_S + C$ lowest-energy states from the CASSCF calculation in step 3 and uses these quantities to calculate MC-PDFT energies. The N_S lowest-energy MC-PDFT energies are used to calculate N_{EE} excitation energies.

Notice that A denotes the number of doubly occupied active orbitals, B denotes the number of unoccupied orbitals in step 2, and C denotes how many more states we include above the number targeted. These are the only three parameters because V is determined by the ground state of the chosen molecule, and N_S depends on the user's goal for the molecule chosen to be studied, i.e., how many states are of interest.

The parameter A determines how deep one goes into doubly occupied molecular orbitals (DOMOs) that possibly participate in electronic excitation. The recommended value used here is 3. If the third highest DOMO (i.e., HDOMO–2) were degenerate with HDOMO–3, then A would be increased to 4, but this increase was not needed for any of the molecules in the present tests.

The parameter B determines how high one goes into the virtual space in step 2. The recommended value for B is 25. (In a test case for BO radical, the CASSCF results in the third step for B = 25 and B = 30 are the same, whereas B = 20 was not sufficient, so we used B = 25 for all the other systems. If desired, users can check whether the results are sensitive to increasing B for their problem.)

The parameter C is the number of extra states calculated in step 2. The currently recommended value of C is 2, but if states $N_S + C$ and $N_S + C + 1$ are degenerate, we should keep $N_S + C + 1$ states for that case. For example, if N_S = 6 and C = 2, we would nominally consider 8 roots, but if the eighth and ninth roots are degenerate, we consider 9 so as not to break a degeneracy.

We will use ethylene (which has V = 0) as an example for how the scheme works. For the example, we set N_{EE}, A, B, and C equal to 2, 3, 25, and 2, respectively. Because the ground state is a closed-shell singlet and A is 3, the first step is a ground-state CASSCF

calculation with 6 electrons in 3 orbitals, namely, a RHF calculation (because there is only one configuration in the calculation). The second step is a RASSCF calculation averaged over five states (the ground state plus $N_{\mathrm{EE}} + C$ excited states) using the orbitals from the first step as the initial guess. The number of RAS1 orbitals, RAS2 orbitals, and RAS3 orbitals in step 2 are 3, 0, and 25, respectively. The maximum number of holes in RAS1 and maximum number of particles in RAS3 are both 1 in the current form of the ABC scheme. The third step picks all the transition orbitals in step 2 and put them in the RAS2 space, and then runs a CASSCF calculation averaging over 5 states.

The question we will address is, can we obtain good results for a range of systems with a single set of values of A, B, and C? If we identify such a set of values, it can be used for well-defined predictions on other systems. Of course, if one has specific knowledge about new systems, one can perhaps use that knowledge to find even better active spaces. In fact, even for the systems treated here, we can obtain more accurate results, as discussed below; but that is not our goal here. The present scheme involves no system-specific knowledge and no visualization of orbitals; it is an attempt to find an automatic scheme for active space selection.

4.3 Computational Details

We consider ten doublet radicals in this work. The geometries used for the ten molecules are listed in Table 4.1.

The benchmark results were calculated using MR-CISD+Q as implemented in Molpro.[76, 77] The +Q corrections are computed by the rotated wave function method of Werner *et al.*[190] The active spaces that we used for the benchmark calculations are similar to the ones used by Li and Liu,[200] with possibly small changes in the number of occupied or virtual orbitals. These active spaces are listed in Table 4.2. For radicals in the first two

Table 4.1: Geometries in Terms of Bond Lengths ($\mathring{\mathrm{A}}$) and Bond Angles (degree)

radical	geometry	reference
CH_3	R(CH) = 1.0767, $\angle$(HCH) = 120	[195]
NH_2	R(NH) = 1.024,[a] $\angle$(HNH) = 103.4	[196]
C_2H	R(CC) = 1.208, R(HC) = 1.058, $\angle$(HCC) = 180	[197]
CN	R(CN) = 1.1718	[30]
CO^+	R(CO) = 1.1151	[30]
N_2^+	R(NN) = 1.1164	[30]
BO	R(BO) = 1.2045	[30]
C_2H_3	R(C1C2) = 1.340, R(C1H1) = 1.101, R(C2H2) = 1.112, R(C2H3) = 1.107, $\angle$(H1C1C2) = 135.1, $\angle$(C1C2H2) = 121.7, $\angle$(H2C2H3) = 116.5	[198]
CH_2N	R(CN) = 1.2655, R(CH) = 1.1010, $\angle$(HCH) = 121.27	[199]
CH_2O^+	R(CO) = 1.2078, R(CH) = 1.1161, $\angle$(HCH) = 116.57	[38]

[a]The value for r_0 is used for r_e of the N–H bond.

groups, C_1 symmetry is used for the benchmark calculations; for radicals in the last group, C_s symmetry is used for C_2H_3 and $C_{2\mathrm{v}}$ for CH_2N and CH_2O^+.

In keeping with Pople's recommendation for a theoretical model chemistry, the MC-PDFT and MS-CASPT2 calculations use C_1 symmetry for all the three groups of radicals. For all ten cases tested in the present work, V is 1, and our targeted number of excitation energies is 5. The number of roots (N_S) and the parameters used in the ABC scheme are listed in Table 4.3. Also listed in Table 4.3 are the active spaces arrived at by the ABC scheme and the orbitals contained in the ABC active space. The orbitals in the active spaces are illustrated in Figures S1–S10 in the Supporting Information.

The CASSCF, MS-CASPT2, and MC-PDFT calculations were carried out using Molcas 8.1.[74] The MS-CASPT2 calculations include the default ionization potential–electron affinity (IPEA) shift[148] of 0.25 $\mathrm{E_h}$, and they were done with an imaginary shift[149] of 0.20 $\mathrm{E_h}$.

The MC-PDFT calculations were carried out with tPBE[19] on-top density functional and the trevPBE[47] on-top density functional.

Table 4.2: Details of the SA-CASSCF Calculations Used as the Starting Points for MR-CISD+Q Calculations

radical	active space	SA[a]	dominant configuration in ground state
CH_3	(5,10)	10	$(1e'_x)^2(1e'_y)^2(1a''_2)^1(2e'_x)^0(2e'_y)^0(3a'_1)^0(2a''_2)^0(3e'_x)^0(3e'_y)^0(4a'_1)^0$
NH_2	(5,6)	10	$(1b_1)^2(3a_1)^2(1b_2)^1(4a_1)^0(2b_1)^0(5a_1)^0$
C_2H	(9,9)	9	$(3\sigma)^2(4\sigma)^2(5\sigma)^1(1\pi_x)^2(1\pi_y)^2(2\pi_x)^0(2\pi_y)^0(6\sigma)^0(7\sigma)^0$
CN	(7,7)	10	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0(6\sigma)^0$
CO^+	(7,7)	10	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0(6\sigma)^0$
N_2^+	(7,8)	10	$(2\sigma_u)^2(2\pi_{u,x})^2(2\pi_{u,y})^2(3\sigma_g)^1(1\pi_{g,x})^0(1\pi_{g,y})^0(3\sigma_u)^0(4\sigma_g)^0$
BO	(7,8)	10	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0(6\sigma)^0(7\sigma)^0$
C_2H_3	(7,10)	7,4[b]	$(5a')^2(6a')^2(7a')^1(8a')^0(9a')^0(10a')^0(11a')^0(1a'')^2(2a'')^0(3a'')^0$
CH_2N	(9,11)	6,6,6,6[c]	$(4a_1)^2(5a_1)^2(6a_1)^0(7a_1)^0(8a_1)^0(1b_1)^2(2b_1)^1(3b_1)^0(1b_2)^2(2b_2)^0(3b_2)^0$
CH_2O^+	(9,11)	5,5,5,5[c]	$(4a_1)^2(5a_1)^2(6a_1)^0(7a_1)^0(8a_1)^0(1b_1)^2(2b_1)^1(3b_1)^0(1b_2)^2(2b_2)^0(3b_2)^0$

[a]The SA column gives the number of states averaged. When a series of numbers is given, they denote the number of states in each irrep.
[b]The two irreps used are A′ and A″.
[c]The four irreps are A_1, B_1, B_2, and A_2.

4.4 Results and Discussion

Our goal is not to show that we can get low-energy CASSCF states or accurate excitation energies by SA-CASSCF but rather to show that we can get active spaces that serve as useful references for post-SCF steps like those in MS-CASPT2 and MC-PDFT. We tested the ABC scheme on three groups of radicals classified according to the number of electrons, as specified in the Introduction. The errors of MC-PDFT calculations using the tPBE and trevPBE on-top functionals are shown in Tables 4.4-4.13. The mean unsigned deviation (MUD) for each state is listed in Table 4.14. Readers interested in specific molecules can consult the individual tables, but because the tables for individual molecules largely speak for themselves, we will keep the discussion short.

Table 4.14 shows that the MUD over all 50 excitation energies is 0.14 eV for MS-CASPT2,

Table 4.3: Details of ABC Active Spaces

	target		parameters			ABC active	dominant configuration
no.	radical	N_S	A	B	C	space	in ground state
1	CH_3	6	3	25	2	(5,8)	$(1e'_x)^2(1e'_y)^2(1a''_2)^1(2e'_x)^0(2e'_y)^0(2a''_2)^0(3e'_x)^0(3a'_1)^0$
2	NH_2	6	3	25	2	(5,5)	$(1b_1)^2(3a_1)^2(1b_2)^1(4a_1)^0(2b_1)^0$
3	C_2H	6	3	25	3[c]	(5,5)	$(5\sigma)^1(1\pi_x)^2(1\pi_y)^2(2\pi_x)^0(2\pi_y)^0$
4	CN	6	3	25	3[c]	(7,6)	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0$
5	BO	6	3	25	3[c]	(7,6)	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0$
6	CO^+	6	3	25	3[c]	(7,6)	$(4\sigma)^2(1\pi_x)^2(1\pi_y)^2(5\sigma)^1(2\pi_x)^0(2\pi_y)^0$
7	N_2^+	6	3	25	2	(7,6)	$(2\sigma_u)^2(2\pi_{u,x})^2(2\pi_{u,y})^2(3\sigma_g)^1(1\pi_{g,x})^0(1\pi_{g,y})^0$
8	C_2H_3	6	3	25	2	(3,6)	$(7a')^1(8a')^0(9a')^0(10a')^0(1a'')^2(2a'')^0$
9	CH_2N	6	3	25	2	(5,6)	$(5a_1)^2(6a_1)^0(2b_1)^1(3b_1)^0(1b_2)^2(2b_2)^0$
10	CH_2O^+	6	3	25	2	(7,5)	$(5a_1)^2(1b_1)^2(2b_1)^1(1b_2)^2(2b_2)^0$

[a]The SA column gives the number of states averaged. When a series of numbers is given, they denote the number of states in each irrep.
[b]The two irreps used are A′ and A″.
[c]The four irreps are A_1, B_1, B_2, and A_2.

which shows that the method is very successful. The MUD is a bit higher for MC-PDFT, for which the error depends slightly on the choice of on-top density functional. The MUD is 0.19 eV for tPBE, and 0.21 eV for trevPBE. Previously[201, 202] we defined functionals with MUDs of 0.30 eV or less to be “successful”. Using this criterion, we judge that all MS-CASPT2 and MC-PDFT calculations with the ABC scheme can be classified as very successful. The fact that we obtained good results for MC-PDFT as well as MS-CASPT2 is especially encouraging because we have found for some problems that MC-PDFT is more sensitive to the quality of the active space than is MS-CASPT2.

Among the ten radicals, MS-CASPT2 works extremely well (MUD over five excitation energies is no more than 0.10 eV) for CH_3, BO, N_2^+, CO^+, and CH_2O^+. Using the same standard, MC-PDFT with tPBE and trevPBE functionals works extremely well for CN and N_2^+, in both of which cases these two functionals perform even better than MS-CASPT2.

Table 4.4: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CH_3

state[a]			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2A_1'$	$(1a_2'')^0(3a_1')^1$	5.867	–0.76	–0.21	–0.24	0.030
2	$1^2E'$	$(1e_x')^1(1a_2'')^2$	6.943	0.85	–0.24	–0.21	0.068
3	$1^2E'$	$(1e_y')^1(1a_2'')^2$	6.943	1.10	–0.15	–0.11	0.13
4	$2^2E'$	$(1a_2'')^0(2e_x')^1$	7.181	–0.90	–0.45	–0.51	–0.013
5	$2^2E'$	$(1a_2'')^0(2e_y')^1$	7.181	–0.87	–0.45	–0.51	0.009
			MSD	–0.11	–0.30	–0.32	0.05
			MUD	0.89	0.30	0.32	0.05

[a]In the first column, the lowest doublet is labeled 0, and the excited doublets are labeled 1–5. The second column gives the symmetry. The third column gives the difference of the dominant configuration from the ground-state dominant configuration.
[b]The MR-CISD+Q columns of Tables 4.4–4.13 contain the benchmark results.

Although the parameters used in the ABC scheme are the same for all ten radicals, the final active space used for each radical can differ even for radicals in the same group, showing the power of the ABC scheme to be adaptive to different systems even with a fixed set of A, B, and C. Very interestingly, the active space generated by the scheme uses only 3 electrons for C_2H_3 radical, and the MUD over the all the five excited states is 0.17 eV for tPBE, although MS-CASPT2 does not perform quite as well using the same active space. It is also noteworthy that the active spaces generated by the scheme are different from each other within the first and the third groups, even though in each group the systems are isoelectronic (and we also note the similarity in symmetry for CH_2N and CH_2O^+); this shows how the ABC scheme successfully adapts to different systems.

If we look only at the vertical excitation energy of the first excited state, the largest unsigned errors (out of ten cases) are only 0.21 eV for MS-CASPT2 and 0.22 eV for tPBE. If we look only at the first two excitations, the largest errors (out of 20 cases) are only 0.35 eV for MS-CASPT2 and 0.35 eV for tPBE. Most of the results are already good enough for

Table 4.5: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for NH_2

	state		MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	1^2A_1	$(3a_1)^1(1b_2)^2$	2.089	0.10	–0.081	–0.057	–0.029
2	1^2B_1	$(1b_1)^1(1b_2)^2$	6.365	0.96	0.35	0.39	0.15
3	2^2A_1	$(1b_2)^0(4a_1)^1$	7.686	–0.94	–0.40	–0.43	0.07
4	2^2B_1	$(3a_1)^1(4a_1)^1$	7.766	–0.90	–0.10	–0.053	–0.29
5	1^2A_1	$(3a_1)^1(2b_1)^1$	9.186	–0.93	–0.33	–0.31	–0.26
			MSD	–0.34	–0.11	–0.09	–0.07
			MUD	0.77	0.25	0.25	0.16

practical use for many problems.

The tables do not show the MR-CISD results obtained without the +Q correction, but it is interesting to note that the MR-CISD calculations have an MUD over the 50 electronic excitations of 0.15 eV, only slightly better than MS-CASPT2 calculations, even though the MS-CASPT2 calculations are based on the automatic active space (i.e., the ABC scheme) whereas the MR-CISD results, being the first stage of the benchmark MR-CISD+Q calculations, are based on manually optimized active spaces. Before we developed the ABC scheme, we manually optimized active spaces for each of the ten molecules. The MUDs obtained for the 50 states in this way (i.e., with laboriously optimized active spaces) were 0.13 eV for trevPBE and MS-CASPT2 and 0.14 eV for tPBE. Although these results (0.13–0.14 eV) are better than the ABC results, we think they are less interesting than the ABC results (0.14–0.21 eV) for three reasons. First, the manually optimized active spaces have an average of 8.4 orbitals, whereas the ABC optimized active spaces have an average of only 5.9 orbitals. Second, the ABC method, being a theoretical model chemistry, can be applied systematically. Third, the ABC method, being an automatic method (after the three parameters are set), and without the need to visualize orbitals, is much less time-consuming

Table 4.6: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for C_2H

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2\Pi$	$(5\sigma)^2(1\pi_y)^1$	0.684	0.070	–0.19	–0.25	0.21
2	$1^2\Pi$	$(5\sigma)^2(1\pi_x)^1$	0.684	0.070	–0.19	–0.25	0.21
3	$2^2\Sigma^+$	$(1\pi_y)^1(2\pi_y)^1$	6.973	0.95	–0.13	–0.13	0.24
4	$1^2\Sigma^-$	$(1\pi_x)^1(2\pi_y)^1$	7.335	0.54	–0.19	–0.20	0.18
5	$2^2\Delta$	$(1\pi_x)^1(2\pi_y)^1$	7.657[a]	0.54	–0.25	–0.26	0.19
			MSD	0.43	–0.19	–0.21	0.21
			MUD	0.43	0.19	0.21	0.21

[a]This is one state of a pair of degenerate states

in human time.

To make the scheme work in other cases, it may be necessary to add new elements, for example, Kohn–Sham orbitals in step 1, double excitations in step 2, or a different value of C in step 3 from that used in step 2. As another example, in step 2, we may also choose the transition orbitals on the basis of two or more most dominant configurations for each state or on the basis of natural orbital occupation numbers. But it is very encouraging that the present scheme already does quite well for a range of systems.

4.5 Conclusion

In common parlance, a major detriment to the application of multireference methods is that they are not “black-box” methods. The present effort is a step in the direction of trying to mitigate that impediment by defining a scheme for finding active spaces. As mentioned in the Introduction, we expect that the scheme will need refinement to be applied to a larger class of systems, for example, to apply it to systems with conjugated bonds or systems containing transition metals or to apply it along reaction paths, which are notoriously challenging for

Table 4.7: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CN

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2\Pi$	$(1\pi_x)^1(5\sigma)^2$	1.284	0.30	–0.006	–0.051	0.18
2	$1^2\Pi$	$(1\pi_y)^1(5\sigma)^2$	1.284	0.30	–0.006	–0.051	0.18
3	$2^2\Sigma^+$	$(4\sigma)^1(5\sigma)^2$	3.130	0.56	0.15	0.17	0.11
4	$2^2\Pi$	$(5\sigma)^0(2\pi_y)^1$	7.791	0.33	–0.014	0.000	0.09
5	$2^2\Pi$	$(5\sigma)^0(2\pi_x)^1$	7.791	0.33	–0.014	0.000	0.09
			MSD	0.36	0.02	0.01	0.13
			MUD	0.36	0.04	0.05	0.13

multireference methods because the wave function may change character along a reaction coordinate. Nevertheless, it is very promising that we have found a scheme that qualifies as a theoretical model chemistry and that gives satisfactory results for the first five excitation energies of ten different open-shell molecules, all treated with the same procedure and the same three parameters. One big advantage of the scheme is that it is noniterative, whereas other schemes one might propose can involve repeating various steps until the orbitals are sufficiently refined. The ABC scheme has only three steps and three parameters. We have demonstrated the success of the scheme for two different kinds of post-MCSCF calculation, namely, MS-CASPT2 and MC-PDFT. The mean unsigned error over 50 excitations is only 0.19 eV for MS-CASPT2 and 0.14 eV with the MC-PDFT functionals; to put these in context, we note that we have previously[201, 202] stated that any method that gives an MUD of 0.30 eV or better for electronic excitation is “successful”.

Some key issues in selecting an active space are determining what hole and particle orbitals to include, how to obtain initial guesses for the SCF procedures that will place these hole and particle orbitals in the active space where they can be optimized, and what states to average over during the optimization. Although these decisions are often made on a

Table 4.8: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for BO

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2\Pi$	$(1\pi_{\mathrm{x}})^1(5\sigma)^2$	3.542	–0.55	–0.16	–0.22	0.001
2	$1^2\Pi$	$(1\pi_{\mathrm{y}})^1(5\sigma)^2$	3.542	–0.55	–0.16	–0.22	0.001
3	$2^2\Sigma^+$	$(4\sigma)^1(5\sigma)^2$	5.549	–0.22	–0.02	–0.07	0.064
4	$2^2\Pi$	$(5\sigma)^0(2\pi_{\mathrm{y}})^1$	7.149	0.14	–0.39	–0.41	0.141
5	$2^2\Pi$	$(5\sigma)^0(2\pi_{\mathrm{x}})^1$	7.149	0.14	–0.39	–0.41	0.141
			MSD	–0.21	–0.22	–0.27	0.07
			MUD	0.32	0.22	0.27	0.07

case-by-case basis by using previous knowledge of the excitation spectra or by physical reasoning about what kind of states should be in the low-energy spectrum, the results can be very sensitive to the choices made, and case-by-case methods do not constitute a model chemistry. We have shown here that one can also obtain successful active spaces by a procedure that proceeds automatically after one has specified three parameters: A for selecting the hole orbitals, B for selecting the particle orbitals, and C for determining the states over which to average.

Kohn–Sham DFT has made great progress in the last 30 years, and a recent survey[203] through the end of the year 2014 found 114138 documents matching DFT-related index terms. This is one illustration of how important user-friendly computational methods are. However, KS-DFT, like almost all scientific methods, can also be a dangerous tool in the sense that the nonexpert user can possibly generate meaningless results. Multireference calculations can be even more dangerous in this respect because the calculations are more complex. The automatic scheme presented in this chapter cannot prevent the generation of meaningless results, if it is misused, but it is a step toward making the calculations more user-friendly.

Table 4.9: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for N_2^+

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2\Pi_u$	$(1\pi_{u,x})^1(3\sigma_g)^2$	1.316	0.40	–0.008	–0.051	0.18
2	$1^2\Pi_u$	$(1\pi_{u,y})^1(3\sigma_g)^2$	1.316	0.40	–0.008	–0.051	0.18
3	$2^2\Sigma_u^+$	$(2\sigma_u)^1(3\sigma_g)^2$	3.219	0.55	0.018	0.071	–0.001
4	$1^2\Pi_g$	$(3\sigma_g)^0(1\pi_{u,x})^1$	8.756	0.25	–0.048	–0.013	0.020
5	$1^2\Pi_g$	$(3\sigma_g)^0(1\pi_{u,y})^1$	8.756	0.25	–0.048	–0.013	0.020
			MSD	0.37	–0.02	–0.01	0.08
			MUD	0.37	0.03	0.04	0.08

Although we have presented the ABC scheme as a way to carry out practical applications, it could also serve another role; in particular, it could be useful as a way to generate data for machine learning algorithms that would help us to tackle even more complex systems.

Table 4.10: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CO^+

	state		MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2\Pi$	$(1\pi_x)^1(5\sigma)^2$	3.235	0.085	–0.051	–0.13	0.13
2	$1^2\Pi$	$(1\pi_y)^1(5\sigma)^2$	3.235	0.085	–0.051	–0.13	0.13
3	$2^2\Sigma^+$	$(4\sigma)^1(5\sigma)^2$	5.675	0.51	0.23	0.20	0.22
4	$2^2\Pi$	$(5\sigma)^0(2\pi_y)^1$	9.240	0.099	–0.46	–0.46	0.008
5	$2^2\Pi$	$(5\sigma)^0(2\pi_x)^1$	9.240	0.099	–0.46	–0.46	0.008
			MSD	0.17	–0.16	–0.20	0.10
			MUD	0.17	0.25	0.28	0.10

Table 4.11: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for C_2H_3

	state		MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	$1^2A''$	$(7a')^2(1a'')^1$	3.038	–0.30	–0.072	–0.12	0.20
2	$2^2A''$	$(7a')^0(2a'')^1$	4.742	0.86	–0.21	–0.23	0.29
3	$2^2A'$	$(1a'')^1(2a'')^1$	5.354	–0.10	–0.29	–0.28	–0.009
4	$3^2A'$	$(7a')^0(8a')^1$	6.301	–0.046	0.094	0.058	0.48
5	$4^2A'$	$(7a')^0(10a')^1$	6.974	–0.095	–0.19	–0.25	0.40
			MSD	0.06	–0.13	–0.16	0.27
			MUD	0.28	0.17	0.19	0.27

Table 4.12: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CH_2N

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	1^2B_2	$(1b_2)^1(2b_1)^2$	3.822	0.22	–0.12	–0.13	0.18
2	1^2A_1	$(5a_1)^1(2b_1)^2$	4.320	0.14	0.078	0.082	0.088
3	2^2B_2	$(2b_1)^0(2b_2)^1$	4.472	0.58	–0.43	–0.43	0.20
4	2^2B_1	$(1b_2)^1(2b_2)^1$	6.090	–0.229	–0.055	–0.047	0.13
5	2^2A_1	$(2b_1)^0(6a_1)^1$	6.904	0.27	0.36	0.33	0.54
			MSD	0.20	–0.03	–0.04	0.23
			MUD	0.29	0.21	0.20	0.23

Table 4.13: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CH_2O^+

state			MR-CISD+Q	CASSCF	tPBE	trevPBE	MS-CASPT2
1	1^2B_2	$(2b_1)^2(1b_2)^1$	3.688	–0.12	0.20	0.14	0.10
2	1^2A_1	$(5a_1)^1(2b_1)^2$	5.158	0.60	0.028	0.087	0.059
3	2^2B_2	$(2b_1)^0(2b_2)^1$	5.559	–0.024	–0.40	–0.42	–0.005
4	2^2B_1	$(1b_1)^1(2b_1)^2$	6.128	0.34	0.21	0.24	0.20
5	3^2B_1	$(1b_2)^1(2b_2)^1$	7.853	0.26	–0.47	–0.43	–0.068
			MSD	0.21	–0.09	–0.07	0.06
			MUD	0.27	0.26	0.26	0.09

Table 4.14: Signed Errors (eV), Mean Signed Deviation (MSD), and Mean Unsigned Deviation (MUD) of MC-PDFT and MS-CAPST2 Excitation Energies with Respect to MR-CISD+Q for CH_2O^+

state	CASSCF	tPBE	trevPBE	MS-CASPT2
1	0.29	0.11	0.14	0.13
2	0.48	0.13	0.17	0.14
3	0.55	0.22	0.23	0.11
4	0.38	0.20	0.20	0.16
5	0.38	0.30	0.30	0.17
av[a]	0.42	0.19	0.21	0.14

[a]The last row gives the average over all 50 excitation energies.

Chapter 5

Automatic Active Space Selection for Calculating Electronic Excitation Energies Based on High-Spin Unrestricted Hartree-Fock Orbitals

5.1 Introduction

Excited states are usually inherently multiconfigurational, and the most widely used method for generating the multiconfigurational reference wave functions is complete active space self-consistent field theory[145] (CASSCF). A CASSCF calculation does not provide quantitative energetics on its own. It is usually used as the first step of a two-step calculation in which the second step is a configuration interaction,[41, 42, 187, 188, 189] perturbation theory[15, 75, 16, 17, 204, 205, 206, 207, 208, 209] (e.g., multi-state complete active space second-order perturbation theory (MS-CASPT2),[75] or multiconfiguration pair-density functional theory[19, 20] (MC-PDFT) calculation with the CASSCF wave function serving

as a reference function. Because CASSCF is multiconfigurational, such calculations are usually called multireference calculations, and their results depend on the composition of the active space. An active space is determined by both the number of active electrons and the quantity and identity of active orbitals, but the most troublesome aspect of using an active space method is usually finding or choosing the best active space orbitals, especially the orbitals that are unoccupied in the most dominant configuration. One sometimes finds a strong dependence of the results of multireference calculations on this orbital choice.

The difficulty of selecting an active space for such calculations has been recognized for a long time, and therefore there is great interest in developing automated schemes for selecting the active orbitals so that one can define a model chemistry[186] that can be tested systematically. Originally, the goal of automating the selection of the active space was seen as a desirable but perhaps unattainable objective,[210] but recent work has seen considerable progress in this area. A variety of methods are in use, and they are employed by choosing appropriate starting orbitals for the SCF iterations (we will follow the usual convention of calling these "guess orbitals"); some examples are natural orbitals[169, 211, 212, 213, 214, 215] (NOs, from a previous calculation) whose occupation numbers range within a threshold, correlated participating orbitals[80, 47, 147] (CPOs), entangled orbitals from density matrix renormalization group[216, 217] (DRMG) calculations[218, 183, 219, 220] or other calculations, dominant transition orbitals from a large RASSCF calculation[221] (as in the ABC scheme), or other schemes.[182, 222, 182, 223, 224]

The present work attempts to make further progress on this problem by reducing the cost of selecting active orbitals for multireference methods while maintaining good performance. Two very recent automatic active space selection methods, one using entangled orbitals from DMRG calculations and the other being the ABC scheme, require excited-state calculations, in which electrons are excited to a large number of candidate virtual orbitals, to find active orbitals. This is because the desired active orbitals that are unoccupied or occupied with

only a small occupation number are not well optimized in a ground-state calculation, and therefore, the orbital energies are in a wrong order for use in excited-state calculations, or the desired orbital is not even present or is present only in the sense that it may be recovered by taking a linear combination of canonical orbitals. Therefore, initialization of an excited-state calculation may require searching for and defining active orbitals within a large set of virtual orbitals, as if trying to find and identify an unknown item by finding shreds of it in a heap containing fragments of many items.

Our goal is to simplify and automatize the troublesome step of finding starting orbitals for the CASSCF process. We propose to do this by using a high-spin-state (HS-state) unrestricted Hartree–Fock[54] (UHF) calculation to find orbitals to serve as guesses to start the CASSCF calculation. The motivation for this idea is that the HS-state UHF calculation optimizes a set of low-energy singly occupied orbitals that may be good starting points for the singly occupied orbitals into which an electron is excited in low-spin transitions to excited states. In contrast, a ground-state-singlet calculation has no occupancy of these orbitals, and consequently, they are not optimized in such a calculation. HS states are used in other methods as well, for example, spin-flip time-dependent density functional theory[225, 226] (SF-TD-DFT), spin-flip equation-of-motion coupled cluster theory with single and double excitations [227] (SF-EOM-CCSD), and spin-flip configuration interaction[228] (SF-CI), but the purpose of using open-shell spin states in those methods is to allow a convenient way to account for multiconfigurational character. It has been previously noted that HS-state UHF orbitals can be used for generating valence virtual orbitals,[229] and our results will show that some needed Rydberg orbitals can also be occupied in a HS state.

We use the active space generated from HS-state UHF calculations for state-averaged CASSCF calculations, whose wave functions are used as references for MS-CASPT2 and MC-PDFT calculations. The state-averaged CASSCF calculations are denoted SA(N)-CASSCF, where N is the number of states averaged.

5.2 ABC2 Sceheme

Our first automatic scheme was called ABC, and the new automatic scheme is called ABC2. As was the case for the ABC scheme, the ABC2 scheme employs three parameters, and we will call them A, B, and C, similar to the way that we named the three parameters of the ABC scheme. (Of course the parameters have new meanings in ABC2.) We neglect spin–orbit coupling; therefore, total electron spin S is a good quantum number. A key step in ABC2 is to promote A down-spin electrons into the up-spin manifold to obtain (for a closed-shell ground state) a state with M equal to ($2A$ + 1). This will be called the HS state.

Using a singlet molecule as an example, the ABC2 scheme has three steps:

In step 1, we run a UHF calculation for the HS state and get its NOs. The A NOs that are doubly occupied in the ground state but singly occupied in the HS state calculations are called the promotion orbitals, and the A NOs that are unoccupied in the ground state but singly occupied in the HS state are called the particle orbitals. The A promotion orbitals, the A particle orbitals, and ($B - A$) most correlated doubly occupied orbitals from the HS UHF calculation are used in the state-averaged CASSCF calculation in the second step. The most correlated doubly occupied orbital is the one whose NO occupation number is greater than 1 but deviates the most from 2.

In the second step, we use these UHF NOs as the guess orbitals for an SA(N)-CASSCF calculation in which we set $N = N_{\mathrm{EE}} + 1 + C$. (Notice that N_{EE} + 1 equals the desired number of excitation energies plus the ground state, and therefore, C is the number of "extra" states included so that the desired lowest N_{EE} excitations are not missed.) All averaged states are given the same weight in the average. In the SA(N)-CASSCF calculation for a closed-shell ground state, we set the number of active electrons equal to $2B$ and the number of active orbitals equal to $B + A$. The SA(N)-CASSCF wave function obtained in step 2 is the reference orbital for an MC-PDFT calculation or an MS-CASPT2 calculation. We take the

lowest N_{EE} excitation energies as the result. As stated in the Introduction, in this chapter, the target N_{EE} is 2.

Because the new scheme can only occupy B more orbitals to find unoccupied orbitals for active space calculations, it should not be applied to systems where more active orbitals are needed. To predict whether the excitation energies are reliable, we check whether both of the excitation energies are with 1.1 eV of the CASSCF prediction in step 3. If not, the method is considered to not make a prediction of excitation energies but rather to make a recommendation to improve the active space.

Next consider a doublet ground state. For a system with n up-spin electrons and n_α – 1 down-spin electrons in its ground state, the multiplicity equals 2. Promoting A down-spin electrons into the up-spin manifold gives a state with M equal to (2A + 2). This will again be called the HS state. In the SA(N)-CASSCF calculation, we set the number of active electrons equal to 2B + 1 and the number of active orbitals equal to B + A + 1. Thus, for doublet ground states, we have 2B + 1 active electrons in A + B + 1 active orbitals. The active orbitals include A particle orbitals, A promotion orbitals, the singly occupied orbital of the ground spin state, and (B – A) most correlated doubly occupied orbitals from the HS calculation.

The generalization to an arbitrary ground spin state is as follows. For a system with n_α up-spin electrons and n_β down-spin electrons in its ground state, the multiplicity equals (1 + n_α – n_β). Promoting A down-spin electrons into the up-spin manifold yields the HS state with M equal to (2A + 1 + n_α – n_β). In the SA(N)-CASSCF calculation, we set the number of active electrons equal to 2B + n – n and the number of active orbitals equal to B + A + n_α – n_β. The active orbitals include A particle orbitals, A promotion orbitals, (n_α – n_β) singly occupied orbitals of the ground spin state, and (B – A) most correlated doubly occupied orbitals in the HS calculation.

After considerable experimentation, we chose A = 2, B = 3, and C = 2 as default values.

In this chapter, we test only the default values because our goal is to define a full theoretical model chemistry for the prediction of the two lowest excited states (i.e., N_{EE} = 2), and a key feature of a full theoretical model chemistry is that, having selected a target, "particular procedures for particular molecules or particular symmetries should be avoided."[186]

A few refinements (independent of symmetries and particular molecules) have to be added to handle degeneracies. If the default value of A would excite only one of a pair of degenerate promotion orbitals or would occupy only one of a pair of degenerate particle orbitals, we increase A by 1 to avoid unequal treatment of a degenerate pair of orbitals. If the default value of B would select only one of a pair of degenerate doubly occupied orbitals or result in only one of a pair of degenerate orbitals in the active space in the converged CASSCF wave function in step 2, we increase B by 1 to avoid unequal treatment of a degenerate pair of orbitals. If the default value of C would select only one of a pair of degenerate states, we increase C by 1 to avoid unequal treatment of a degenerate pair of states.

5.3 Computational Details

We tested the performance of MC-PDFT and MS-CASPT2 with active spaces selected by ABC2. The MS-CASPT2 calculations used an ionization energy–electron affinity (IPEA) shift[148] of 0.25 au. The MC-PDFT calculations used the translated PBE (tPBE) on-top density functional.[62] We used the jun-cc-pVTZ basis set.[230]

All excitation energy calculations were performed in OpenMolcas v.18.09.[231] Some geometry optimization calculations were performed in Gaussian 16.[232] The references for geometries and excitation energies are shown in Table 5.1. The geometries are in Tables S2 and S3 in the Supporting Information.

Table 5.1: References for Molecular Geometries and Excitation Energies

molecule	geometry	excitation energy reference
Doublets		
CH_3 (methyl radical)	[195]	[221]
NH_2 (amino radical)	[196]	[221]
C_2H (ethynyl radical)	[197]	[221]
BO (boron oxide)	[30]	[221]
CN (cyanide radical)	[30]	[221]
CO^+ (carbon monoxide cation)	[30]	[221]
N_2^+ (dinitrogen cation)	[30]	[221]
C_2H_3 (vinyl radicalc)	[198]	[221]
CH_2N (methyleneamino radical)	[199]	[221]
CH_2O^+ (formaldehyde cation)	[38]	[221]
Singlets without π Bonds		
H_2O (water)	[233]	[234]
H_2S (hydrogen sulfide)	[233]	[233, 235]
HCl (hydrogen chloride)	[233]	[236]
NH_3 (ammonia)	[233]	[237]
Singlets with One or Two π Bonds		
C_2H_2 (acetylene)	[233]	[238]
CH_2N_2 (diazomethane)	[233]	[239]
C_2H_3F (vinyl fluorided)	a	[240]
C_2H_4 (ethylene)	b	[241]
C_3H_4 (cyclopropane)	[233]	[242, 243]
C_4H_6 (trans-butadiene)	[244]	[245, 246]
C_6H_6 (benzene)	[247]	[248]
CH_2O (formaldehyde)	[233]	[249]
CH_2NH (methanimine)	[233]	[233]
CH_3CHO (acetaldehyde)	[233]	[250]
N_2 (dinitrogen)	[233]	[251]
CO (carbon monoxide)	[233]	[252]

^{a}Optimized using the M06-2X[64, 253] density functional with the aug-cc-pVTZ basis set in Gaussian 16. ^{b}Optimized using the M06-L[63, 64] density functional with the aug-cc-pVTZ basis set in Gaussian 16. ^{c}Also called the ethenyl radical. ^{d}Also called fluoroethylene.

5.4 Systems Studied

We studied the first two excited states (N_{EE} is 2) for 10 doublet radicals, 4 singlet molecules without π bonds, and 12 singlet molecules with π bonds.

Table 5.2 also gives the leading terms in the ground-state wave functions and their ground-state configurations. The multireference characters of the ground states were judged by the M diagnostics,[80] which are presented in Table 5.2. The systems that we studied range from being weakly multiconfigurational ($M \leqq 0.01$) to being strongly multiconfigurational ($M \geqslant 0.10$).

5.5 Computational Details

We tested the performance of MC-PDFT and MS-CASPT2 with active spaces selected by ABC2. The MS-CASPT2 calculations used an ionization energy–electron affinity (IPEA) shift[148] of 0.25 au. The MC-PDFT calculations used the translated PBE (tPBE) on-top density functional.[62] We used the jun-cc-pVTZ basis set.[230]

All excitation energy calculations were performed in OpenMolcas v.18.09.[231] Some geometry optimization calculations were performed in Gaussian 16.[232] The references for geometries and excitation energies are shown in Table 5.1. The geometries are in Tables S2 and S3 in the Supporting Information.

Table 5.2: References for Molecular Geometries and Excitation Energies

system	A	B	C	active space[a]	ground-state wave function[b]	M
Doublets						
CH_3	3	3	3	(7,7)	$0.99(2a')^2(1e_x)^2(1e_y)^2(1a''1)^1(3a')^0(2e_x)^0(2e_y)^0$	0.002
NH_2	2	3	3	(7,6)	$0.98(2a_1)^2(1b_1)^2(3a_1)^2(1b_2)^1(4a_1)^0(2b_1)^0$	0.02
C_2H	2	3	3	(7,6)	$0.93(4a_1)^2(5a_1)^1(1\pi_x)^2(1\pi_y)^2(2\pi_y)^0(2\pi_x)^0$	0.09
BO	2	3	2	(7,6)	$0.90(4a_1)^2(1\pi_x)^2(1\pi_y)^2(5a_1)^1(2\pi_y)^0(2\pi_x)^0$	0.07
CN	2	3	2	(7,6)	$0.93(4a_1)^2(5a_1)^1(1\pi_x)^2(1\pi_y)^2(2\pi_y)^0(2\pi_x)^0$	0.11
CO^+	2	3	2	(7,6)	$0.91(4a_1)^2(1\pi_x)^2(1\pi_y)^2(5a_1)^1(2\pi_y)^0(2\pi_x)^0$	0.09
N_2^+	2	3	3	(7,6)	$0.93(2a_{1u})^2(1e_{g,x})^2(1e_{g,y})^2(3a_{1g})^1(1e_{u,x})^0(1e_{u,y})^0$	0.13
C_2H_3	2	3	2	(7,6)	$0.96(5a')^2(6a')^2(1a'')^2(7a')^1(2a'')^0(8a')^0$	0.11
CH_2N	2	3	2	(7,6)	$0.97(4a_1)^2(5a_1)^2(2b_1)^1(1b_2)^2(2b_2)^0(6a_1)^0$	0.10
CH_2O^+	2	3	2	(7,6)	$0.94(5a_1)^2(1b_1)^2(1b_2)^2(2b_1)^1(2b_2)^0(3b_1)^0$	0.07
Singlets without π Bonds						
H_2O	2	3	2	(6,5)	$0.98(1b_1)^2(3a_1)^2(1b_2)^2(2b_1)^0(4a_1)^0$	0.006
H_2S	2	3	2	(6,5)	$0.99(2b_1)^2(5a_1)^2(2b_2)^2(3b_1)^0(6a_1)^0$	0.02
HCl	2	3	2	(6,5)	$0.99(5a_1)^2(2e_{1,x})^2(2e_{1,y})^2(6a_1)^0(7a_1)^0$	0.01
NH_3	3	3	2	(6,6)	$0.98(1e_x)^2(1e_y)^2(3a_1)^2(4a_1)^0(2e_x)^0(2e_y)^0$	0.004
Singlets with One or Two π Bonds						
C_2H_2	3	3	3	(6,6)	$0.97(5a_1)^2(1\pi_{1,x})^2(1\pi_{1,y})^2(6a_1)^0(2\pi_{1,x})^0(2\pi_{1,y})^0$	0.06
CH_2N_2	2	3	2	(6,5)	$0.94(2b_1)^2(1b_2)^2(2b_2)^2(3b_1)^0(3b_2)^0$	0.09
C_2H_3F	2	3	2	(6,5)	$0.98(10a')^2(11a')^2(1a'')^2(2a'')^0(12a')^0$	0.06
C_2H_4	2	3	2	(6,5)	$0.99(1b_{2u})^2(1b_{3g})^2(1b_{3u})^2(1b_{2g})^0(4a_g)^0$	0.06
C_3H_4	2	3	2	(6,5)	$0.99(6a_1)^2(3b_1)^2(3b_2)^2(1a_2)^0(7a_1)^0$	0.03
C_4H_6	2	3	2	(6,5)	$0.98(8a_g)^2(1a_u)^2(1b_g)^2(2a_u)^0(2b_g)^0$	0.07
C_6H_6	2	3	2	(6,5)	$0.96(a_2u)^2(e_{1g})^2(e_{1g})^2(e_{2u})^0(e_{2u})^0$	0.06
CH_2O	2	3	2	(6,5)	$0.96(1b_1)^2(1b_2)^2(2b_1)^2(2b_2)^0(6a_1)^0$	0.09
CH_2NH	2	3	2	(6,5)	$0.94(6a')^2(7a')^2(a_1'')^2(2a'')^0(8a')^0$	0.07
CH_3CHO	2	3	2	(6,5)	$0.96(9a')^2(2a'')^2(10a')^2(3a'')^0(11a')^0$	0.08
CO	3	3	2	(6,5)	$0.96(5a_1)^2(1\pi)^2(1\pi)^2(2\pi)^0(2\pi)^0$	0.05
N_2	2	4	2	(8,6)	$0.96(2a_1u)^2(1e_{g,x})^2(1e_{g,y})^2(3a_1g)^2(1e_{u,x})^0(1e_{u,y})^0$	0.07

[a](Number of electrons, number of orbitals).
[b]Only leading terms are shown.

Systems with C_{3v}, D_{3h}, or linear symmetries may need the degeneracy-related refinements of the last paragraph of section 2. These refinements generate the final values of the three parameters, which are also given in Table 5.2. Systems that have Rydberg orbitals in the resulting active spaces are C_2H_3 (8a′), CH_2N (6a′), CH_2O (6a_1), H_2O (4a_1), H_2S (6a_1), HCl (7a_1), C_2H_3F (12a′), C_2H_4 (4a_g), C_3H_4 (7a_1), CH_2NH (8a′), and CH_3CHO (11a′). The active Rydberg orbitals of CH_2O and C_2H_4 are illustrated in Figures 5.3d and 5.2a.

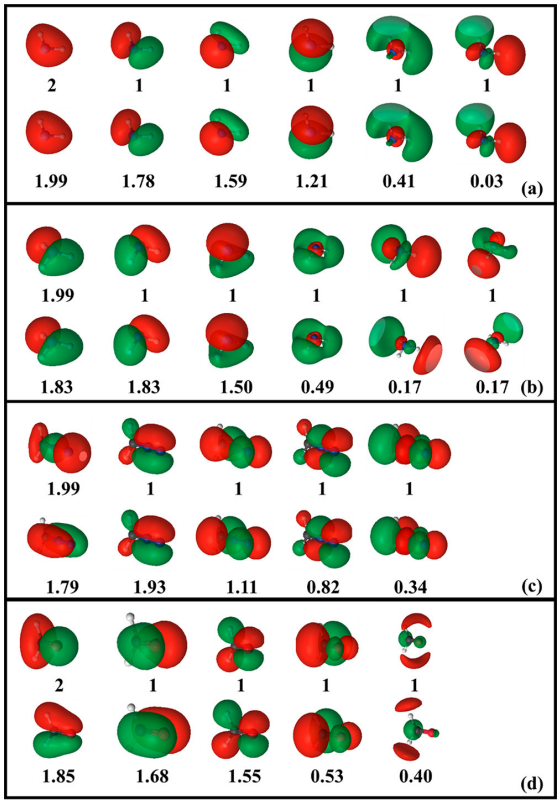

Figure 5.1: HS-state UHF NOs (upper rows) as guess orbitals for active orbitals and converged SA-CASSCF NOs in the active space (lower rows) with the NO occupation numbers (NOON) and state-averaged (SA) NOON labeled below each corresponding orbital for (a) NH_2, (b) NH_3, (c) CH_2N_2, and (d) CH_2O.

5.6 Methods Compared

The results of the new ABC2 scheme will be compared to three other ways of selecting guess orbitals: (i) the ABC scheme, which was explained in a previous paper,[221] (ii) using ground-state UHF NOs or RHF orbitals, which needs no explanation, and (iii) the default orbital scheme built into OpenMolcas. The OpenMolcas default guess orbitals are the eigenvectors of a model Fock matrix, which is generated by projecting the Fock matrix of atomic orbital energies onto a molecular orbital basis that is obtained by diagonalizing the overlap matrix.

5.7 Results and Discussion

All doublet molecules studied here pass the 1.1 eV test of step 3. Among 16 singlet molecules, 9 pass the 1.1 eV test for MC-PDFT and 9 (but not the same 9) pass the 1.1 eV test for MS-CASPT2.

5.7.1 Doublet Excited States

The best estimates for the doublets are taken as the values calculated in our previous work[221] by multireference configuration interaction with single and double excitations and a Davidson correction[43, 190] (MRCI+Q), and these results are given in the last column of Table 3, which also gives results calculated with the ABC2 scheme. We see from Table 3 that ABC2 gives very good excitation energies for the doublet systems. The only exception is CH_2O^+, where the second state of ABC2 is the third state with MRCI+Q.

Table 5.3: Excitation Energies (eV) for the Ten Doublet Systems Calculated by CASSCF, MC-PDFT, and MS-CASPT2 Using the Active Spaces in Table 5.2 and Reference Values and Mean Errors

system	excitation	sym.	CASSCF	MC-PDFT	MS-CASPT2	ref
CH_3	$1a_2'' \rightarrow 3a_1'$	$2A_1'$	5.43	5.78	6.01	5.87
	$1e' \rightarrow 1a_2''$	$2E'$	7.58	6.75	7.29	6.94
NH_2	$3a_1 \rightarrow 1b_2$	$2A_1$	2.20	2.03	2.10	2.09
	$1b_1 \rightarrow 1b_2$	$2B_1$	7.14	6.58	6.58	6.37
C_2H	$1\pi_x \rightarrow 5a_1$	2Π	0.71	0.37	0.82	0.68
	$1\pi_y \rightarrow 5a_1$	2Π	0.71	0.37	0.82	0.68
BO	$1\pi_x \rightarrow 5a_1$	2Π	3.09	3.27	3.49	3.54
	$1\pi_y \rightarrow 5a_1$	2Π	3.18	3.29	3.51	3.54
CN	$1\pi_x \rightarrow 5a_1$	2Π	1.49	1.17	1.44	1.28
	$1\pi_y \rightarrow 5a_1$	2Π	1.49	1.17	1.44	1.28
CO^+	$1\pi_x \rightarrow 5a_1$	2Π	3.23	3.02	3.29	3.24
	$1\pi_y \rightarrow 5a_1$	2Π	3.23	3.02	3.29	3.24
N_2^+	$1\pi_{u,x} \rightarrow 3a_{1,g}$	$2\Pi_u$	1.72	1.33	1.46	1.32
	$1\pi_{u,y} \rightarrow 3a_{1,g}$	$2\Pi_u$	1.72	1.33	1.46	1.32
C_2H_3	$1a'' \rightarrow 7a'$	$2A''$	3.23	3.09	3.22	3.04
	$7a' \rightarrow 2a''$	$2A''$	5.29	4.44	4.79	4.74
CH_2N	$1b_2 \rightarrow 2b_1$	$2B_2$	4.03	3.92	4.08	3.82
	$5a_1 \rightarrow 2b_1$	$2A_1$	4.30	4.39	4.43	4.32
CH_2O^+	$1b_2 \rightarrow 2b_1$	$2B_2$	3.50	3.77	3.77	3.69
	$2b_1 \rightarrow 2b_2$	$2B_2$	5.89	5.24	5.61	5.56a
MSEb			0.13	–0.11	0.12	
MUEb			0.28	0.17	0.13	

[a](Number of electrons, number of orbitals).
[b]Only leading terms are shown.

Among the MC-PDFT results, the maximum absolute error is 0.32 eV, which is obtained for the second $2B_2$ state of CH_2O^+ (the third MRCI+Q excited state), and the minimum absolute error is 0.01 eV, which is obtained for the 2πu state of N_2^+. Among the MS-CASPT2 results, the maximum absolute error is 0.35 eV for the second excited state of CH_3, and the minimum absolute error is 0.01 eV for the first excited state of NH_2. The mean unsigned error (MUE) is only 0.17 eV for MC-PDFT and only 0.13 eV for MS-CASPT2. We conclude that ABC2 is successful for both MC-PDFT and MS-CASPT2 on the doublet systems.

5.7.2 Singlet Excited States

Table 5.4 summarizes the excitation energies of 16 singlet systems, which are further divided into systems without a π bond and those with one or two π bonds.

Table 5.4: Excitation Energies (eV) for the Ten Doublet Systems Calculated by CASSCF, MC-PDFT, and MS-CASPT2 Using the Active Spaces in Table 5.2 and Reference Values and Mean Errors

system	excitation	sym.	CASSCF	MC-PDFT	step 3	MS-CASPT2	step 3	ref
H_2O	$1b_2 \rightarrow 3a_1$	$1B_2$	6.39	7.26	pass	7.95	not	7.41
	$1b_2 \rightarrow 2b_1$	$1A_2$	8.24	8.88		9.87		9.20
H_2S	$2b_2 \rightarrow 4a_1$	$1B_2$	6.23	6.04	pass	6.25	pass	6.12
	$2b_2 \rightarrow 3b_1$	$1A_2$	6.17	6.31		6.43		6.33
HCl	$2e_1 \rightarrow 6a_1$	$1E_1$	7.66	7.81	pass	7.99	pass	7.88
	$2e_1 \rightarrow 6a_1$	$1E_1$	7.66	7.81		7.99		7.88
NH_3	$3a_1 \rightarrow 4a_1$	$1A_1$	5.58	6.33	pass	6.72	not	6.38
	$3a_1 \rightarrow 2e$	$1E_1$	7.13	7.80		8.42		7.90
C_2H_2	$1\pi \rightarrow 2\pi$	$1\Sigma^-$	7.25	6.77	pass	7.06	pass	7.1
	$1\pi \rightarrow 2\pi$	1Δ	7.62	6.86		7.35		7.2
CH_2N_2	$2b_2 \rightarrow 3b_1$	$1A_2$	2.87	3.32	not	3.13	not	3.14
	$2b_2 \rightarrow 3b_2$	$1A_1$	7.36	6.09		8.49		5.90
CH_2O	$2b_1 \rightarrow 2b_2$	$1A_2$	3.75	3.91	pass	3.82	pass	4.07
	$2b_1 \rightarrow 6a_1$	$1B_1$	6.78	7.27		7.29		7.11
C_2H_3F	$\pi(1a'') \rightarrow 3s(12a')$	$1A''$	7.21	7.75	not	7.80	not	6.98
	$\pi(1a'') \rightarrow \pi^*(2a'')$	$1A'$	9.37	5.82		7.54		7.44
C_2H_4	$\pi(1b_{3u}) \rightarrow 3s(4a_g)$	$1B_{3u}$	7.06	7.79	not	7.86	not	7.11
	$\pi(1b_{3u}) \rightarrow \pi^*(1b_{2g})$	$1B_{1u}$	9.06	5.78		7.59		7.60
C_3H_4	$\sigma \rightarrow \pi$	$1B_1$	6.75	6.69	not	5.92	pass	6.45
	$\pi \rightarrow \pi^*$	$1B_2$	7.71	6.30		6.78		7.00
C_4H_6	$b_g \rightarrow a_u$	$1B_u$	7.35	4.28	not	5.94	not	5.92
	$b_g \rightarrow b_g$	$1A_g$	6.35	6.75		6.59		6.39
C_6H_6	$e_{1g} \rightarrow e_{2u}$	$1B_{2u}$	5.72	5.29	not	4.93	not	4.9
	$e_{1g} \rightarrow e_{2u}$	$1B_{1u}$	7.12	5.37		6.10		6.2
N_2	$n \rightarrow \pi^*$	1Π	9.71	8.97	pass	9.21	pass	9.31
	$n \rightarrow \pi^*$	1Π	9.71	8.97		9.21		9.31
CO	$n \rightarrow \pi^*$	1Π	9.71	7.81	not	8.62	pass	8.51
	$n \rightarrow \pi^*$	1Π	9.71	7.81		8.62		8.51
CH_2NH	$7a' \rightarrow 2a''$	$1A''$	5.15	5.08	pass	5.24	pass	5.21
CH_3CHO	$10a' \rightarrow 3a''$	$1A''$	4.14	4.31	pass	4.16	pass	4.27
MSE for cases that pass step 3[a]				–0.06		0.09		
MUE for cases that pass step 3[a]				0.20		0.17		

[a]Mean signed error (MSE) and mean unsigned error (MUE) are computed for systems that pass the 1.1 eV test in step 3. Only one state of a pair of degenerate states is used for MSE and MUE calculations when the pair of degenerate states correspond to the second and third excited states. This applies to NH_3 and C_2H_2.

The MC-PDFT calculations pass the step 3 diagnostic for all four singlet systems without π bonds and five of the systems with π bonds. This implies that the ABC2 scheme with MC-PDFT does not make a reliable prediction for all of the systems with π bonds, but an advantage of the step 3 test is that the method has an internal built-in criterion for whether its results should be considered reliable. The large differences between CASSCF and MC-PDFT energies are mainly from $\pi \rightarrow \pi^*$ excitations, and MC-PDFT tends to underestimate these excitation energies. Averaging MC-PDFT excitation energies over those where ABC2 makes a prediction, that is, where the results pass the step 3 test, we get 0.20 eV for the MUE.

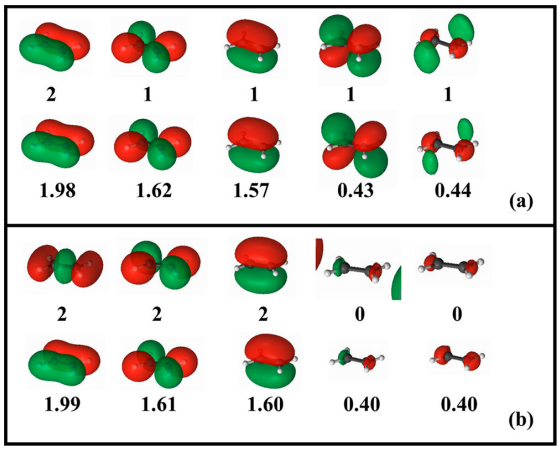

Figure 5.2: HS-state UHF NOs (upper rows) as guess orbitals for active orbitals and converged SA-CASSCF NOs in the active space (lower rows) with the NOON and SA-NOON labeled below each corresponding orbital for C_2H_4 (ethylene), from (a) HS-state UHF calculations and (b) low-spin-state UHF calculations.

The MS-CASPT2 calculations with the ABC2 scheme pass the step 3 diagnostic for two of the singlets without π bonds and for seven of those with π bonds. Thus, again, the method makes a prediction for only a subset of the states. For the π-bonded systems, the large differences between MS-CASPT2 and CASSCF are again from $\pi \rightarrow \pi^*$ excitations, but the errors are mainly from the $\pi \rightarrow$ Rydberg excitations. The main source of error is also from

Rydberg excitations even for systems without π bonds.

The MSEs of MC-PDFT and MS-CASPT2 for systems that pass the step-3 diagnostic are less than 0.1 eV, suggesting that among these systems both MC-PDFT and MS-CASPT2 have an even distribution of results above and below excitation energies (i.e., the errors are not systematic). The MUEs are both less than 0.30 eV, which we consider is the criterion we have labeled as successful in previous work[201, 202, 254] on approximate predictions of excitation energies.

The MC-PDFT results that fail the 1.1 eV diagnostic do in fact give large errors for excitation energies, showing that the reliability test is successful.

In a few cases, the MS-CASPT2 results are reasonably accurate even when the test failed; for example, MS-CASPT2 gives 7.86 eV ($\pi \rightarrow$ 3s, vs 7.11 eV experimentally) and 7.59 eV ($\pi \rightarrow \pi^*$, vs 7.60 eV experimentally) for the first two excitation energies of ethylene. The $\pi \rightarrow \pi^*$ excitation energy very well reproduces the experimental result, and the MUE over the two transition energies is 0.38 eV. Encouragingly, the $\pi \rightarrow \pi^*$ excitation energy is closer to the experimental result than is the MS-CASPT2 result in the classic first MS-CASPT2 paper,[75] where the calculations yielded 7.98 eV. Another encouraging comparison is to the results obtained with the more expensive entangled orbital scheme, which yielded 7.52 eV for the $\pi \rightarrow$ 3s transition and 8.31 eV for the $\pi \rightarrow \pi^*$ transition of ethylene, giving an MUE over the two states of 0.52 eV, which is larger than the present 0.38 eV.

5.7.3 Comparison to the ABC scheme

In Table S1, we summarize the first two excitation energies of MC-PDFT calculations with active spaces generated by the ABC scheme (note that with the ABC scheme we used the aug-cc-pVTZ[255, 44] basis set). The four error quantifiers of ABC2, namely, the mean signed error (MSE), MUE, minimum absolute error, and maximum absolute error, are very similar to those of the ABC scheme, although the minimum absolute error and maximum

absolute error are for different systems. This means that the ABC scheme and the ABC2 scheme perform about equally as well for the doublet systems.

However, when we test the two schemes in a broader range of systems, we find that the ABC2 scheme is more robust than the ABC scheme. A comparison is summarized in Table 5.5, which shows the performance of the ABC scheme and the ABC2 scheme on some of the large systems that pass the step 3 diagnostic for both MC-PDFT and MS-CASPT2. We see that both MC-PDFT and MS-CASPT2 are improved by using the ABC2 scheme as compared to using the ABC scheme. The main problem with the ABC scheme is that the only occupied orbital included in the active space is the HOMO for the singlets. The lack of more occupied orbitals in the active spaces results in inaccurate MC-PDFT results. Although the previous ABC scheme could be improved by including more occupied orbitals in the active space, the ABC2 scheme would still be favored because of its lower cost, which is discussed below.

Table 5.5: Excitation Energies (eV) of CASSCF, MC-PDFT, and MS-CASPT2 for Some Large Molecules That Pass the Step 3 Test and Their MSE (eV) and MUE (eV)

system	ABC			ABC2		
	CASSCF	MC-PDFT	MS-CASPT2	CASSCF	MC-PDFT	MS-CASPT2
C_2H_2	2.92	2.97	3.24	3.23	3.09	3.22
	5.65	4.53	5.03	5.29	4.44	4.79
CH_2N_2	4.30	3.70	4.00	4.03	3.92	4.08
	4.57	4.40	4.41	4.30	4.39	4.43
CH_2O^+	3.65	3.89	3.79	3.50	3.77	3.77
	5.58	5.16	5.55	5.89	5.24	5.61
C_2H_2	7.16	6.70	7.05	7.23	6.69	6.98
	7.66	6.66	7.37	7.67	6.73	7.20
CH_2O	1.98	2.07	3.41	3.75	3.91	3.82
	5.07	5.55	6.96	6.78	7.27	7.29
CH_2NH	3.67	3.63	4.76	5.15	5.08	5.24
CH_3CHO	3.36	2.99	3.69	4.14	4.31	4.16
MSE	–0.38	–0.66	–0.07	0.07	–0.11	0.04
MUE	0.74	0.70	0.24	0.24	0.19	0.12

5.7.4 Comparison to other guess schemes

We next compare to MC-PDFT calculations for C_2H_3, CH_2N, CH_2O^+, C_2H_2, CH_2O, CH_2NH, and CH_3CHO using ground-spin-state UHF orbitals (with M_S = 0 for singlets and M_S = 1/2 for doublets) and the OpenMolcas default guess orbitals. These two methods use the same size of active space (for each molecule) as that predicted by the ABC2 scheme to show that the HS UHF orbitals are more robust. The results are presented in Table 5.6, and the $\langle S^2 \rangle$ values for the ground-state doublets are presented in Table S4. We see that some calculations using these two sets of orbitals do not converge to correct states; some of them result from the absence of orbitals to which electrons are excited in some of the low-lying excited states. That is a disadvantage of those methods as compared to ABC2, and that kind of problem was one of the motivations for proposing ABC2. We calculated the MUEs for these two older methods without including the states to which the wave functions are not converged. The MUEs show that even when the older methods are averaged over only the subset of cases for which the$_y$ do not fail, they have a higher MUE than the ABC2 method.

Table 5.6: Excitation Energies (eV), MSE, and MUE of CASSCF, MC-PDFT, and CASPT2 for Systems in Table 5.5, with the Same Number of Active Electrons and Active Orbitals as the ABC2 Scheme but with Orbital Identities Coming from OpenMolcas Default and Ground-Spin-State HF Calculations, Respectively

system	OpenMolcas default			ground-state HF[a]		
	CASSCF	MC-PDFT	MS-CASPT2	CASSCF	MC-PDFT	MS-CASPT2
C_2H_3.30	2.96	3.19	3.30	2.96	3.19	
	5.42	4.35	4.84	5.42	4.35	4.84
CH_2N	4.15	3.58	wrong state	4.16	3.52	3.87
	4.99	4.32	wrong state	wrong state		
CH_2O^+.75	3.75	3.79	3.52	3.93	3.80	
	6.02	5.34	5.37	6.03	5.27	5.54
C_2H_2.23	6.69	6.98	6.59	6.61	7.24	
	7.67	6.73	7.12	wrong state		
CH_2O	4.52	3.53	3.62	3.74	3.81	3.89
	wrong state, missing 3s orbital			6.98	7.39	7.48
CH_2NH	5.30	5.07	5.10	wrong state, missing $\pi*$ orbital		
CH_3CHO	4.86	3.86	3.93	4.14	4.31	no convergence
MSE	0.38[b]	–0.26[b]	–0.10[c]	0.05[d]	–0.14[d]	0.09[e]
MUE	0.38[b]	0.27[b]	0.18[c]	0.34[d]	0.26[d]	0.14[e]

[a]The ground-state wave function is an RHF wave function for singlets and UHF for doublets. See Table S4.
[b]Excluding the second excited state of CH_2O.
[c]Excluding both excited states of CH_2N and the second excited state of CH_2O.
[d]Excluding the second excited state of CH_2N, the second excited state of C_2H_2, and the first excited state of CH_2NH.
[e]Excluding the second excited state of CH_2N, the second excited state of C_2H_2, the first excited state of CH_2NH, and the first excited state of CH_3CHO.

We compared the UHF NOs for a HS state and RHF orbitals for a low-spin state for C_2H_4 in Figure 2. We see that with the present choice of basis the first two RHF virtual orbitals do not include a π^* orbital, while the HS UHF calculation gives a π^* orbital and a Rydberg orbital, which is what is needed for further optimization in the following SA(N)-CASSCF calculation. This is an example of how the HS UHF calculation gives better guesses for active orbitals than does a ground-state calculation.

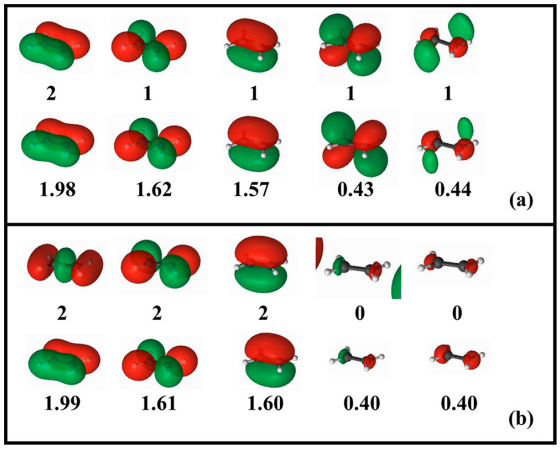

Figure 5.3: HS-state UHF NOs (upper rows) as guess orbitals for active orbitals and converged SA-CASSCF NOs in the active space (lower rows) with the NOON and SA-NOON labeled below each corresponding orbital for C_2H_4 (ethylene), from (a) HS-state UHF calculations and (b) low-spin-state UHF calculations.

On systems that both MC-PDFT and CASPT2 pass the step 3 diagnostic, we tested the scheme with the two most correlated doubly occupied orbitals included in the active space; this corresponds to B = 4. For these systems, the ABC2 scheme with B = 3 gives 0.06, 0.11, and 0.08 eV for the MSE of CASSCF, MC-PDFT, and MS-CASPT2, respectively, and 0.26, 0.16, and 0.13 eV for the MUE of the three calculations, respectively. Each of these six quantities is different from the corresponding ones with B = 4 (shown in Table 7) by at most only 0.01 eV. It means that using 3 for B is adequate for the ABC2 scheme, and the results are stable with respect to B.

Table 5.7: Excitation Energies (eV) and Mean Errors Calculated with B = 4 for Systems[a] That Pass Step 3 for Both MC-PDFT and CASPT2 in Tables 5.3 and 5.4

system	CASSCF	MC-PDFT	MS-CASPT2
CH_3	5.43	5.78	6.03
	7.58	6.74	7.31
NH_2	2.20	2.03	2.09
	7.14	6.58	6.58
C_2H	0.90	0.56	0.85
	0.90	0.56	0.85
BO	3.01	3.25	3.48
	3.01	3.25	3.48
CN	1.50	1.24	1.41
	1.50	1.24	1.41
CO^+	3.23	3.02	3.29
	3.23	3.02	3.29
N_2^+	1.62	1.31	1.46
	1.62	1.31	1.46
C_2H_3	3.21	3.08	3.21
	5.29	4.46	4.78
CH_2N	4.14	3.73	3.97
	4.46	4.33	4.36
CH_2O^+	3.47	3.76	3.75
	5.81	5.24	5.59
C_2H_2	7.25	6.72	6.90
	7.64	6.80	7.05
CH_2O	3.77	3.92	3.83
	6.76	7.24	7.29
H_2S	6.27	6.07	6.23
	6.29	6.30	6.40
HCl	7.68	7.80	7.98
	7.68	7.80	7.98
CH_2NH	5.17	5.07	5.22
CH_3CHO	4.16	4.31	4.17
MSE	0.07	–0.11	0.07
MUE	0.27	0.14	0.12

[a]N_2 is excluded here because B is already 4 in Table 5.4 to include a pair of degenerate orbitals.

5.7.5 Timing

We use the largest molecule among the systems that we have tested, namely, acetaldehyde, as an example to compare ABC2 to ABC for the active spaces generated, the MC-PDFT and MS-CAPST2 errors compared with experiment, and the timings for the orbital guess step and the CASSCF+MC-PDFT and CASSCF+MS-CASPT2 processes. The comparison is given in Table 5.8. For the step of getting orbital guesses, ABC2 includes getting integrals that are used throughout the calculations and the HS UHF calculation, while the ABC scheme includes calculating the integrals, a ground-state CASSCF calculation (equivalent to a restricted Hartree–Fock calculation for acetaldehyde), and a state-averaged RASSCF calculation to find transition orbitals in the second step. The CASSCF+MC-PDFT/MS-CASPT2 process includes an SA(5)-CASSCF calculation with the active space generated by the scheme and a subsequent MC-PDFT/MS-CASPT2 calculation. We see that ABC2 needs only 1/3 of the time that is consumed by the ABC scheme for MS-CAPST2 and one-half of the time for MC-PDFT. The ABC2 scheme yields a more reasonable active space with a more accurate result at a lower cost. We conclude that ABC2 is preferred over the previous ABC scheme.

Table 5.8: Comparison of Automated Schemes for Calculating the First Excited State of CH_3CHO

	ABC	ABC2
N in SA(N)-CASSCF	5	5
active space	(2,4)	(6,5)
MC-PDFT error (eV)	–1.28	0.04
MS-CASPT2 error (eV)	–0.59	–0.11
timing (seconds) to get orbital guesses	555	107
timing (seconds) for CASSCF+MC-PDFT	208	212
timing (seconds) for CASSCF+MS-CASPT2	165	177

5.7.6 Active spaces in other methods

The active space selection scheme we proposed here could be useful for any methods attempting to calculate electronic excitation energies by a post-MCSCF step. This includes not only multireference perturbation theory and MC-PDFT, both of which are considered here, but also multireference configuration interaction.[256, 257, 258, 188, 42, 259] Active spaces can also be used with single-reference coupled-cluster[260, 261, 262] (CC) calculations and equation-of-motion coupled-cluster calculations[48, 49] (EOM-CC) to restrict excitations to an active space to reduce computational cost,[263, 264, 265, 266, 267] but active space selection is less important in those single-reference calculations because their costs grow polynomially rather than exponentially with the numbers of active orbitals.

5.8 Conclusion

Our goal was to define, refine, and test a model chemistry for selecting orbitals for excited-state calculations beginning with a CASSCF wave function. We therefore proposed a well-defined scheme that we call the ABC2 scheme for automating active space selection for excited-state calculations when the goal is to predict the first two spin-conserving excitation energies; the scheme is based on using guess orbitals from a HS-state UHF calculation. We proposed a set of parameters such that the scheme is independent of symmetries and particular molecules and it satisfies Pople's criterion for a testable model chemistry. A key step in the ABC2 scheme is step 3, which provides a test to be passed for the results to be considered reliable. Of the 24 systems studied (10 doublet systems and 16 singlet systems), 19 (a different subset of 19 systems for MC-PDFT and MS-CASPT2) passed the test, and for these systems, we obtained very good results with the recommended parameters both for MC-PDFT calculations with the tPBE functional and for MS-CASPT2 calculations with the IPEA shift. We also showed that ABC2 is more robust than our previous ABC scheme,

than the use of low-spin-state guess orbitals, or than the use of the default orbitals provided by the built-in algorithm of OpenMolcas.

Chapter 6

Extended multi-state Pair-Density Functional Theory

6.1 Introduction

Kohn–Sham density functional theory (KS-DFT)[8] has been successful in treating many chemical problems, but it is less accurate for treating inherently multi-configurational electronic states – which are called strongly correlated states – than for treating states well represented by a single Slater determinant – which are called weakly correlated.[269] Strong correlation usually arises from near degeneracy of two or more states, and excited electronic states are usually strongly correlated, and often they strongly interact with other states. Thus, the accurate treatment of strongly correlated states is necessary for spectroscopy and photochemistry.[270, 271] Furthermore, the accurate treatment of strongly correlated sets of states is also required to properly describe magnetic effects.[272, 273]

Although KS-DFT has lower accuracy for strongly correlated states than for weakly correlated ones, for large molecules it is much less expensive than wave function theory (WFT)

methods of comparable accuracy. We have proposed multi-configuration pair-density functional theory (MC-PDFT) as a method that builds on a multi-configurational self-consistent-field (MCSCF) reference wave function and is more innately suitable for strongly correlated systems than KS-DFT; MC-PDFT also has the advantage of being computationally less expensive compared with WFT methods in terms of computer time and memory with comparably accurate treatments of correlation energy.[19, 20] We refer the reader to a recent review article[21] that compares PDFT to other ways to combine wave function methods and density functional methods for excited-state calculations.

When states are nearly degenerate and have the same symmetry, they interact strongly with each other, and they should be treated by a method that gives the correct topography[271] of adiabatic potential energy surfaces (PESs) at conical intersections; such methods are called multi-state (MS) methods. For example, in WFT, multireference Møller–Plesset perturbation theory[274] is a state-specific method because it calculates the final approximation of the energy of each state separately, whereas multi-configuration quasi-degenerate perturbation theory (MC-QDPT)[207] and extended MC-QDPT (XMC-QDPT)[209] are multi-state methods because the final energies are eigenvalues of the same matrix (hence they interact through the off-diagonal elements of that matrix). Similarly, complete active space perturbation theory (CASPT2)[15, 16, 17] is a state-specific method, and multi-state CASPT2 (MS-CASPT2)[75] and extended MS-CASPT2 (XMS-CASPT2)[209] are multi-state methods.

The original MC-PDFT is a state-specific method. We recently proposed state-interaction PDFT (SI-PDFT) as a multi-state generalization;[275] SI-PDFT yields the correct topography of adiabatic PESs for conical intersections and it has been applied successfully to several problems;[275, 276, 277] but it is inconvenient because two MCSCF calculations and two sets of orbitals are required, and it puts the ground state on an unequal footing with the excited states, which is sometimes undesirable (for example, for treating magnetic states). In this chapter we present a new multi-state method that eliminate these drawbacks of SI-PDFT.

This is called extended multi-state-PDFT (XMS-PDFT) because it uses the intermediate basis proposed by Granovsky[209] for XMC-QDPT.

A key aspect of all the above-mentioned MS methods is that they determine a model space spanned by the states to be treated as strongly interacting. Similar to XMC-QDPT or XMS-CASPT2, XMS-PDFT builds up a model space that spans the N lowest-energy states optimized in a state-averaged CASSCF (SA-CASSCF) calculation. (Generalizations to incomplete active spaces and smaller model spaces are straightforward but are not considered here.) The model space states are called the intermediate basis and are obtained by unitary transformation from the SA-CASSCF states.

6.2 Theory

6.2.1 MC-PDFT

The MC-PDFT method may be based on single-state CASSCF (SS-CASSCF) calculations or on SA-CASSCF calculations. In the present article we consider the latter type of calculation, in which case one starts with a reference wave function obtained by performing an SA-CASSCF calculation and given by

$$|\Psi_I\rangle = \sum_i c_i^I \, |\mathrm{CSF}_i\rangle \tag{6.1}$$

where i is the index of a configuration state function (CSF), and I is the index of a reference state. The MC-PDFT energy for state I is

$$E_{\mathrm{MC-PDFT}} = T_\mathrm{e} + V_\mathrm{e} + E_\mathrm{ot}(\rho_I, \Pi_I) \tag{6.2}$$

where the terms are the electronic kinetic energy, the classical electrostatic energy (which is the sum of the nuclear–nuclear repulsion, the electron–nuclear attraction energy, and the classical electron–electron repulsion), and the on-top energy computed as a functional of

the density ρ_I and the on-top density Π_I, both computed from $|\Psi_I\rangle$, with the latter given by

$$\Pi_I(\mathbf{r}) = \int \dots \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)\, d\sigma_1\, d\sigma_2\, d\mathbf{x}_3 \dots\, d\mathbf{x}_n \Big|_{\mathbf{r}_1=\mathbf{r}_2=\mathbf{r}} \tag{6.3}$$

where n is the number of electrons in the system, and

$$\mathbf{x}_k = \mathbf{r}_k\sigma_k \tag{6.4}$$

is the collectively the spatial coordinate r and the spin coordinate σ of an electron

Eq. 6.2 applies to MC-PDFT calculations starting with either SS-CASSCF or SA-CASSCF. We note that it does not separate the energy into an uncorrelated component, a static correlation component, and a dynamic correlation component. Because the original MC-PDFT method computes the state energies independently, it is a state-specific method in the sense that the final energy of each state is computed separately, even if one starts with SA-CASSCF kinetic energies, densities, and on-top densities.

6.2.2 multi-state MC-PDFT

To obtain the correct topography of PESs at conical intersections, we have proposed the SI-PDFT method1[275] as an MS extension of MC-PDFT. In SI-PDFT, we generate a set of intermediate states with the reference SA-CASSCF states and an auxiliary state from a state-specific ground-state CASSCF calculation. The ground intermediate state is obtained by projecting the SS-CASSCF state into the space spanned by the SA-CASSCF states, and the other intermediate states are obtained by performing Schmidt orthogonalization of the excited states obtained by the SA-CASSCF calculation to the ground intermediate state. Then one constructs an effective Hamiltonian in the intermediate state basis and diagonalizes it to get the SI-PDFT energy for each state. This treats the ground and excited states unequally. Moreover, using different orbital sets (i.e., using both the orbitals from the SS-CASSCF calculation and those from the SA-CASSCF calculation) is inconvenient. To

avoid these problems, we next propose two new multi-state MC-PDFT methods that use only one set of orbitals.

In general, the intermediate states are obtained by a unitary transformation:

$$|\Phi_I\rangle = \sum_J U_{JI} |\Psi_J\rangle \tag{6.5}$$

where $|\Phi_I\rangle$ is an intermediate state, and $|\Psi_J\rangle$ is an SA-CASSCF state. The electronic Hamiltonian of the molecule is diagonal in the SA-CASSCF states but usually not in the intermediate basis.

We construct an effective Hamiltonian in the intermediate-state basis with diagonal elements defined as

$$H_{II}^{\text{eff}} = E_I^{\text{MC-PDFT}} \tag{6.6}$$

where EMC-PDFTI is the MC-PDFT energy for the intermediate state $|\Phi_I\rangle$. The off-diagonal elements of the effective Hamiltonian are defined as

$$H_{IJ}^{\text{eff}} = \langle \Phi_I | \, \text{H} \, | \Phi_J \rangle \tag{6.7}$$

with I, J = 1, 2, ..., N, where N is the number of states in the model space. (In the present work, the number of states in the model space is always the same as the number of states averaged in the SA-CASSCF calculation.) The effective Hamiltonian is then diagonalized to give the multi-state MC-PDFT energies for each adiabatic state.

Following the above scheme, we next introduce how to generate the matrix **U** in XMS-PDFT.

6.2.3 XMS-PDFT

The intermediate basis in XMS-PDFT diagonalizes the effective Hamiltonian suggested by Granovsky for XMC-QDPT in ref. [209], where he stressed that “the effective Hamiltonian should be a function of the subspace spanned by the selected CI vectors, rather than a

function of any particular choice of basis in this subspace" and that "the computed energies must be uniquely defined, continuous and smooth functions of the molecular geometry and any other external parameters, with possible exceptions at the manifolds of their accidental degeneracy such as conical intersections". The XMS-CASPT2 method also uses this intermediate basis. We use the XMS-CASPT2 procedure[278] to explain this, and the explanation starts by recalling the procedure for MS-CASPT2.

In MS-CASPT2, the unperturbed Hamiltonian is defined as

$$\mathbf{H}_0 = \mathbf{PFP} + \mathbf{QFQ} \tag{6.8}$$

where

$$P = \sum_I |\Psi_I\rangle\langle\Psi_I| \tag{6.9}$$

is the projection operator onto the SA-CASSCF state space and

$$\mathbf{Q} = \mathbf{1} - \mathbf{P} \tag{6.10}$$

is the projection operator onto the complementary state space. In MS-CASPT2, the state-space Fock operator is defined as

$$F = \sum_{pq} f_{pq} E_{pq} = \sum_{pq} f_{pq} \sum_{\sigma} a^+_{p\sigma} a_{q\sigma} \tag{6.11}$$

where $E_{pq} = \sum_\sigma a^+_{p\sigma} a_{q\sigma}$ is a single-excitation operator, $a^+_{p\sigma}$ and $a_{q\sigma}$ are creation and annihilation operators on molecular orbitals p and q for spin σ, and f_{pq} is an element in the orbital Fock matrix

$$f_{pq} = h_{pq} + \sum rsd_{rs}(J^{sr}_{pq} - \frac{1}{2}K^{sr}_{pq}) \tag{6.12}$$

where h_{pq} contains the electronic kinetic energy and electron-Coulomb interaction, d_{rs} is a state-averaged density matrix element, and J^{sr}_{pq} and K^{sr}_{pq} are two-electron integrals. The matrix elements of the state-space Fock matrix are defined as

$$F_{IJ} = \langle\Psi_I|\,\mathrm{F}\,|\Psi_J\rangle = \sum_{pq}\sum_{ij} f_{pq} c^J_i c^J_j \langle\mathrm{CSF}_i|\,\mathrm{E}_{pq}\,|\mathrm{CSF}_j\rangle \tag{6.13}$$

The state-space Fock matrix defined in eq. 6.13 is not necessarily diagonal, because the reference wave functions (i.e., the SA-CASSCF wave functions) are the eigenstates of the Hamiltonian operator, not necessarily the eigenstates of the state-space Fock operator or the zeroth-order Hamiltonian.

The MS-CAPST2 method neglects the off-diagonal elements of the state-space Fock matrix, but following the prescription used in the XMS-CASPT2 method, the XMS-PDFT method diagonalizes the state-space Fock matrix by a transformation matrix $\mathbf{U}^{\mathrm{X}}$:

$$(\mathbf{U}^{\mathrm{X}})^{\dagger}\mathbf{F}(\mathbf{U}^{\mathrm{X}}) = \tilde{\mathbf{F}} \tag{6.14}$$

The $\mathbf{U}^{\mathrm{X}}$ matrix determined this way then yields the intermediate states defined by

$$|\Phi_I\rangle = \sum_J U^{\mathrm{X}}_{JI} |\Psi_J\rangle \tag{6.15}$$

where $|\Phi_I\rangle$ is an intermediate state in XMS-PDFT (and also in XMS-CASPT2). With the same transformation, we get a Hamiltonian matrix in the intermediate basis,

$$(\mathbf{U}^{\mathrm{X}})^{\dagger}\mathbf{H}(\mathbf{U}^{\mathrm{X}}) = \tilde{\mathbf{H}} \tag{6.16}$$

where $\mathbf{H}$ is the Hamiltonian matrix in the basis of the SA-CASSCF reference states, and $\tilde{\mathbf{H}}$ is the Hamiltonian matrix in the basis of the intermediate states.

After the intermediate states are obtained, XMS-PDFT defines an effective Hamiltonian in the intermediate basis such that the diagonal element H^{eff}_{II} is the MC-PDFT energy of intermediate state $|\Phi_I\rangle$, and the off-diagonal element H^{eff}_{IJ} is $\tilde{H}_{IJ}$. The XMS-PDFT energies ($\mathrm{E}^{\mathrm{XMS-PDFT}}_I$) and eigenvectors are obtained by diagonalizing the effective Hamiltonian matrix.

We notice that the off-diagonal elements in the state-space Fock matrix are zero for states with different symmetries. This suggests that the XMS-PDFT method is identical to MC-PDFT if all states in the model space belong to different irreps (similarly, XMS-CASPT2

is identical to MS-CASPT2 or even single-state CASPT2 for such a case). This is not a problem, but we have found that the off-diagonal elements in the state-space Fock matrix are almost zero for many geometries in some mixed-valence[279] systems (see Fig. S3,† where sections and figures with the prefix S are in the ESI†) even when the states have the same symmetry, and we will see that XMS-PDFT does not always give good results for such systems.

6.3 Computational Details

The calculations are performed in OpenMolcas v18.09, tag 548-g19e2926-dirty,[231] with codes modified to perform XMS-PDFT and FMS-PDFT calculations.

In XMS-CAPST2 calculations, an ionization-potential-electron-affinity (IPEA) shift[148] of 0.25 a.u. is used. In the PDFT calculations, we used the translated PBE (tPBE) on-top functional.

Table 6.1 presents the wave function symmetry, basis set, number of averaged states, number of active electrons, and identities of active MOs for each system studied. The internal coordinates that are scanned for each system are shown in Table 6.2. The geometries are available in Section S1.

Table 6.1: Systems Studied, Symmetry Enforced on the Wave Function (Sym), Basis Set, Number of States in the SA Calculation (N_{states}), Number of Active Electrons (n) and Active Molecular Orbitals (Active MOs)

System	Sym	Basis Set	N_{states}	n	Active MOs
LiF	C_1	jun-cc-pVQZ[44, 230]	2	8	2s and $2p_z$ of F, 2s of Li
LiH	C_{2v}	aug-cc-pVQZ[255, 44]	4	2	2s, $2p_z$, 3s, $3p_z$ of Li, 1s of H
					N, and O atoms and 1s of H atom)
C_6H_5OH	C_1	jul-cc-pVDZ[255, 44, 230]	2	12	3π, $3\pi^*$, σ^*_{OH},σ_{CO}, σ^*_{CO}, and p_z of O

Table 6.2: Systems Studied and the Internal Coordinates Scanned for Potential Energy Curves

System	Internal coordinates scanned
LiF	r(LiF) = [1.0–9.0] $\mathring{\mathrm{A}}$
LiH	r(LiH) = [1.0–12.0] $\mathring{\mathrm{A}}$
C_6H_5OH	r(OH) = [0.5–3.0] $\mathring{\mathrm{A}}$ and τ(C–C–O–H) = 1 or 10°

6.4 Results and Discussion

6.4.1 Lithium fluoride (LiF)

Lithium fluoride has an avoided crossing of the ground state and first excited state that has been widely studied.[280, 281, 282, 283, 284, 285, 286, 287] The ground state at the equilibrium distance is ionic, corresponding to the $(2p_{z,F})^2(2s_{Li})^0$ configuration. The ground state has A_1 symmetry in the C_{2v} group. This state interacts with another A1 state that corresponds to two neutral ground-state atoms, namely $(2p_{z,F})^1(2s_{Li})^1$. The accurate value of the distance of the avoided crossing is about 7.4 $\mathring{\mathrm{A}}$.[280] However, theoretical calculations usually underestimate the distance by 1.0 $\mathring{\mathrm{A}}$, with an exception being the calculation in ref. [282].

The MC-PDFT method gives an unphysical double crossing between 4 $\mathring{\mathrm{A}}$ and 6 $\mathring{\mathrm{A}}$, associated with a “dip” of the energy curve, as shown in Fig. 6.1(a). The XMS-PDFT and FMS-PDFT methods, however, remove the incorrect double crossing and also recover the expected shape of the avoided crossing at a larger distance. Additionally, the two new multi-state PDFT methods preserve the correct asymptotic character of the two states, and they work well for the whole potential energy curve.

Fig. 6.1(b) shows that the XMS-PDFT results agree with XMS-CASPT2 for the overall

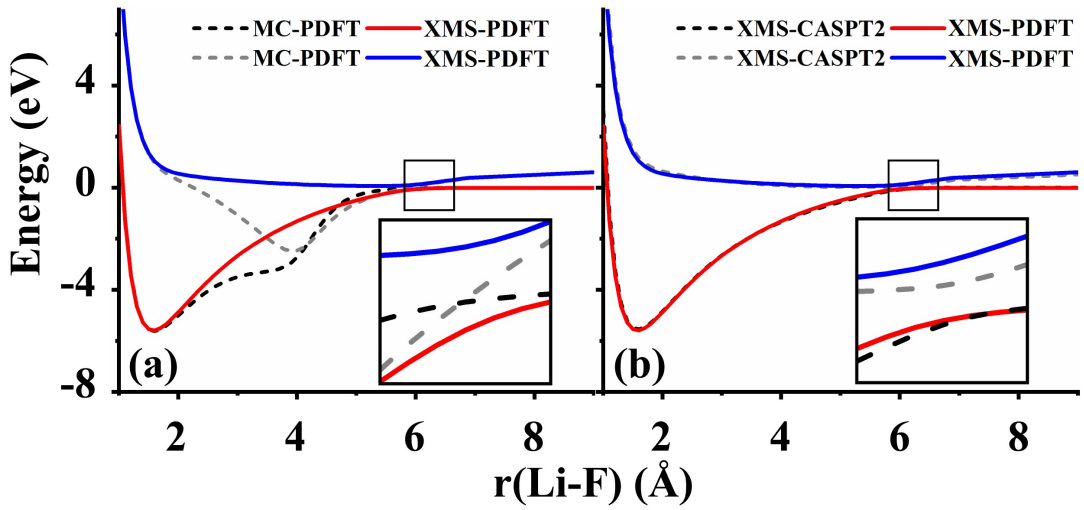

Figure 6.1: Comparison of the potential energy curves of the two lowest A_1 electronic states of LiF with (a) XMS-PDFT and MC-PDFT, (b) XMS-PDFT and XMS-CASPT2. The area near the avoided crossing (indicated by a small box) for each curve is also shown magnified. J. J. Bao, C. Zhou, Z. Varga, S. Kanchanakungwankul, L. Gagliardi, and D. G. Truhlar.Multi-state pair-density functional theory.Faraday Discuss., 224:348–372, 2020. Reproduced by permission of The Royal Society of Chemistry.

shapes of the two curves. The minimum separation of the two curves is 0.18 eV at 5.97 $\mathring{\mathrm{A}}$ by XMS-PDFT, and 0.11 eV at 6.11 $\mathring{\mathrm{A}}$ by XMS-CASPT2. The bond lengths with the minimum energy separation obtained by the two methods are significantly shorter than 7.4 $\mathring{\mathrm{A}}$ because the calculations underestimate the electron affinity of F, which is a very hard[288] problem.

6.4.2 Lithium hydride (LiH)

The ground state of lithium hydride is an ionic state near the equilibrium geometry, but this state interacts with three covalent states, corresponding to $(2s_{Li})^1(1s_H)^1$, $(2p_{z,Li})^1(1s_H)^1$, and $(3s_{Li})^1(1s_H)^1$ configurations, as the Li–H bond dissociates. All four states have A_1 symmetry in the C_{2v} point group.

Despite the complexity of the ionic state of LiH crossing with at least three other states

as shown in Fig. 6.2, a similar pattern to LiF is still found for the third and fourth states of LiH beyond 10 Å. The zoomed-in region in Fig. 6.2(a) shows that the MC-PDFT curves for the third and fourth states still have a dip and a double crossing, while XMS-PDFT recovers the avoided crossing of the two states and also remove the dip.

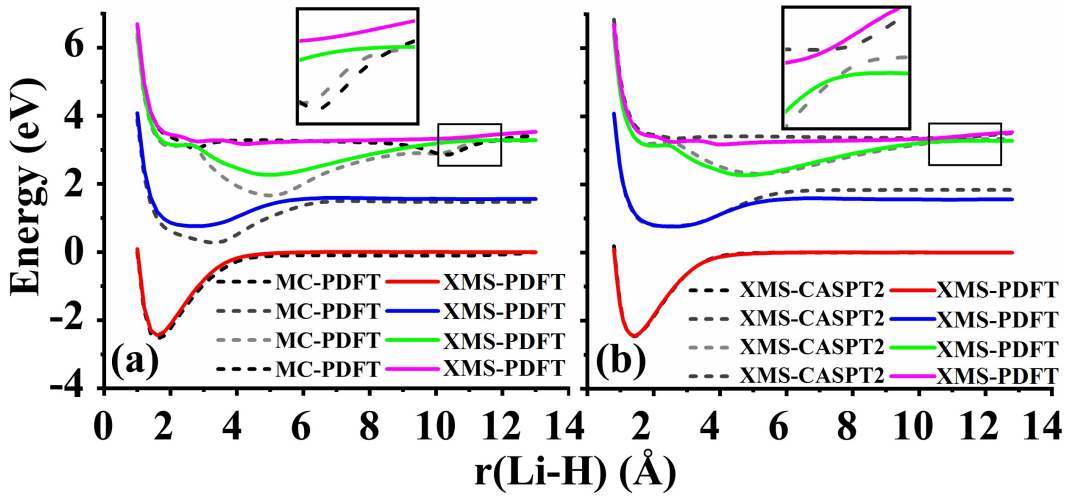

Figure 6.2: Comparison of the potential energy curves of LiH with (a) XMS-PDFT and MC-PDFT, (b) XMS-PDFT and XMS-CASPT2. The zoomed-in area near the avoided crossing is shown for the ionic state and the highest covalent state calculated. J. J. Bao, C. Zhou, Z. Varga, S. Kanchanakungwankul, L. Gagliardi, and D. G. Truhlar.Multi-state pair-density functional theory.Faraday Discuss., 224:348–372, 2020. Reproduced by permission of The Royal Society of Chemistry.

The first (red) and second (blue) states calculated by XMS-PDFT agree very well with those calculated by XMS-CASPT2. The minimum energy separation between the third and fourth states is 0.10 eV at 10.66 Å by XMS-PDFT, and 0.07 eV at 11.28 Å by XMS-CASPT2. The shapes of the XMS-CASPT2 curves match much better with XMS-PDFT than with MC-PDFT, especially for the energy minima of the excited states.

6.4.3 Phenol (C_6H_5OH)

The O–H bond dissociation in phenol has been well studied in the past and it can be used as a model system for testing whether a method gives a proper description of potential energy curves for photodissociation. We tested MC-PDFT and XMS-PDFT for the O–H dissociation in phenol with an H–O–C–C dihedral angle of 1° (nearly planar) or 10°.

Fig. 6.3 shows that the MC-PDFT potential energy curves are qualitatively wrong at both angles, with a double crossing when the dihedral angle is 1° and a lack of avoidance at 10°. The XMS-PDFT method successfully produces avoided crossings near 2.2 Å for both torsion angles with minimum energy separations of 0.04 and 0.28 eV for 1° and 10°, respectively. The corresponding O–H distances are 2.21 and 2.15 Å.

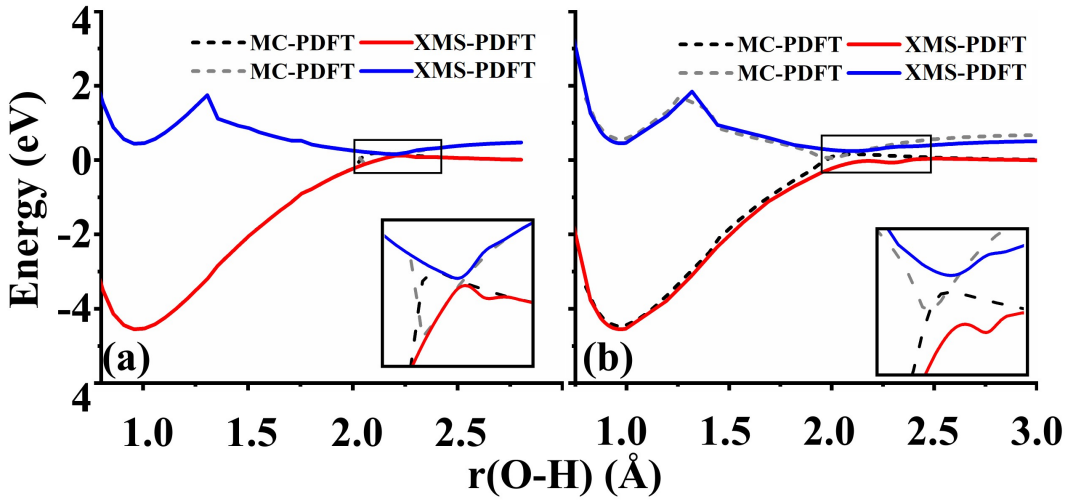

Figure 6.3: (a) and (b) Potential energy curves of two states for O–H dissociation in phenol with H–O–C–C dihedral angles of 1° and 10°, calculated by MC-PDFT (dashed and dotted lines) and XMS-PDFT (solid lines). J. J. Bao, C. Zhou, Z. Varga, S. Kanchanakungwankul, L. Gagliardi, and D. G. Truhlar.Multi-state pair-density functional theory.Faraday Discuss., 224:348–372, 2020. Reproduced by permission of The Royal Society of Chemistry.

In regions that are far away from the avoided crossings for each dihedral angle, the XMS-PDFT curves agree very well with the MC-PDFT ones. However, we noticed that for

this molecule XMS-PDFT presents a noticeable “bump” after the avoided crossing. This is apparently because the geometry dependence of the off-diagonal elements of the effective Hamiltonian matrix is not consistent enough with the geometry dependence of the diagonal elements. However, the bumps are no greater than 0.07 eV, corresponding to 1.6 kcal/mol, which is usually accurate enough for treating electronically excited states.

6.5 Conclusion

A general scheme for multi-state MC-PDFT is proposed in this chapter. In this scheme, the CASSCF reference states are rotated to a set of intermediate states via a unitary transformation, and an effective Hamiltonian matrix in the intermediate-state basis is constructed using the MC-PDFT method for the diagonal elements and wave function theory for the off-diagonal ones. As a practical method, XMS-PDFT is proposed for the unitary transformation in this chapter, and it is tested on three systems exhibiting avoided crossings of two and four states. The XMS-PDFT method uses the transformation proposed by Granovsky for XMC-QDPT. Tests are performed on systems with avoided crossings to compare XMS-PDFT to state-specific MC-PDFT and the more expensive multi-state method, XMS-CASPT2. We found that XMS-PDFT gives good results for all systems except the mixed-valence spiro cation. The new multi-state method proposed here is preferred to the previous SI-PDFT because they treat the ground state and excited states on an equal footing and they require only a single SA-CASSCF calculation and a single set of orbitals.

Chapter 7

Compressed-State Multi-State Pair-Density Functional Theory

7.1 Introduction

Multiconfiguration pair-density functional theory[19, 20] (MC-PDFT) provides an efficient method to compute potential energy surfaces by starting with multiconfiguration self-consistent-field (MCSCF) calculations to generate orbitals and using multiconfigurational spin eigenfunctions to allow bond dissociation and to represent open-shell states without spin mixing. One may use various kinds of MCSCF calculations, and in the present paper the examples will be based on state-averaged complete-active-space SCF[289, 290] (SA-CASSCF).

Single-state MC-PDFT involves a final energy that is the sum of the classical energy of a configuration interaction (CI) eigenvector (usually from the last step of an MCSCF calculation) plus an on-top energy contribution computed from a pair-density functional,

where the classical CI energy is defined as the sum of the kinetic energy of the CI state plus its classical electrostatic energy. Single-state MC-PDFT, just like single-state complete active space second-order perturbation theory[15, 16] (CASPT2), single-state multireference Moller–Plesset theory[274] (MR-MP2), or single-state n-electron valence state perturbation theory[204, 205] (NEVPT2), which can also build on MCSCF starting points, does not yield the correct topology of potential energy surfaces at conical intersections, which are very important for both spectroscopy[270] and photochemistry.[291] To solve this problem in CASPT2, MR-MP2, or NEVPT2 and to obtain a more consistent treatment of nearly degenerate states, one turns to quasi-degenerate perturbation theory in which the final state is diagonalization of a model-space effective Hamiltonian; this yields multi-state CASPT2[75] (MS-CASPT2), extended MS-CASPT2[209, 278] (XMS-CASPT2), multiconfiguration quasi-degenerate perturbation theory[207] (MC-QDPT), extended MC-QDPT[209] (XMC-QDPT), and quasi-degenerate NEVPT2[208] (QD-NEVPT2). To solve this problem in MC-PDFT, we previously proposed three methods of increasing convenience: state-interaction PDFT[275] (SI-PDFT), variational multi-state PDFT[268] [268], and extended multi-state PDFT[268] (XMS-PDFT). In the present paper, we propose a new method that improves on all of these (as explained next) and is recommended as the method of choice for future applications to calculating excited electronic states by MC-PDFT.

The SI-PDFT method requires two calculations in different molecular orbital bases, one from a state-specific calculation and the other from a state-averaged (SA) calculation. The VMS-PDFT and the XMS-PDFT methods belong to a class of methods that we call multi-state PDFT (MS-PDFT). These methods are more convenient in that they involve only a single orbital basis, and they work by defining a rotation matrix that transforms the eigenstates of an SA-CASSCF calculation into a set of intermediate states whose energy is calculated by MC-PDFT; the difference of these two methods is how to obtain the rotation matrix. The VMS-PDFT intermediate states maximize the sum of the MC-PDFT energies for

these intermediate states. This requires computing the MC-PDFT energies at each step of the maximization process that yields the transformation matrix. The task was made more manageable by using of Fourier analysis and carrying out the variational maximization within a low-order Fourier approximation; this version of VMS-PDFT is called FMS-PDFT. The XMS-PDFT intermediate states diagonalize the generalized Fock operator[209, 37, 292] in the space of the SA-CASSCF states. This does not require additional MC-PDFT energy calculations to obtain the transformation matrix as in VMS-PDFT, and so it is much more efficient. As we have shown previously,[275, 268] the SI-PDFT and VMS-PDFT methods perform well in all cases that we have tested, but the XMS-PDFT works well in only seven out of eight cases tested.[268]

In this paper, we propose a new MS-PDFT method, namely, compressed-state multi-state PDFT (CMS-PDFT), that is as convenient (even slightly more convenient) as XMS-PDFT, but works as well as SI-PDFT and VMS-PDFT. It requires only one orbital set, and the transformation to intermediate states is based entirely on data available from the SA-CASSCF calculation without requiring additional MC-PDFT calculations. The essential characteristic of the new method is that the rotation to intermediate states maximizes the sum over states of the classical Coulomb energies of the active electrons. This is equivalent to maximizing the sum over states of the wave function contributions to the MC-PDFT calculations, whereas the VMS-PDFT method maximizes the sum over states of the total MC-PDFT energies. The primary goal for CMS-PDFT is to find a set of intermediate states for which it is appropriate to calculate the MC-PDFT energy by adding an on-top energy to a classical contribution. We will see that a successful strategy is to choose intermediate states that maximize the sum over states of the classical electron–electron Coulomb energies; this leads to a more compressed electronic distribution in the CMS intermediate states than for the reference CI states, and this is why the method is called compressed-state MS-PDFT. We will see that maximizing the sum over states of the classical electron–electron Coulomb

energies is equivalent to maximizing the classical CI energy (defined above), whereas the previously presented VMS-PDFT method maximizes the sum of the classical CI energy and the on-top energy, which is less computationally convenient. Nevertheless both methods may work well in part because they find an intermediate states whose energy—prior to the configuration mixing of the diagonalization of the model space effective Hamiltonian—is least already lowered by configuration mixing; this minimizes the possibility of double counting of configuration mixing effects.

7.2 Theory

7.2.1 MC-PDFT

The MC-PDFT energy for a reference state K with wave function $|\Psi_K\rangle$, which is a CI eigenstate, is defined as

$$E_I^{\mathrm{MC-PDFT}} = V_{\mathrm{NN}} + \sum_{pq} D_{pq}^{KK} h_{pq} + \frac{1}{2} \sum_{pqrs} D_{pq}^{KK} D_{rs}^{KK} g_{pqrs} + E_{\mathrm{ot}}(\rho_K, \Pi_K) \tag{7.1}$$

where V_{NN} is the nuclear repulsion, p, q, r, and s are generic-orbital indices, D_{pq}^{KK} is the spinless reduced 1-particle density matrix for state K, h_{pq} is the one-electron energy that includes the electronic kinetic energy and the electron–nuclear attraction, g_{pqrs} is a two-electron Coulomb integral given in Mulliken notation by $(pq|rs)$, and $\mathrm{E_{ot}}$ is the on-top energy functional of the density, ρ_K, and the on-top density, Π_K. The sum of all of the terms except the on-top energy equals the sum of the kinetic energy plus the classical electrostatic energy of the CI state, and (as defined in the Introduction) is the classical CI energy; the classical CI energy is the wave function contribution to the MC-PDFT energy.

The trace of the two-electron Coulomb energy is written as

$$Q = \frac{1}{2} \sum_{K=1}^{N} \sum_{pqrs} D_{pq}^{KK} D_{rs}^{KK} g_{pqrs} \tag{7.2}$$

where N is the number of states in the model space. In the notation of eq 7.2, the indices p and q are for one electron and r and s are for the other electron; therefore, we can rewrite Q as a sum of three terms, one with both electrons in inactive (i) orbitals, one with one electron in an inactive orbital and one in an active (a) orbital, and the last one with both electrons in active orbitals

$$Q = Q_{\mathrm{i-i}} + Q_{\mathrm{i-a}} + Q_{\mathrm{a-a}} \tag{7.3}$$

$$Q_{\mathrm{i-i}} = 2N \sum_{ij} g_{iijj} \tag{7.4}$$

$$Q_{\mathrm{i-a}} = 2 \sum_{K=1}^{N} \sum_{itu} D_{tu}^{KK} g_{tuii} \tag{7.5}$$

$$Q_{\mathrm{a-a}} = \frac{1}{2} \sum_{K=1}^{N} \sum_{tuvx} D_{tu}^{KK} D_{vx}^{KK} g_{tuvx} \tag{7.6}$$

where t, u, v, and x are indices of active orbitals, and i and j are indices of inactive orbitals (which are orbitals that are doubly occupied in all configurations in all states).

7.2.2 MS-PDFT

In MS-PDFT methods,[268] we first obtain a set of rotated states, $|\Phi_K\rangle$, such that

$$|\Phi_K\rangle = \sum_J U_{KI} |\Psi_J\rangle \tag{7.7}$$

where K, L, ... are intermediate-state indices and U_{KI} is a rotation matrix element, and **U** is an orthogonal matrix. (We do not need a complex unitary transformation because we are neglecting spin–orbit coupling and have no external magnetic field, so we take all wave functions to be real). Then, we construct an effective Hamiltonian, whose diagonal elements are calculated as

$$H_{KK}^{\mathrm{eff}} = E_K^{\mathrm{MC-PDFT}} \tag{7.8}$$

where EKMC-PDFT is the MC-PDFT energy for the intermediate state $|\Phi_K\rangle$. The off-diagonal elements in the effective Hamiltonian are defined as

$$H_{KL}^{\text{eff}} = \langle \Phi_K | \, \mathrm{H} \, | \Phi_L \rangle \tag{7.9}$$

We then diagonalize the effective Hamiltonian matrix to obtain the MS-PDFT energies. Because the last step is a diagonalization, MS-PDFT methods give the correct topology of potential energy curves at and near conical intersections.

7.2.3 CMS-PDFT

In CMS-PDFT, we chose the intermediate states to maximize the trace of the wave function contribution to the MC-PDFT energies. It is easily shown that this is equivalent to maximizing Q, which is the classical electrostatic electron–electron energy (Coulomb energy), and that it is also equivalent to maximizing $Q_{\mathrm{a^\vee a}}$; the latter equivalence is because the traces of the other terms are invariant to an orthogonal transformation. Therefore, we simply maximize the active–active Coulomb energy $Q_{\mathrm{a^\vee a}}$. Maximizing $Q_{\mathrm{a^\vee a}}$ means that the electronic densities for the intermediate states are compressed, which we indicate in the name of the new method.

Maximizing the classical Coulomb energy was used in orbital localizations decades ago by Edmiston and Ruedenberg.[293] However, the Edmiston–Ruedenberg orbital localization by maximizing the classical Coulomb energy does not change the wave function of a state; while in CMS-PDFT, we change the state basis on which we evaluate the PDFT energies. The compressed electronic densities give us better PDFT energies and thus better MS-PDFT potential energy surfaces because in MC-PDFT, the classical Coulomb energy and the on-top energy are lower when the electronic densities are more spread out. The issue of localization and delocalization also arises in KS-DFT. However, the compression achieved here is not motivated by decreasing delocalization error,[294] by which KS-DFT with many

approximate exchange–correlation functionals predicts too delocalized an electronic density because the classical Coulomb term and local functionals give lower energies with more delocalized densities. We have previously argued that MC-PDFT is free from delocalization error because MC-PDFT uses electronic densities from the CASSCF wave functions or other MCSCF wave functions, which are free from density functional delocalization error.[295]

CMS-PDFT for Two-State Problems

In MS-PDFT for a two-state problem (N = 2), the sum of active–active Coulomb energy for the intermediate states can be expressed exactly as

$$Q_{\mathrm{a-a}} = A + B + \sin(4\theta_{KJ}) + C\cos(4\theta_{KJ}) \tag{7.10}$$

where θ_{KJ} is the rotation angle that defines the transformation matrix according to

$$\mathbf{U}_{KJ}^{\mathrm{CMS}}(\theta_{KJ}) = \begin{pmatrix} \cos(4\theta_{KJ}) & \sin(4\theta_{KJ}) \\ -\sin(4\theta_{KJ}) & \cos(4\theta_{KJ}) \end{pmatrix} \tag{7.11}$$

To obtain the CMS-PDFT intermediate states, we maximize $Q_{\mathrm{a\breve{}a}}$ by using eq. 7.10 rather than a local quadratic function because eq. 7.10 allows us to guarantee convergence to a maximum, whereas using a quadratic algorithm can lead to either a maximum or a minimum. The optimal rotation angle is found iteratively. First, we calculate $Q_{\mathrm{a\breve{}a}}$ at the XMS rotation angle, $\theta_{KJ}^{\mathrm{XMS}}$ and at $\theta_{KJ}^{\mathrm{XMS}}$ – 5° and $\theta_{KJ}^{\mathrm{XMS}}$ + 5°. Then, these quantities are used to find A, B, and C in eq. 7.10. We then find the angle, KJCMS1 that maximizes eq. 7.10. Then, we calculate $Q_{\mathrm{a\breve{}a}}$ at $\theta_{KJ}^{\mathrm{CMS1}}$ and $\theta_{KJ}^{\mathrm{CMS1}} \pm 5°$ and repeat the maximization to find a new angle, KJCMS2. We repeat this process until the change in $Q_{\mathrm{a\breve{}a}}$ is less than $10^{\breve{}6}$ hartree. For the cases considered in the present paper, only two iterations were required, that is, $\theta_{KJ}^{\mathrm{CMS2}}$ is already converged.

CMS-PDFT for More Than Two States

With N states, there will be $\frac{N(N\check{}1)}{2}$ pairs of states and each pair is indexed as I_{pair}. In CMS-PDFT for a multiple-state problem ($N \geqslant 3$), we use successive two-by-two rotations to maximize $Q_{\mathrm{a\check{}a}}$. First, we define a rotation matrix, $R_{KL}(\theta_{KL})$, such that a pair of states, K and L, are rotated by an angle of θ_{KL} to obtain a new set of intermediate states, namely,

$$|\Xi_M\rangle = \sum_N (R_{KL})_{MN} \, |\Phi_N\rangle \tag{7.12}$$

$$(R_{KL})_{MN} = \begin{cases} \delta_{MN}, & M \neq K \text{ or } L \\ \delta_{MN}\cos(\theta_{KJ}), & M = K \text{ or } L \\ \sin(\theta_{KJ}), & M = K \text{ and } N = L \\ -\sin(\theta_{KJ}), & M = L \text{ and } N = K \end{cases} \tag{7.13}$$

The trace of the active–active Coulomb energy of the states in the set $|\Xi_M\rangle$ is denoted as $Q^{\Xi}_{\mathrm{a\check{}a}}(\theta_{KL})$ and that of the states in set $|\Xi_M\rangle$ is denoted as $Q_{\mathrm{a\check{}a}}\phi$. In cases where $Q^{\Phi}_{\mathrm{a\check{}a}}$ is already a maximum, the condition for $Q^{\Xi}_{\mathrm{a\check{}a}}(\theta_{KL})$ to be maximized is that θ_{KL} = 0, namely, $|\Xi_M\rangle$ is identical to $|\Xi_M\rangle$. In cases where $Q^{\Phi}_{\mathrm{a\check{}a}}$ is not already a maximum, we first express $Q^{\Xi}_{\mathrm{a\check{}a}}(\theta_{KL})$ for a rotation involving one pair of states (K and L) as an equation analogous to eq. 7.12

$$Q_{\mathrm{a-a}} = A + B + \sin(4\theta_{KJ}) + C\cos(4\theta_{KJ}) \tag{7.14}$$

and we find the maximum and its corresponding rotation angle in the same way as explained above for the two-state case. Then, we replace $|\Phi_M\rangle$ with $|\Xi_M\rangle$, and we repeat the previous process for another pair of states. A cycle is defined as maximizing $Q^{\Xi}_{\mathrm{a\check{}a}}$ for state pairs from I_{pair} = 1 to $I_{\mathrm{pair}} = \frac{N(N\check{}1)}{2}$ (forwardly) and from $I_{\mathrm{pair}} = \frac{N(N\check{}1)}{2}$ to I_{pair} = 1 (backwardly). This cycle is repeated until $Q^{\Xi}_{\mathrm{a\check{}a}}$ changes by less than $10^{\check{}6}$ hartree from the previous cycle.

Although we use the change in $Q^{\Xi}_{\mathrm{a}\breve{}\mathrm{a}}$ as the only convergence criterion to obtain the rotated states, we observed that θ_{KL} is always less than $10^{\breve{}6}$ deg for all pairs of states when convergence is achieved.

7.3 Computational Details

We studied a four-state case and all five two-state cases that were studied in our previous work;[268] the methods are tested and illustrated by studying the adiabatic potential energy curves along various paths through configuration space. The active spaces and the basis set used are listed in Table 1, and the reaction coordinates of the paths and the ranges over which they are scanned are in Table 2.

Table 7.1: Systems Studied, Symmetry Enforced on the Wave Function (Sym), Basis Set, Number of States in the SA Calculation (N_{states}), Number of Active Electrons (n) and Active Molecular Orbitals (Active MOs)

System	Sym	Basis Set	N_{states}	n	Active MOs
LiF	C_1	jun-cc-pVQZ[44, 230]	2	8	2s and $2p_z$ of F, 2s of Li
LiH	C_{2v}	aug-cc-pVQZ[255, 44]	4	2	2s, $2p_z$, 3s, $3p_z$ of Li, 1s of H
HNCO	C_1	cc-pVDZ[255, 44]	2	16	valence shell (2s and 2p of C, N, and O atoms and 1s of H atom)
CH_3NH_2	C_1	6-31++G(d,p)[296, 297]	2	6	2σ, $1\sigma^*$, $2p_z$, 3s, and $3p_z$ of N
C_6H_5OH	C_1	jul-cc-pVDZ[255, 44, 230]	2	12	3π, $3\pi^*$,σ_{OH}, σ^*_{OH},σ_{CO}, σ^*_{CO}, and p_z of O
spiro[a]	C_{2v}	6-31G(d)[298]	2	11	see ref. [276]

[a]Spiro denotes 2,2′,6,6′-tetrahydro-4H,4′H-5,5′-spirobi[cyclopenta[c]pyrrole] cation.

The CMS-PDFT calculations to generate energies are performed in OpenMolcas[231] version 18.09, tag 2501-g65d5cf6-dirty, which is a local version. The computational times given in Section 4.7 were collected with OpenMolcas version 19.11, tag 1975-g36ac7380, which is a public version, on the Intel Core i7-8650U CPU @ 2.11 GHz processor with 8.0

Table 7.2: Systems Studied and the Internal Coordinates Scanned for Potential Energy Curves

System	Internal coordinates scanned
LiF	r(LiF) = [1.0–9.0] $\mathring{\mathrm{A}}$
LiH	r(LiH) = [1.0–12.0] $\mathring{\mathrm{A}}$
HNCO	r(NC) = [1.25–3.00] $\mathring{\mathrm{A}}$ and τ(HNCO) = 150 or 175 $^\circ$
CH_3NH_2	r(NH) = [0.8–3.6] $\mathring{\mathrm{A}}$ and τ(H6–C4–N1–H3) = 0, 90, 95, or 100° [a]
C_6H_5OH	r(OH) = [0.5–3.0] $\mathring{\mathrm{A}}$ and τ(C–C–O–H) = 1 or 10°
spiro	see Section 7.4.6

[a] Each of the two H dissociation pathways on the amino group is studied when τ is 0.

GB memory (not all 8.0 GB memory was requested during the computation).

7.4 Results and Discussion

7.4.1 LiF

The avoided crossing of the ionic and covalent potential energy curves of LiF has been studied by many previous workers, and it is widely used to test new methods.[280, 281, 282, 283, 284, 285, 286, 287, 299]

7.1a shows that the CMS-PDFT potential curves are almost indistinguishable from the XMS-PDFT ones even in the avoided crossing region. This stability with respect to method is very encouraging because the CMS ones are based on a variational criterion and the XMS ones are based on a generalized Fock matrix.

7.1b compares the results to XMS-CASPT2. Our previous experience has been that the Li–F distance at the minimum gap (the gap is the energetic separation of the two states) is shorter when calculated by calculated by MS-PDFT than when calculated by XMS-CASPT2,[209] and the present calculations are consistent with that. However, the most

important point shown in 7.1b is that CMS-PDFT and XMS-CASPT2 agree very well. This is very encouraging because for practical problems with larger active spaces, CMS-PDFT is much less computationally expensive than XMS-CASPT2.

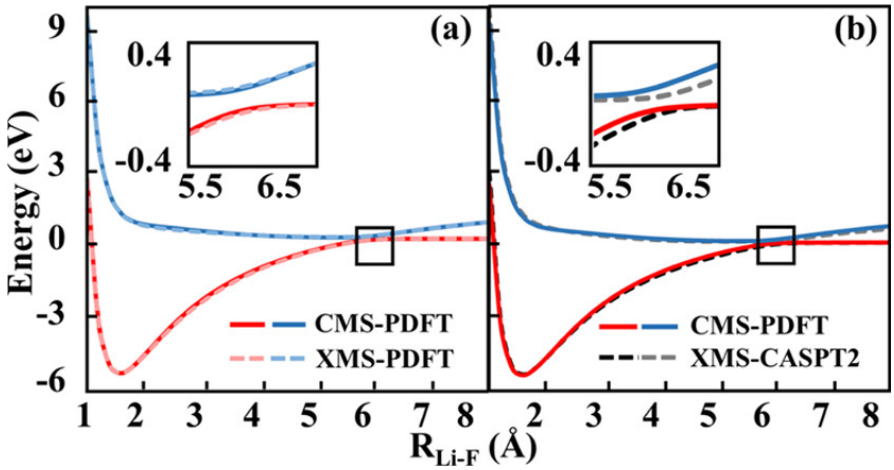

Figure 7.1: CMS-PDFT, XMS-PDFT, and XMS-CASPT2 potential energy curves for LiF.

7.4.2 LiH

LiH is another good test system that has been widely studied.[300, 301, 302, 303, 304] Here we test CMS-PDFT on the lowest four 1Σ states of LiH, and the results are in 7.2. 7.2a shows that the CMS results are smoother than those obtained with FMS in the region between 3 and 5 $\mathring{\text{A}}$; this may be a consequence of the simpler nature of the CMS maximization, which allows us to carry the iterative maximization to completion (as indicated above in Section 2.3.2), whereas the FMS calculations employ a one-shot maximization because of the cost in performing the maximization process.

7.2b shows that the CMS-PDFT curves overlap with the XMS-CASPT2 ones very well for most regions. The region between 11 and 13 $\mathring{\text{A}}$ contains the minimum gap for both CMS-PDFT and XMS-CASPT2, and it is shown in the zoomed insert of 7.2b; there is a slight quantitative difference in the results, but both methods show a smooth avoided crossing.

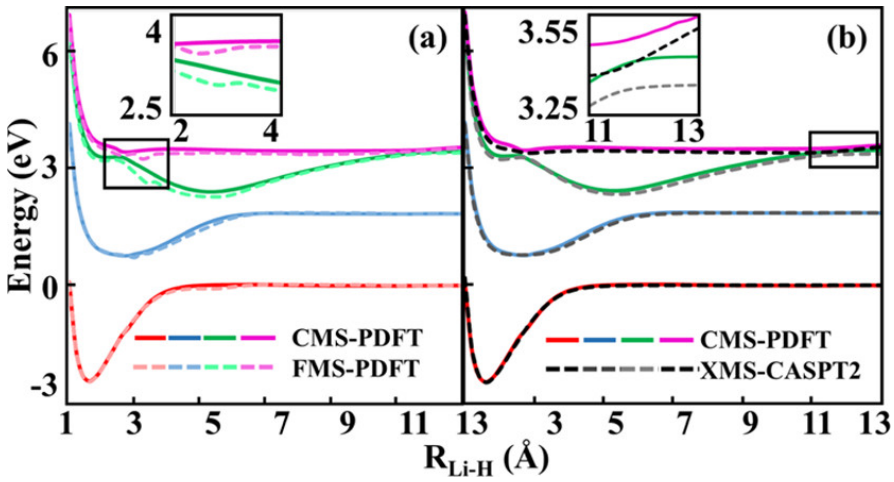

Figure 7.2: CMS-PDFT, XMS-PDFT, and XMS-CASPT2 potential energy curves for LiF.

7.4.3 HNCO

The CMS-PDFT potential curves for isocyanic acid (HNCO → HN···CO) are in 7.3, and they show that CMS-PDFT works very well for this molecule. The separation of the two states obtained by CMS-PDFT is slightly greater for CMS-PDFT than for XMS-PDFT. The difference arises from the fact that the XMS intermediate states and the CMS-PDFT intermediate states are very different in the zoomed-in regions, as can be seen from Figure S1; (figures with a prefix S are in the Supporting Information) the stability of the methods to this kind of change in the intermediate states is encouraging.

7.4.4 CH_3NH_2

As stated in ref [305], “Study of the excited electronic photochemistry of simple alkyl amines has a number of motivations.... A number of important systems, such as energetic materials..., contain the alkyl amino units and their properties and behavior depend on alkyl amino excited states and dynamics.” Methylamine, shown in 7.4, is the simplest amine, and it has been the subject of many experimental and theoretical studies.[305, 306, 307, 308, 309,

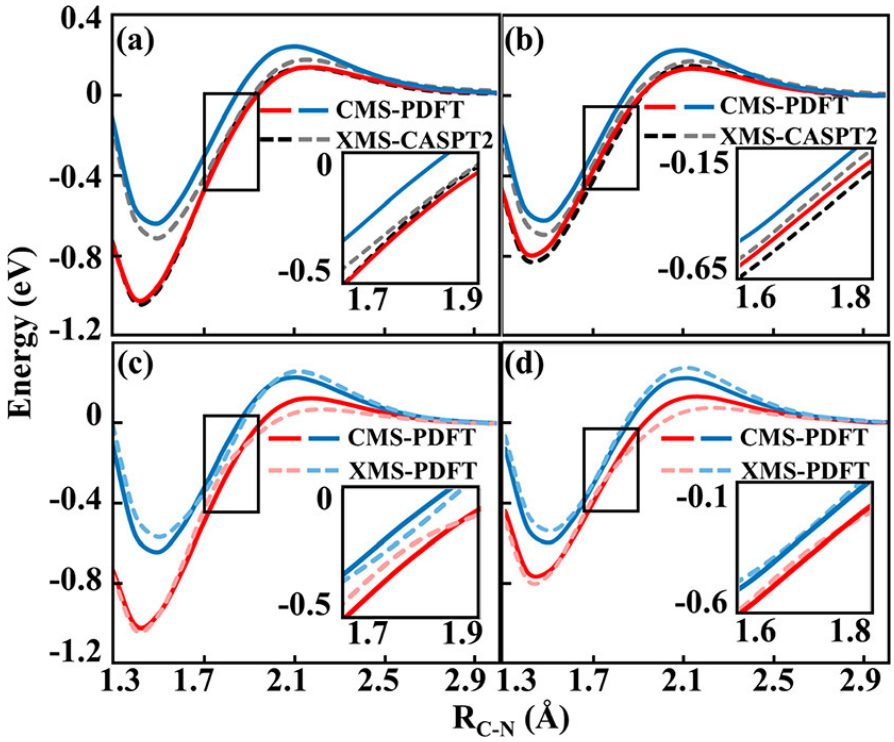

Figure 7.3: CMS-PDFT and XMS-PDFT potential energy curves for isocyanic acid (HNCO) with two dihedral angles as in Table 7.2; the dihedral angle is 175° in (a,c) and 150° in (b,d).

310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320] Michael and Noyes[306] reported that the N–H bond dissociation is a dominant reaction channel with a yield of more than 75% from the first-excited methylamine dissociation. Butler and co-workers[308] measured translational energies of different products from several reaction channels of the CH_3NH_2 dissociation at 222 nm and concluded that the major channel is N–H bond dissociation.

We studied three pathways for dissociation of the N–H3 bond, where the labeling is explained by 7.4. The first pathway has the dihedral angle ϕ at 90°, and the potential energy curves are shown in 7.5a. 7.5b is a similar path, which has ϕ at 100°, which passes farther from the conical intersection and so has a larger gap. We see excellent results, encouragingly very similar for both methods, although they have different intermediate states. 7.5c,d are both for the pathway with ϕ equal to 0 deg. 7.5c shows that the CMS-PDFT curves are

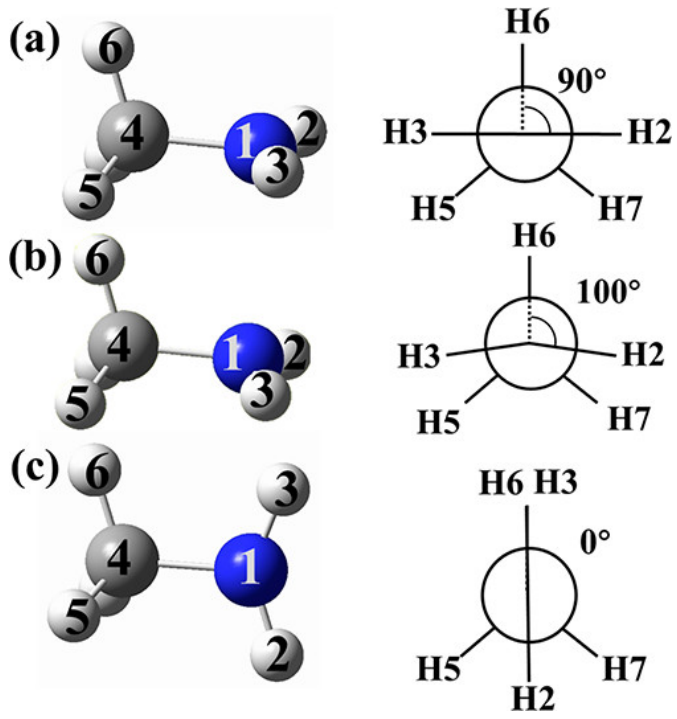

Figure 7.4: Lowest energy structures along each three pathways studied by CMS-PDFT. The figure illustrates the H2–N–C–H6 dihedral angle ϕ that is fixed along the dissociation pathways studied here. On all paths, all internal coordinates except the N–H3 distance are frozen along the dissociation pathway.

also very similar to the XMS-PDFT ones for this pathway; 7.5d compares CMS-PDFT with XMS-CASPT2, and we notice that the second excited state of CMS-PDFT is not as smooth as that of XMS-CASPT2 near 1.2 $\mathring{\text{A}}$.

In general, the agreement of CMS-PDFT with XMS-CASPT2 is satisfactory. In any given case, it is not known which of these two methods is more accurate (just as, for single-reference systems, sometimes MP2 is more accurate than density functional theory, but in many other cases density functional theory is more accurate). We note, in this context, that multireference perturbation theory sometimes suffers from intruder-state problems due to small energy denominators,[321] but MS-PDFT has no energy denominators and hence no intruder-state problems.

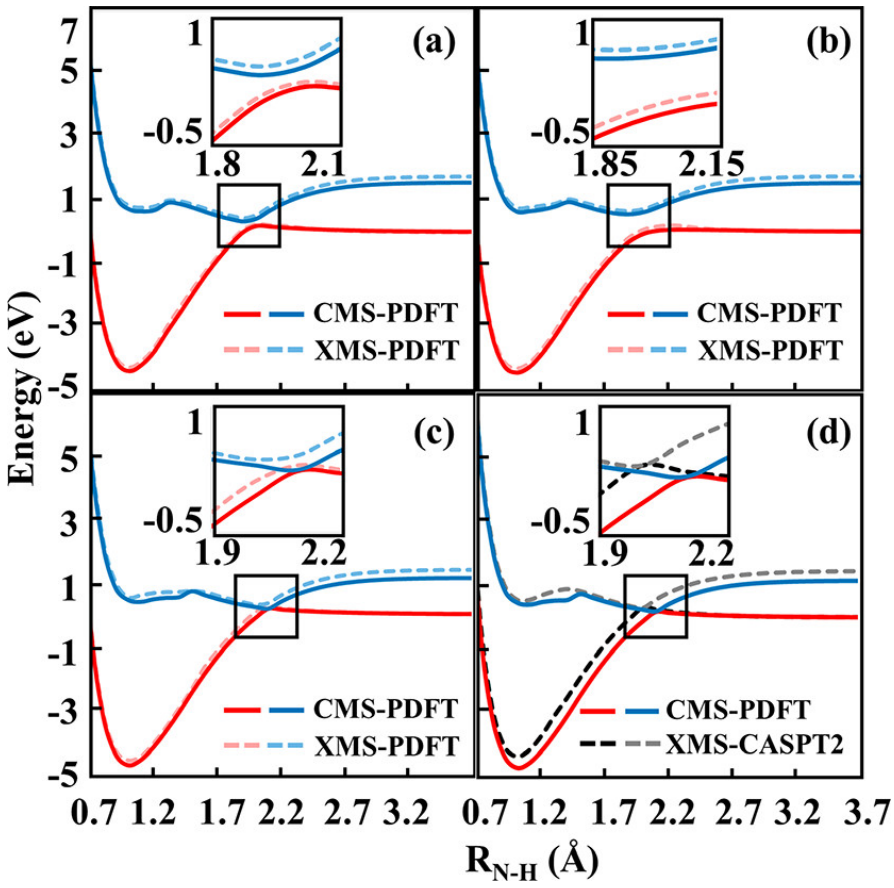

Figure 7.5: CMS-PDFT and XMS-PDFT potential energy curves for methylamine with the dihedral angle ϕ (defined in Figure 4) fixed at various values. (a) ϕ = 90°. (b) ϕ = 100°. (c,d) ϕ = 0.

7.4.5 Phenol

The potential energy surfaces and the dynamics of photodissociation of phenol have been studied many times[322, 323, 324, 325, 326, 327, 328, 329] not only for their own interest but also because phenol is a prototype of the $1\pi\sigma^*$ motif that is very important for biomolecules and other aromatic molecules such as thiophenol.[330, 331, 332, 333, 334, 335, 336]

7.6a,b show that XMS-PDFT and CMS-PDFT are also qualitatively similar. The CMS-PDFT method removes the bumpiness of the XMS-PDFT curves in 7.6d. As we discussed

before,[268] the bumpiness is due to a mismatch between where the diagonal elements are equal to each other and where the off-diagonal elements are maximal in the XMS-PDFT effective Hamiltonian matrix. This is not a surprising result in XMS-PDFT because XMS-PDFT intermediate states diagonalize the generalized Fock matrix; on the contrary, the CMS intermediate states do not suffer from the bumpiness because the CMS intermediate states contain some physics, namely, the sum of the classical Coulomb energies are maximized. Note that VMS-PDFT is also free of the bumpiness in phenol for a similar reason as shown in ref [268], but also note that CMS-PDFT is much more efficient than VMS-PDFT.

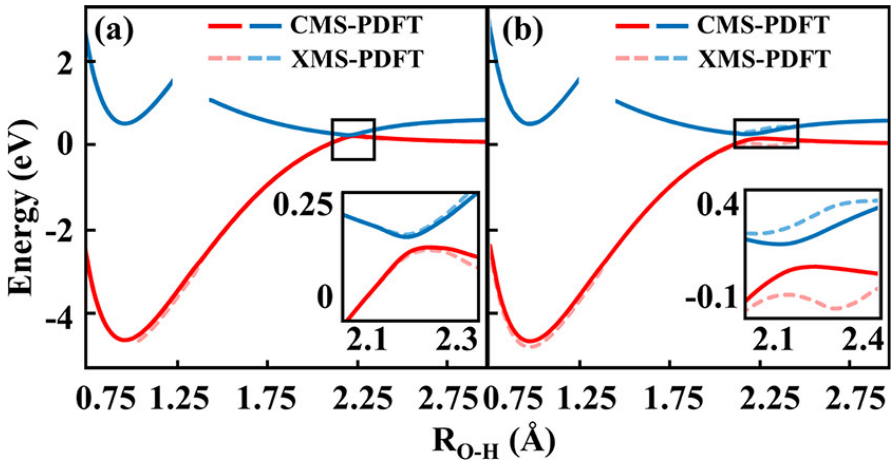

Figure 7.6: CMS-PDFT and XMS-CASPT2 potential energy curves for dissociation of the O–H bond of phenol with the two dihedral angles specified in Table 2; the dihedral angle is 1° in (a,c) and 10° in (b,d). We show a gap in the excited state potential curve in the region, where there is an avoided crossing with a third state; one would need a three-state calculation to get accurate results in that region.

7.4.6 Spiro

The 2,2′,6,6′-tetrahydro-4H,4′H-5,5′-spirobi[cyclopenta[c]pyrrole] cation, which will simply be called the spiro cation, contains two subsystems that share a bridging carbon. As seen in 7.7a,b, the carbon–nitrogen frame of the left-hand subsystem lies in the xz plane, whereas

that of the right-hand system is in the yz plane. The cation has an electron hole, and because of the Jahn–Teller effect, the ground-state equilibrium geometry locates the hole on either the left subsystem or the right subsystem. We then study a reaction pathway starting from where the hole is on the left subsystem to where the hole is on the right subsystem. For an atom A, where A is a carbon, nitrogen, or hydrogen atom of the molecule, we denote the Cartesian coordinates as x_A, y_A, an z_A. In the ground electronic state there are two minima in the potential energy surface, leading to twin equilibrium geometries. At the first of these, whose coordinates are labeled $x_A^{(1)}$, $y_A^{(1)}$, and $z_A^{(1)}$, the hole is localized on the left subsystem, and at the second, with coordinates $x_A^{(2)}$, $y_A^{(2)}$, and $z_A^{(2)}$, it is localized on the right subsystem. We consider a linear synchronous transit path[337] from geometry 1 to geometry 2. Along this path, the coordinates of the atoms vary linearly. For example

$$x_A(\xi) = (\frac{1}{2} - \xi)x_A^{(1)} + (\frac{1}{2} + \xi)x_A^{(2)} \tag{7.15}$$

where ξ is a unitless reaction coordinate. We previously studied the energies along this path by SI-PDFT[275] and VMS-PDFT[268] and both methods were shown to work well; however, XMS-PDFT did not give good results for this system.

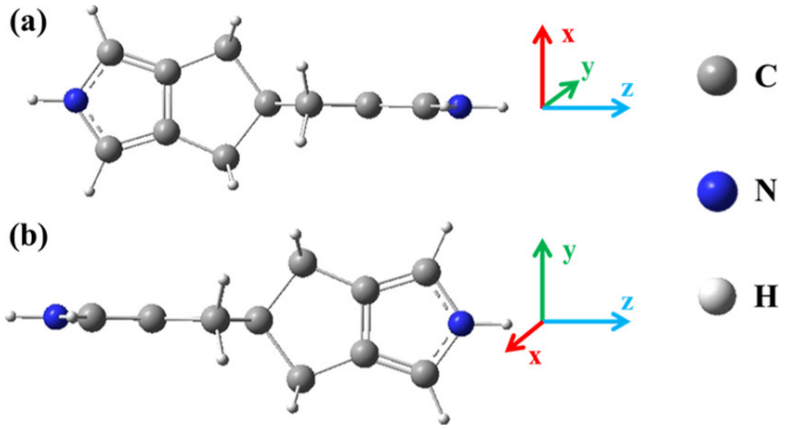

Figure 7.7: Two views of 2,2′,6,6′-tetrahydro-4H,4′H-5,5′-spirobi[cyclopenta[c]pyrrole] cation in the xz plane (a) and in the yz plane (b).

The CMS-PDFT potential energy curves and the XMS-CASPT2 potential energy curves

are shown in 7.8a. The CMS rotation angles and the XMS rotation angles are shown in Figure S2. The figure shows that the CMS-PDFT minimum energy separation is greater than that of XMS-CASPT2, although the results are qualitatively similar. Details of the intermediate states are given in Figure S2.

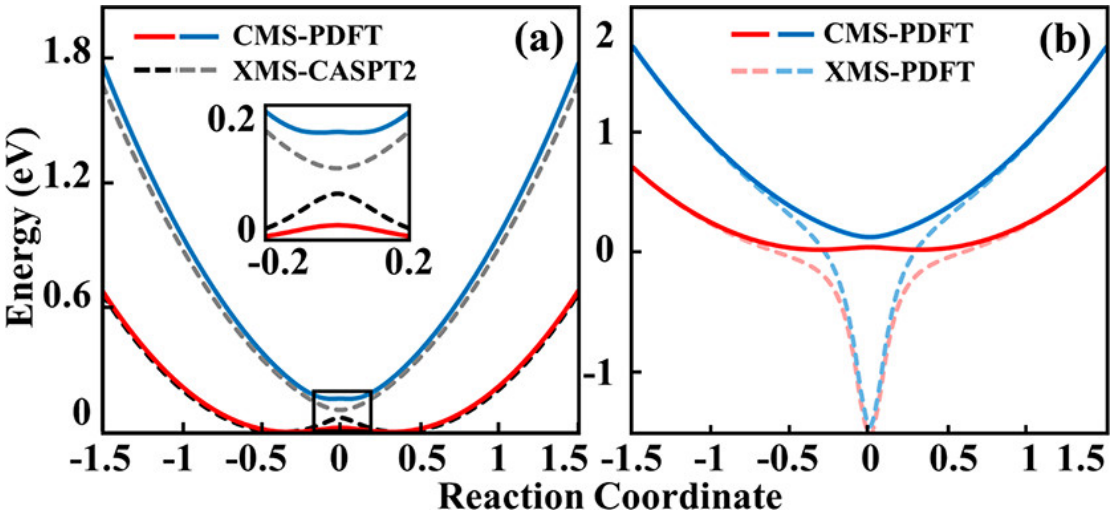

Figure 7.8: CMS-PDFT, XMS-CASPT2 and XMS-PDFT potential energy curves for the spiro cation.

The spiro cation is the only case among those tested where XMS-PDFT fails to work.[268] Because the hole is delocalized onto the two subsystems in the XMS intermediate states, XMS-PDFT gives a dip on the energy curves. 7.8b shows that CMS-PDFT corrects this problem.

7.4.7 Computational Cost of Using CMS-, XMS- and FMS- intermediate states

Because CMS-PDFT, XMS-PDFT, and FMS-PDFT generate intermediate states differently, we briefly discuss the cost to generate the intermediate states. Given that there are N states calculated and n_{a} active orbitals, the cost to generate XMS intermediate states is the cost to build the zero-order Hamiltonian matrix in XMS-CASPT2, which scales as

$(N+1)Nn_{\mathrm{a}}^2$. The cost to generate CMS intermediate states is the number of CMS iterations multiplied (because we use three-point fitting) by three times the cost to compute $Q_{a^\smile a}$, which scales as (N - 1)Nn_{a}^4. The cost to generate FMS intermediate states is three times the number of grid points times the cost to compute the on-top energies, which scales as (N - 1)Nn_{o}^4 times the number of grid points, where n_{o} is the number of occupied (including inactive and active) orbitals. Note that the magnitude of N is at most several tens, and n_{a} is also at most several tens, while n_{o} is a much greater number depending on the number of atoms and the size of the basis set. This indicates that the cost of generating CMS intermediate states is similar to that of generating XMS intermediate states but less than that for FMS intermediate states.

The times consumed in the program for obtaining intermediate states for the three methods are listed in Table 3. The times consumed for generating XMS or CMS intermediate states are significantly shorter than those for generating FMS intermediate states, which require three PDFT calculations in the systems shown in Table 3. The time to generate intermediate states is also much shorter than the time for obtaining the reference wave function, which is obtained from a complete-active-space CI[338] (CASCI) calculation.

7.5 Conclusion

We proposed a new MS-PDFT method, namely, CMS-PDFT, which maximizes the sum-over-intermediate-states of the classical electrostatic two-electron energies among active electrons. The new method, similar to the previous XMS-PDFT, computes the transformation matrix from the CASSCF wave functions without requiring additional PDFT energy calculations; however, it has the advantage of the VMS-PDFT method that it works in all cases tried. By comparing the CMS-PDFT potential energy curves with the XMS-CASPT2 potential energy curves, we showed that the new method works well. This method is as efficient as the

Table 7.3: Time (Seconds) Consumed for Generating Reference Wave Function, Generating Intermediate States and Computing PDFT Energies in OpenMolcas for Spiro and Phenol with Two States (N = 2)

System	N	time[a]				
		CASCI[b]	CMS	XMS	FMS[c]	PDFT
Spiro	2	5(2)	0[d]	0[d]	102	34
Phenol	2	21(12)	2	0[d]	330	110

[a]The time consumed is taken from the time consumed in corresponding modules in OpenMolcas. Reference wave functions (the CASCI column) and the state rotations are obtained from the RASSCF module in OpenMolcas, and the PDFT energies are calculated in the MCPDFT module in OpenMolcas.
[b]The numbers before the parentheses are the times consumed for the RASSCF module, and the numbers in the parentheses are the times consumed diagonalizing the CASCI Hamiltonian matrix.
[c]The FMS intermediate states need three PDFT calculations for two states.
[d]The time consumed for generating XMS or CMS intermediate states is taken as the difference of the time consumed for one RASSCF module that computes the intermediate states and for another RASSCF module that does not compute the intermediate states. Zero second means the seconds consumed for a RASSCF module that computes the intermediate states are the same as seconds needed for the RASSCF module that does not compute the intermediate states.

XMS-PDFT method in generating the intermediate states, and the CMS intermediate states are as robust as the VMS intermediate states for giving the correct topology for potential energy surfaces at and near conical intersection.

Our goal in the present work has been to develop an MC-PDFT method for calculating excited electronic states with all of the following desirable features: (i) it yields potential energy surfaces with the correct topology at and near conical intersections. (ii) It yields adiabatic states with the correct spatial, spin, and permutational symmetries. (iii) It does not involve a transformation of orbitals after the SA-CASSCF optimization, and in particular it only does state transformations using a single set of SA-CASSCF orbitals. (iv) It works as well or better than FMS-PDFT. (v) It does not require multiple PDFT calculations at each geometry (as are required in FMS-PDFT). The CMS-PDFT method presented here satisfies

all of these criteria. Furthermore, just as was the case for VMS-PDFT, the transformation to intermediate states is based on a physically motivated variational criterion, whereas the XMS method is motivated by the mathematical transformation properties associated with the generalized Fock operator rather than by a physical criterion. Therefore, we recommend the CMS-PDFT method for future MS-PDFT studies.

Chapter 8

Concluding Remarks

Our work has shown that MC-PDFT successfully predicts excitation energies and bond distances for both ground states and excited states. The ABC scheme is successful but is limited to the doublet molecules. The ABC2 scheme is more robust for studying low-lying excitation energies. Two multistate methods, XMS-PDFT and CMS-PDFT, provide correct topologies of potential energy surfaces at reasonable costs. Both are good candidates for studying photochemical problems. Because two automatic active space selection schemes are proposed for MC-PDFT, these schemes can be tested with XMS-PDFT or CMS-PDFT for excitation energies and even potential energy surfaces in the future work. The gradient of XMS-PDFT and CMS-PDFT can also be developed so they can be used to study photochemistry dynamics.

References

[1] B. G. Hunt. Photochemistry of ozone in a moist atmosphere. *J. Geophys. Res.*, 71:1385–1398, 1966.

[2] S. Han, C. E. Gunthardt, R. Dawes, D. Xie, S. W. North, and H. Guo. Origin of the "odd" behavior in the ultraviolet photochemistry of ozone. *Proc. Natl. Acad. Sci. U. S. A.*, 117:21065–21069, 2020.

[3] R. Wayne. The photochemistry of ozone. *Atmos. Environ.*, 21:1683–1694, 1987.

[4] R. Cox and G. Hayman. The stability and photochemistry of dimers of the Clo radical and implications for Antarctic ozone depletion. *Nature*, 332:796–800, 1988.

[5] R. Hu, S. Seager, and W. Bains. Photochemistry in terrestrial exoplanet atmospheres. II. H_2S and so_2 photochemistry in anoxic atmospheres. *Astrophys. J.*, 769:6, 2013.

[6] M. T. Martins-Costa, J. M. Anglada, J. S. Francisco, and M. F. Ruiz-Lopez. Photochemistry of so_2 at the air–water interface: A source of OH and HOSO radicals. *J. Amer. Chem. Soc.*, 140:12341–12344, 2018.

[7] Y. Zuo and J. Hoigné. Evidence for photochemical formation of H_2O_2 and oxidation of SO_2 in authentic fog water. *Science*, 260:71–73, 1993.

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

[9] M. Petersilka, U. Gossmann, and E. Gross. Excitation energies from time-dependent density-functional theory. *Phys. Rev. Lett.*, 76:1212, 1996.

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

[11] M. E. Casida. Time-dependent density functional response theory for molecules. In *Recent Advances In Density Functional Methods: (Part I)*, pages 155–192. World Scientific, 1995.

[12] M. Isegawa and D. G. Truhlar. Valence excitation energies of alkenes, carbonyl compounds, and azabenzenes by time-dependent density functional theory: Linear response of the ground state compared to collinear and noncollinear spin-flip TDDFT with the Tamm-Dancoff approximation. *J. Chem. Phys.*, 138:134111, 2013.

[13] A dressed TDDFT treatment of the 2^1A_g states of butadiene and hexatriene. *Chem. Phys. Lett*, 389:39 – 42, 2004.

[14] N. T. Maitra, F. Zhang, R. J. Cave, and K. Burke. Double excitations within time-dependent density functional theory linear response. *J. Chem. Phys.*, 120:5932–5937, 2004.

[15] K. Andersson, P. A. Malmqvist, B. O. Roos, A. J. Sadlej, and K. Wolinski. Second-order perturbation theory with a casscf reference function. *J. Chem. Phys.*, 94:5483–5488, 1990.

[16] B. O. Roos, P. Linse, P. E. M. Siegbahn, and M. R. A. Blomberg. A simple method for the evaluation of the second-order-perturbation energy from external double-excitations with a CASSCF reference wavefunction. *Chem. Phys.*, 66:197, 1982.

[17] K. Andersson, P. Å. Malmqvist, and B. O. Roos. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.*, 96:1218, 1992.

[18] J. K. Yu, C. Bannwarth, R. Liang, E. G. Hohenstein, and T. J. Martínez. Nonadiabatic dynamics simulation of the wavelength-dependent photochemistry of azobenzene excited to the nπ* and $\pi\pi$* excited states. *J. Amer. Chem. Soc.*

[19] G. Li Manni, R. K. Carlson, S. Luo, D. Ma, J. Olsen, D. G. Truhlar, and L. Gagliardi. Multiconfiguration pair-density functional theory. *J. Chem. Theory Comput.*, 10:3669–3680, 2014.

[20] L. Gagliardi, D. G. Truhlar, G. Li Manni, R. K. Carlson, C. E. Hoyer, and J. L. Bao. Multiconfiguration pair-density functional theory: A new way to treat strongly correlated systems. *Acc. Chem. Res.*, 50:66–73, 2017.

[21] S. Ghosh, P. Verma, C. J. Cramer, L. Gagliardi, and D. G. Truhlar. Combining wave function methods with density functional theory for excited states. *Chem. Rev.*, 118:7249, 2018.

[22] A. M. Sand, D. G. Truhlar, and L. Gagliardi. Efficient algorithm for multiconfiguration pair-density functional theory with application to the heterolytic dissociation energy of ferrocene. *J. Chem. Phys.*, 146:034101, 2017.

[23] C. E. Hoyer, S. Ghosh, D. G. Truhlar, and L. Gagliardi. Multiconfiguration pair-density functional theory is as accurate as CASPT2 for electronic excitation. *J. Phys. Chem. Lett.*, 7:586, 2016.

[24] J. J. Bao, L. Gagliardi, and D. G. Truhlar. Multiconfiguration pair-density functional theory for doublet excitation energies and excited state geometries: the excited states of CN. *Phys. Chem. Chem. Phys.*, 19:30089–30096, 2017.

[25] G. Poletto and M. Rigutti. The $A^2\Pi$ and $X^2\Sigma$ states of the CN molecule from the berkeley analysis of the CN red system. *Il Nuovo Cimento*, 39:519–530, 1965.

[26] N. Balucani, O. Asvany, L. Huang, Y. Lee, R. Kaiser, Y. Osamura, and H. Bettinger. Formation of nitriles in the interstellar medium via reactions of cyano radicals, CN($X^2\Sigma^+$), with unsaturated hydrocarbons. *Astrophys. J.*, 545:892, 2000.

[27] S. B. Morales, C. J. Bennett, S. D. Le Picard, A. Canosa, I. R. Sims, B. Sun, P. Chen, A. H. Chang, V. V. Kislov, A. M. Mebel, et al. A crossed molecular beam, low-temperature kinetics, and theoretical investigation of the reaction of the cyano radical (CN) with 1, 3-butadiene (C_4H_6). A route to complex nitrogen-bearing molecules in low-temperature extraterrestrial environments. *Astrophys. J.*, 742:26, 2011.

[28] F. Jenkins, Y. K. Roots, and R. S. Mulliken. The red CN band system. *Phys. Rev.*, 39:16, 1932.

[29] R. Ram, S. Davis, L. Wallace, R. Engleman, D. R. Appadoo, and P. Bernath. Fourier transform emission spectroscopy of the $B^2\sigma^+$–$X^2\sigma^+$ system of CN. *J. Mol. Spectrosc.*, 237:225–231, 2006.

[30] K.-P. Huber. *Molecular spectra and molecular structure: IV. Constants of diatomic molecules*. Springer Science & Business Media, 1966.

[31] S. E. Bradforth, E. H. Kim, D. W. Arnold, and D. M. Neumark. Photoelectron spectroscopy of CN^-, NCO^-, and NCS^-. *J. Chem. Phys.*, 98:800–810, 1993.

[32] P. A. Pieniazek, S. E. Bradforth, and A. I. Krylov. Spectroscopy of the cyano radical in an aqueous environment. *J. Phys. Chem. A*, 110:4854–4865, 2006.

[33] A. I. Krylov. The quantum chemistry of open-shell species. *Rev. Comput. Chem.*, 30:151–224, 2017.

[34] F. Guérin. Etats de valence $^2\Pi$ et Δ des molécules isoélectroniques N_2^+, CN et CO^+. *Theor. Chim. Acta.*, 17:97–108, 1970.

[35] H. F. Schaefer III and T. G. Heil. Electronic structures and potential energy curves for the low-lying states of the CN radical. *J. Chem. Phys.*, 54:2573–2580, 1971.

[36] G. Das, T. Janis, and A. C. Wahl. Ground and excited states of the diatoms CN and AlO. *J. Chem. Phys.*, 61:1274–1279, 1974.

[37] J. Hinze. MC-SCF. I. the multi-configuration self-consistent-field method. *J. Chem. Phys.*, 59:6424–6432, 1973.

[38] J. Guan, M. Casida, and D. Salahub. Time-dependent density-functional theory investigation of excitation spectra of open-shell molecules. *J. Mol. Struc.:THEOCHEM*, 527:229–244, 2000.

[39] S. Hirata and M. Head-Gordon. Time-dependent density functional theory for radicals: An improved description of excited states with substantial double excitation character. *Chem. Phys. Lett.*, 302:375–382, 1999.

[40] D.-h. Shi, W.-t. Li, J.-f. Sun, and Z.-l. Zhu. Mrci study on spectroscopic and molecular properties of several low-lying electronic states of the CN radical. *J. Quant. Spectrosc. Radiat. Transf.*, 112:2335–2346, 2011.

[41] P. J. Knowles and H.-J. Werner. An efficient method for the evaluation of coupling coefficients in configuration interaction calculations. *Chem. Phys. Lett.*, 145:514–522, 1988.

[42] H.-J. Werner and P. J. Knowles. An efficient internally contracted multiconfiguration-reference configuration interaction method. *J. Chem. Phys.*, 89:5803–5814, 1988.

[43] S. R. Langhoff and E. R. Davidson. Configuration interaction calculations on the nitrogen molecule. *Int. J. Quantum Chem.*, 8:61–72, 1974.

[44] A. K. Wilson, D. E. Woon, K. A. Peterson, and T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. IX. the atoms gallium through krypton. *J. Chem. Phys.*, 110:7667–7676, 1999.

[45] R. K. Carlson, G. Li Manni, A. L. Sonnenberger, D. G. Truhlar, and L. Gagliardi. Multiconfiguration pair-density functional theory: Barrier heights and main group and transition metal energetics. *J. Chem. Theory Comput.*, 11:82–90, 2015.

[46] S. Ghosh, A. L. Sonnenberger, C. E. Hoyer, D. G. Truhlar, and L. Gagliardi. Multiconfiguration pair-density functional theory outperforms Kohn–Sham density functional theory and multireference perturbation theory for ground-state and excited-state charge transfer. *J. Chem. Theory Comput.*, 11:3643–3649, 2015.

[47] J. L. Bao, A. Sand, L. Gagliardi, and D. G. Truhlar. Correlated-participating-orbitals pair-density functional method and application to multiplet energy splittings of main-group divalent radicals. *J. Chem. Theory Comput.*, 12:4274–4283, 2016.

[48] D. C. Comeau and R. J. Bartlett. The equation-of-motion coupled-cluster method. Applications to open-and closed-shell reference states. *Chem. Phys. Lett.*, 207:414–423, 1993.

[49] J. F. Stanton and R. J. Bartlett. The equation of motion coupled-cluster method. A systematic biorthogonal approach to molecular excitation energies, transition probabilities, and excited state properties. *J. Chem. Phys.*, 98:7029–7039, 1993.

[50] K. Kowalski and P. Piecuch. New coupled-cluster methods with singles, doubles, and noniterative triples for high accuracy calculations of excited electronic states. *J. Chem. Phys.*, 120:1715–1738, 2004.

[51] P. Piecuch, J. R. Gour, and M. Włoch. Left-eigenstate completely renormalized equation-of-motion coupled-cluster methods: Review of key concepts, extension to excited states of open-shell systems, and comparison with electron-attached and ionized approaches. *Int. J. Quantum Chem.*, 109:3268–3304, 2009.

[52] D. Sinha, S. Mukhopadhyay, R. Chaudhuri, and D. Mukherjee. The eigenvalue-independent partitioning technique in Fock space: An alternative route to open-shell coupled-cluster theory for incomplete model spaces. *Chem. Phys. Lett.*, 154:544–549, 1989.

[53] J. F. Stanton and J. Gauss. Analytic energy derivatives for ionized states described by the equation-of-motion coupled cluster method. *J. Chem. Phys.*, 101:8938–8944, 1994.

[54] J. Pople and R. Nesbet. Self-consistent orbitals for radicals. *J. Chem. Phys.*, 22:571–572, 1954.

[55] K. Kowalski and P. Piecuch. Renormalized CCSD(T) and CCSD(TQ) approaches: Dissociation of the N_2 triple bond. *J. Chem. Phys.*, 113:5644–5652, 2000.

[56] M. J. McGuire, K. Kowalski, and P. Piecuch. Renormalized coupled-cluster calculations of reactive potential energy surfaces: A comparison of the CCSD(T), renormalized CCSD(T), and full configuration interaction results for the collinear BeFH system. *J. Chem. Phys.*, 117:3617–3624, 2002.

[57] P. Piecuch, S. A. Kucharski, V. Špirko, and K. Kowalski. Can ordinary single-reference coupled-cluster methods describe potential energy surfaces with nearly spectroscopic accuracy? The renormalized coupled-cluster study of the vibrational spectrum of HF. *J. Chem. Phys.*, 115:5796–5804, 2001.

[58] E. Papajak, H. R. Leverentz, J. Zheng, and D. G. Truhlar. Efficient diffuse basis sets: cc-pV*X*Z+ and maug-cc-pV*X*Z. *J. Chem. Theory Comput.*, 5:1197–1202, 2009.

[59] M. J. Frisch, G. W. Trucks, G. E. R. M. A. Schlegel, H. B.and Scuseria, J. R. Cheeseman, V. Scalmani, G.and Barone, B. Mennucci, G. A. Petersson, and *et al.* Gaussian09 *Revision* C.01 (Gaussian, Inc., Wallingford CT), 2016.

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

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

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

[63] Y. Zhao and D. G. Truhlar. A new local density functional for main-group thermochemistry, transition metal bonding, thermochemical kinetics, and noncovalent interactions. *J. Chem. Phys.*, 125:194101, 2006.

[64] Y. Zhao and D. G. Truhlar. Density functionals with broad applicability in chemistry. *Acc. Chem. Res.*, 41:157–167, 2008.

[65] R. Peverati and D. G. Truhlar. M11-L: A local density functional that provides improved accuracy for electronic structure calculations in chemistry and physics. *J. Chem. Phys. Lett.*, 3:117–124, 2012.

[66] H. S. Yu, X. He, S. L. Li, and D. G. Truhlar. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.*, 7:5032–5051, 2016.

[67] Y. Zhao and D. G. Truhlar. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.*, 120:215–241, 2008.

[68] R. Peverati and D. G. Truhlar. Improving the accuracy of hybrid meta-GGA density functionals by range separation. *J. Chem. Phys. Lett.*, 2:2810–2817, 2011.

[69] S. Y. Haoyu, X. He, S. L. Li, and D. G. Truhlar. MN15: A Kohn–Sham global-hybrid exchange–correlation density functional with broad accuracy for multi-reference and single-reference systems and noncovalent interactions. *Chem. Sci.*, 7:5032–5051, 2016.

[70] M. Valiev, E. J. Bylaska, N. Govind, K. Kowalski, T. P. Straatsma, H. J. Van Dam, D. Wang, J. Nieplocha, E. Apra, T. L. Windus, et al. NWChem: A comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.*, 181:1477–1489, 2010.

[71] M. W. Schmidt, K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, et al. General atomic and molecular electronic structure system. *J. Comput. Chem.*, 14:1347–1363, 1993.

[72] M. S. Gordon and M. W. Schmidt. Advances in electronic structure theory: GAMESS a decade later. In *Theory and applications of computational chemistry*, pages 1167–1189. Elsevier, 2005.

[73] B. O. Roos, P. R. Taylor, and P. E. Sigbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.*, 48:157–173, 1980.

[74] F. Aquilante, J. Autschbach, R. K. Carlson, L. F. Chibotaru, M. G. Delcey, L. De Vico, I. Fdez. Galván, N. Ferré, L. M. Frutos, L. Gagliardi, et al. Molcas 8: New capabilities for multiconfigurational quantum chemical calculations across the periodic table. *J. Comput. Chem.*, 37:506–541, 2016.

[75] J. Finley, P.-Å. Malmqvist, B. O. Roos, and L. Serrano-Andrés. The multi-state CASPT2 method. *Chem. Phys. Lett.*, 288:299–306, 1998.

[76] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, M. Schütz, P. Celani, W. Györffy, D. Kats, T. Korona, R. Lindh, A. Mitrushenkov, G. Rauhut, K. R. Shamasundar, T. B. Adler, R. D. Amos, S. J. Bennie, A. Bernhardsson, A. Berning, D. L. Cooper, M. J. O. Deegan, A. J. Dobbyn, F. Eckert, E. Goll, C. Hampel, A. Hesselmann, G. Hetzer, T. Hrenar, G. Jansen, C. Köppl, S. J. R. Lee, Y. Liu, A. W. Lloyd, Q. Ma, R. A. Mata, A. J. May, S. J. McNicholas, W. Meyer, T. F. Miller III, M. E. Mura, A. Nicklass, D. P. O'Neill, P. Palmieri, D. Peng, K. Pflüger, R. Pitzer, M. Reiher, T. Shiozaki, H. Stoll, A. J. Stone, R. Tarroni, T. Thorsteinsson, M. Wang, and M. Welborn. Molpro, version 2012.1, a package of ab initio programs. see https://www.molpro.net.

[77] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz. Molpro: a general-purpose quantum chemistry program package. *WIREs Comput. Mol. Sci.*, 2:242–253, 2012.

[78] R. K. Carlson, D. G. Truhlar, and L. Gagliardi. Multiconfiguration pair-density functional theory: a fully translated gradient approximation and its performance for transition metal dimers and the spectroscopy of $Re^2Cl_8^{2-}$. *J. Chem. Theory Comput.*, 11:4077–4085, 2015.

[79] Y. Zhang and W. Yang. Comment on "Generalized gradient approximation made simple". *Phys. Rev. Lett.*, 80:890, 1998.

[80] O. Tishchenko, J. Zheng, and D. G. Truhlar. Multireference model chemistries for thermochemical kinetics. *J. Chem. Theory Comput.*, 4:1208–1219, 2008.

[81] S. Kotochigova, T. Zelevinsky, and J. Ye. Prospects for application of ultracold Sr_2 molecules in precision measurements. *Phys. Rev. A*, 79:12504, 2009.

[82] S. De, U. Dammalapati, K. Jungmann, and L. Willmann. Magneto-optical trapping of barium. *Phys. Rev. A*, 79:41402, 2009.

[83] U. Dammalapati, I. Norris, C. Burrows, A. S. Arnold, and E. Riis. Spectroscopy and isotope shifts of the 4s3d $^1D_2$–4s5p $^1P_1$ repumping transition in magneto-optically trapped calcium atoms. *Phys. Rev. A*, 81:23424, 2010.

[84] J. M. Merritt, V. E. Bondybey, and M. C. Heaven. Beryllium dimer–caught in the act of bonding. *Science*, 324:1548, 2009.

[85] A. M. Maniero and P. H. Acioli. Full configuration interaction pseudopotential determination of the ground-state potential energy curves of Li_2 and LiH. *Int. J. Quantum Chem.*, 103:711, 2005.

[86] R. A. Aziz and M. J. Slaman. The Ne-Ne interatomic potential revisited. *Chem. Phys.*, 130:187, 1989.

[87] S. N. Khanna, F. Reuse, and J. Buttet. Stability and observability of charged beryllium clusters. *Phys. Rev. Lett.*, 61:535, 1988.

[88] V. Kumar and R. Car. Structure, growth, and bonding nature of Mg clusters. *Phys. Rev. B: Condens. Matter Mater. Phys.*, 44:8243, 1991.

[89] K. Duanmu, O. Roberto-Neto, F. B. C. Machado, J. A. Hansen, J. Shen, P. Piecuch, and D. G. Truhlar. Geometries, binding energies, ionization potentials, and electron affinities of metal clusters: Mg_nO, $\pm$ 1, n = 1–7. *J. Phys. Chem. C*, 120:13275, 2016.

[90] J. W. Mirick, C.-H. Chien, and E. Blaisten-Barojas. Electronic structure of calcium clusters. *Phys. Rev. A*, 63:23202, 2001.

[91] G. M. Wang, E. Blaisten-Barojas, A. E. Roitberg, and T. P. Martin. Strontium clusters: Many-body potential, energetics, and structural transitions. *J. Chem. Phys.*, 115:3640, 2001.

[92] Y. Wang, H.-J. Flad, and M. Dolg. Ab initio study of structure and bonding of strontium clusters. *J. Phys. Chem. A*, 104:5558, 2000.

[93] V. Boutou, A. R. Allouche, F. Spiegelmann, J. Chevaleyre, and M. A. Frécon. Predictions of geometrical structures and ionization potentials for small barium clusters Ba. *Eur. Phys. J. D*, 2:63, 1998.

[94] A. Ruzsinszky, J. P. Perdew, and G. I. Csonka. Binding energy curves from nonempirical density functionals II. van der Waals bonds in rare-gas and alkaline-earth diatomics. *J. Phys. Chem. A*, 109:11015, 2005.

[95] Y. Zhao and D. G. Truhlar. Comparative DFT study of van der Waals complexes: Rare-gas dimers, alkaline-earth dimers, zinc dimer, and zinc-rare-gas dimers. *J. Phys. Chem. A*, 110:5121, 2006.

[96] M. W. Schmidt, J. Ivanic, and K. Ruedenberg. Electronic structure analysis of the ground-state potential energy curve of Be_2. *J. Phys. Chem. A*, 114:8687, 2010.

[97] G. Herzberg. Zum Aufbau der zweiatomigen Moleküle. *Eur. Phys. J. A*, 57:601, 1929.

[98] L. Herzberg. Über ein neues bandensystem des berylliumoxyds und die struktur des Be O-moleküls. *Eur. Phys. J. A*, 84:571, 1933.

[99] C. F. Bender and E. R. Davidson. Theoretical calculation of the potential curves of the Be_2 molecule. *J. Chem. Phys.*, 47:4972, 1967.

[100] B. Liu and A. D. McLean. Ab initio potential curve for Be_2 ($1\Sigma_g^+$) from the interacting correlated fragments method. *J. Chem. Phys.*, 72:3418, 1980.

[101] B. H. Lengsfield, A. D. McLean, M. Yoshimine, and B. Liu. The binding energy of the ground state of Be_2. *J. Chem. Phys.*, 79:1891, 1983.

[102] J. Stärck and W. Meyer. The ground state potential of the beryllium dimer. *Chem. Phys. Lett.*, 258:421, 1996.

[103] F. Schautz, H. J. Flad, and M. Dolg. Quantum Monte Carlo study of Be_2 and group 12 dimers M_2 (M = Zn, Cd, Hg). *Theor. Chem. Acc.*, 99:231, 1998.

[104] K. D. Jordan. Electronic structure of small metal clusters. I. Anions of Be_2, Be_3, and Be_4. *J. Chem. Phys.*, 67:4027, 1999.

[105] R. A. Chiles and C. E. Dykstra. An electron pair operator approach to coupled cluster wave functions. Application to He_2, Be_2, and Mg_2 and comparison with CEPA methods. *J. Chem. Phys.*, 74:4544, 1981.

[106] Y. S. Lee and R. J. Bartlett. A study of Be_2 with many-body perturbation theory and a coupled-cluster method including triple excitations. *J. Chem. Phys.*, 80:4371, 1984.

[107] J. M. Brom, W. D. Hewett, and W. Weltner. Optical spectra of Be atoms and Be_2 molecules in rare gas matrices. *J. Chem. Phys.*, 62:3122, 1975.

[108] K. Patkowski, V. Spirko, and K. Szalewicz. On the elusive twelfth vibrational state of beryllium dimer. *Science*, 326:1382, 2009.

[109] M. R. A. Blomberg and P. E. M. Siegbahn. Beryllium dimer, a critical test case of MBPT and CI methods. *Int. J. Quantum Chem.*, 14:583, 1978.

[110] S. Evangelisti, G. L. Bendazzoli, and L. Gagliardi. Full configuration interaction calculations on Be_2. *Chem. Phys.*, 185:47, 1994.

[111] I. Magoulas, N. P. Bauman, J. Shen, and P. Piecuch. Application of the CC(P;Q) hierarchy of coupled-cluster methods to the beryllium dimer. *J. Phys. Chem. A*, 122:1350, 2018.

[112] W. J. Balfour and A. E. Douglas. Absorption spectrum of the Mg_2 molecule. *Can. J. Phys.*, 48:901, 1970.

[113] C. R. Vidal and H. Scheingraber. Determination of diatomic molecular constants using an inverted perturbation approach. *J. Mol. Spectrosc.*, 65:46, 1977.

[114] G. D. Purvis and R. J. Bartlett. The potential energy curve for the X $^1\Sigma_g^+$ state of Mg_2 calculated with many-body perturbation theory. *J. Chem. Phys.*, 68:2114, 1978.

[115] H. Partridge, C. W. Bauschlicher, L. G. M. Pettersson, A. D. McLean, B. Liu, M. Yoshimine, and A. Komornicki. On the dissociation energy of Mg_2. *J. Chem. Phys.*, 92:5377, 1990.

[116] P. Li, W. Xie, and K. T. Tang. The van der Waals potential of the magnesium dimer. *J. Chem. Phys.*, 133:84308, 2010.

[117] C. R. Vidal. The molecular constants and potential energy curves of the Ca_2 A$^1\Sigma_u^+$ – X$^1\Sigma_g^+$ system from laser induced fluorescence. *J. Chem. Phys.*, 72:1864, 1980.

[118] O. Allard, C. Samuelis, A. Pashov, H. Knöckel, and E. Tiemann. Experimental study of the Ca_2 ^{1}S + ^{1}S asymptote. *Eur. Phys. J. D*, 26:155–164, 2003.

[119] D. D. Yang, P. Li, and K. T. Tang. The ground state van der Waals potentials of the calcium dimer and calcium rare-gas complexes. *J. Chem. Phys.*, 131:154301, 2009.

[120] T. Bergeman and P. F. Liao. Photoassociation, photoluminescence, and collisional dissociation of the Sr_2 dimer. *J. Chem. Phys.*, 72:886, 1980.

[121] A. Stein, H. Knöckel, and E. Tiemann. Fourier-transform spectroscopy of Sr_2 and revised ground-state potential. *Phys. Rev. A*, 78:42508, 2008.

[122] G. P. Yin, P. Li, and K. T. Tang. The ground state van der Waals potentials of the strontium dimer and strontium rare-gas complexes. *J. Chem. Phys.*, 132:74303, 2010.

[123] A. R. Allouche, M. Aubert-Frécon, G. Nicolas, and F. Spiegelmann. Theoretical study of the electronic structure of the Ba_2 molecule. *Chem. Phys.*, 200:63, 1995.

[124] S. Schäfer, M. Mehring, R. Schäfer, and P. Schwerdtfeger. Polarizabilities of Ba and Ba_2: Comparison of molecular beam experiments with relativistic quantum chemistry. *Phys. Rev. A*, 76:52515, 2007.

[125] K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.*, 157:479, 1989.

[126] W. Kutzelnigg. Perturbation theory of relativistic corrections. *Z. Phys. D: At., Mol. Clusters*, 11:15, 1989.

[127] W. Kutzelnigg. Perturbation theory of relativistic corrections. *Z. Phys. D: At., Mol. Clusters*, 15:27, 1990.

[128] S. Stopkowicz and J. Gauss. Relativistic corrections to electrical first-order properties using direct perturbation theory. *J. Chem. Phys.*, 129:164119, 2008.

[129] D.-D. Yang and F. Wang. Theoretical investigation for spectroscopic constants of ground-state alkaline-earth dimers with high accuracy. *Theor. Chem. Acc.*, 131:1117, 2012.

[130] V. Veryazov, P.-O. Widmark, and B. O. Roos. Relativistic atomic natural orbital type basis sets for the alkaline and alkaline-earth atoms applied to the ground-state potentials for the corresponding dimers. *Theor. Chem. Acc.*, 111:345, 2004.

[131] T. Q. Teodoro, R. L. Haiduke, U. Dammalapati, S. Knoop, and L. Visscher. The ground-state potential energy curve of the radium dimer from relativistic coupled cluster calculations. *J. Chem. Phys.*, 143:84307, 2015.

[132] S. Kristyán and P. Pulay. Can (semi)local density functional theory account for the London dispersion forces? *Chem. Phys. Lett.*, 229:175, 1994.

[133] J. Pérez-Jordá and A. D. Becke. A density-functional study of van der Waals forces: rare gas diatomics. *Chem. Phys. Lett.*, 233:134, 1995.

[134] W. Kohn, Y. Meir, and D. E. Makarov. van der Waals energies in density functional theory. *Phys. Rev. Lett.*, 80:4153, 1998.

[135] Q. Wu and W. Yang. Empirical correction to density functional theory for van der Waals interactions. *J. Chem. Phys.*, 116:515, 2002.

[136] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.*, 132:154104, 2010.

[137] Y. Zhao, B. J. Lynch, and D. G. Truhlar. Doubly hybrid meta DFT: New multi-coefficient correlation and density functional methods for thermochemistry and thermochemical kinetics. *J. Phys. Chem. A*, 108:4786, 2004.

[138] V. R. Cooper, T. Thonhauser, and D. C. Langreth. An application of the van der Waals density functional: Hydrogen bonding and stacking interactions between nucleobases. *J. Chem. Phys.*, 128:204102, 2008.

[139] A. Tarnopolsky, A. Karton, R. Sertchook, D. Vuzman, and J. M. L. Martin. Double-hybrid functionals for thermochemical kinetics. *J. Phys. Chem. A*, 112:3, 2008.

[140] Y. Zhang, X. Xu, and W. A. Goddard. Doubly hybrid density functional for accurate descriptions of nonbond interactions, thermochemistry, and thermochemical kinetics. *Proc. Natl. Acad. Sci. U. S. A.*, 106:4963, 2009.

[141] O. A. Vydrov and T. Van Voorhis. Benchmark assessment of the accuracy of several van der Waals density functionals. *J. Chem. Theory Comput.*, 8:1929, 2012.

[142] J. Klimes and A. Michaelides. Perspective: Advances and challenges in treating van der Waals dispersion forces in density functional theory. *J. Chem. Phys.*, 137:120901, 2012.

[143] Y. Zhao and D. G. Truhlar. Applications and validations of the Minnesota density functionals. *Chem. Phys. Lett.*, 502:1, 2011.

[144] E. Goll, H.-J. Werner, H. Stoll, T. Leininger, P. Gori-Giorgi, and A. Savin. A short-range gradient-corrected spin density functional in combination with long-range coupled-cluster methods: Application to alkali-metal rare-gas dimers. *Chem. Phys.*, 329:276, 2006.

[145] B. O. Roos, P. R. Taylor, and P. E. M. Siegbahn. A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach. *Chem. Phys.*, 48:157, 1980.

[146] B. O. Roos, R. Lindh, P. A. Malmqvist, V. Veryazov, and P. O. Widmark. Main group atoms and dimers studied with a new relativistic ANO basis set. *J. Phys. Chem. A*, 108:2851, 2004.

[147] J. L. Bao, S. O. Odoh, L. Gagliardi, and D. G. Truhlar. Predicting bond dissociation energies of transition-metal compounds by multiconfiguration pair-density functional theory and second-order perturbation theory based on correlated participating orbitals and separated pairs. *J. Chem. Theory Comput.*, 13:616, 2017.

[148] G. Ghigo, B. O. Roos, and P.-Å. Malmqvist. A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2). *Chem. Phys. Lett.*, 396:142, 2004.

[149] N. Forsberg and P. A. Malmqvist. Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.*, 274:196, 1997.

[150] M. Douglas and N. M. Kroll. Quantum electrodynamical corrections to the fine structure of helium. *Ann. Phys.*, 82:89, 1974.

[151] B. A. Hess. Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators. *Phys. Rev. A*, 33:3742, 1986.

[152] P. Sharma, D. G. Truhlar, and L. Gagliardi. Active Space Dependence in Multiconfiguration Pair-Density Functional Theory. *J. Chem. Theory Comput.*, 14:660, 2018.

[153] E. Tiesinga, S. Kotochigova, and P. S. Julienne. Scattering length of the ground-state Mg + Mg collision. *Phys. Rev. A*, 65:42722, 2002.

[154] O. Allard, A. Pashov, H. Knöckel, and E. Tiemann. Ground-state potential of the Ca dimer from Fourier-transform spectroscopy. *Phys. Rev. A*, 66:42503, 2002.

[155] N. C. Handy and A. J. Cohen. Left-right correlation energy. *Mol. Phys.*, 99:403, 2001.

[156] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.*, 46:618, 1934.

[157] R. J. Bartlett and G. D. Purvis. Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem. *Int. J. Quantum Chem.*, 14:561, 1978.

[158] J. M. L. Martin, A. Sundermann, P. L. Fast, and D. G. Truhlar. Thermochemical analysis of core correlation and scalar relativistic effects on molecular atomization energies. *J. Chem. Phys.*, 113:1348, 2000.

[159] M. B. Sullivan, M. A. Iron, P. C. Redfern, J. M. L. Martin, L. A. Curtiss, and L. Radom. Heats of formation of alkali metal and alkaline earth metal oxides and hydroxides: Surprisingly demanding targets for high-level ab initio procedures. *J. Phys. Chem. A*, 107:5617, 2003.

[160] H. Yu and D. G. Truhlar. Components of the bond energy in polar diatomic molecules, radicals, and ions formed by group-1 and group-2 metal atoms. *J. Chem. Theory Comput.*, 11:2968, 2015.

[161] M. Barysz and A. J. Sadlej. Two-component methods of relativistic quantum chemistry: from the Douglas–Kroll approximation to the exact two-component formalism. *J. Mol. Struct.: THEOCHEM*, 573:181, 2001.

[162] W. Liu and D. Peng. Exact two-component Hamiltonians revisited. *J. Chem. Phys.*, 131:31104, 2009.

[163] P. O. Löwdin and H. Shull. Natural orbitals in the quantum theory of two-electron systems. *Phys. Rev.*, 101:1730, 1956.

[164] C. F. Bender and E. R. Davidson. A natural orbital based energy calculation for helium hydride and lithium hydride. *J. Phys. Chem.*, 70:2675, 1966.

[165] K. V. Darvesh and F. Grein. Configuration selection in the MCSCF method. I. Application to the $B^1\Sigma^+$ state of HF. *Int. J. Quantum Chem.*, 28:247, 1985.

[166] F. Illas, J. Rubio, J. M. Ricart, and P. S. Bagus. Selected versus complete configuration interaction expansions. *J. Chem. Phys.*, 95:1877, 1991.

[167] L. Serrano-Andrés, J. Sánchez-Marín, and I. Nebot-Gil. Theoretical study of the low-lying states of trans-1,3-butadiene. *J. Chem. Phys.*, 97:7499, 1992.

[168] M. W. Schmidt and M. S. Gordon. The construction and interpretation of MCSCF wave functions. *Annu. Rev. Phys. Chem.*, 49:233, 1998.

[169] P. Pulay and T. P. Hamilton. UHF natural orbitals for defining and starting MC-SCF calculations. *J. Chem. Phys.*, 88:4926, 1988.

[170] K. Kowalski and P. Piecuch. The active-space equation-of-motion coupled-cluster methods for excited electronic states: Full EOMCCSDt. *J. Chem. Phys.*, 115:643, 2001.

[171] D. Zgid and M. Nooijen. The density matrix renormalization group self-consistent field method: Orbital optimization with the density matrix renormalization group method in the active space. *J. Chem. Phys.*, 128:144116, 2008.

[172] F. Aquilante, T. K. Todorova, L. Gagliardi, T. B. Pedersen, and B. O. Roos. Systematic truncation of the virtual space in multiconfigurational perturbation theory. *J. Chem. Phys.*, 131:34113, 2009.

[173] R. Izsák, M. Szori, P. J. Knowles, and B. Viskolcz. High accuracy ab initio calculations on reactions of OH with 1-alkenes. the case of propene. *J. Chem. Theory Comput.*, 5:2313, 2009.

[174] V. Veryazov, P. A. Malmqvist, and B. O. Roos. How to select active space for multiconfigurational quantum chemistry? *Int. J. Quantum Chem.*, 111:3329, 2011.

[175] K. J. H. Giesbertz. Are natural orbitals useful for generating an efficient expansion of the wave function? *Chem. Phys. Lett.*, 591:220, 2014.

[176] S. S. Ray, A. Ghosh, S. Chattopadhyay, and R. K. Chaudhuri. Taming the electronic structure of diradicals through the window of computationally cost effective multireference perturbation theory. *J. Phys. Chem. A*, 120:5897, 2016.

[177] J. L. Bao, S. O. Odoh, L. Gagliardi, and D. G. Truhlar. Predicting bond dissociation energies of transition metal compounds by multiconfiguration pair-density functional theory and second-order perturbation theory based on correlated participating orbitals and separated pairs. *J. Chem. Theory Comput.*, 13:616, 2017.

[178] S. J. Stoneburner, J. Shen, A. O. Ajala, P. Piecuch, D. G. Truhlar, and L. Gagliardi. Systematic design of active spaces for multi-reference calculations of singlet–triplet gaps of organic diradicals, with benchmarks against doubly electron-attached coupled-cluster data. *J. Chem. Phys.*, 147:164120, 2017.

[179] C. J. Stein and M. Reiher. Automated selection of active orbital spaces. *J. Chem. Theory Comput.*, 12:1760, 2016.

[180] C. J. Stein, V. von Burg, and M. Reiher. The delicate balance of static and dynamic electron correlation,. *J. Chem. Theory Comput.*, 12:3764, 2016.

[181] S. Ghosh, C. J. Cramer, D. G. Truhlar, and L. Gagliardi. Generalized-active-space pair-density functional theory: An efficient method to study large, strongly correlated, conjugated systems. *Chem. Sci.*, 8:2741, 2017.

[182] E. R. Sayfutyarova, Q. Sun, G. K.-L. Chan, and G. Knizia. Automated construction of molecular active spaces from atomic valence orbitals. *J. Chem. Theory Comput.*, 13:4063, 2017.

[183] C. J. Stein and M. Reiher. Automated identification of relevant frontier orbitals for chemical compounds and processes. *Chimia*, 71:170, 2017.

[184] A. C. West. Atom-based strong correlation method: An orbital selection algorithm. *J. Phys. Chem. A*, 121:8912, 2017.

[185] K. Kim, G. Frenking, and C. Dykstra. *Theory and Applications of Computational Chemistry: the First Forty Years*. Elsevier Science Limited, 2005.

[186] J. A. Pople. Nobel lecture: Quantum chemical models. *Rev. Mod. Phys.*, 71:1267, 1999.

[187] C. D. Sherrill and H. F. Schaefer. The configuration interaction method. *Adv. Quantum Chem.*, 34:143, 1999.

[188] F. B. Brown, I. Shavitt, and R. Shepard. Multireference configuration interaction treatment of potential energy surfaces: Symmetric dissociation of H_2O in a double-zeta basis. *Chem. Phys. Lett.*, 105:363, 1984.

[189] P. J. Knowles and H.-J. Werner. Internally contracted multiconfiguration-reference configuration interaction calculations for excited states. *Theor. Chim. Acta*, 84:95, 1992.

[190] H.-J. Werner, M. Kállay, and J. Gauss. The barrier height of the F + H_2 reaction revisited: Coupled-cluster and multireference configuration-interaction benchmark calculations. *J. Chem. Phys.*, 128:34305, 2008.

[191] W. J. Hunt and W. A. Goddard. Excited states of H_2O using improved virtual orbitals. *Chem. Phys. Lett.*, 3:414, 1969.

[192] W. R. Wadt and W. A. Goddard. The low-lying excited states of water, methanol, and dimethyl ether. *Chem. Phys.*, 18:1, 1976.

[193] D. M. Potts, C. M. Taylor, R. K. Chaudhuri, and K. F. Freed. Improved virtual orbital-complete active space configuration interaction method, a 'packageable' efficient ab initio many-body method for describing electronically excited states. *J. Chem. Phys.*, 114:2592, 2001.

[194] P.-Å. Malmqvist, A. Rendell, and B. O. Roos. The restricted active space self-consistent-field method, implemented with a split graph unitary group approach. *J. Phys. Chem.*, 94:5477, 1990.

[195] E. Hirota and C. Yamada. Intramolecular motions and molecular structure of the CH_3 radical. *J. Mol. Spectrosc.*, 96:175, 1982.

[196] G. Herzberg. *Molecular Spectra and Molecular Structure*. 1966.

[197] R. C. Fortenberry, R. A. King, J. F. Stanton, and T. D. Crawford. A benchmark study of the vertical electronic spectra of the linear chain radicals C_2H and C_4H. *J. Chem. Phys.*, 132:144303, 2010.

[198] P. Zhang and K. Morokuma. Ab initio molecular orbital study of the weak $\tilde{\mathrm{C}}^2\mathrm{A}' \leftarrow \tilde{\mathrm{X}}^2\mathrm{A}'$. *Chem. Phys. Lett.*, page 482.

[199] J. Gauss, J. F. Stanton, and R. J. Bartlett. Analytic ROHF–MBPT(2) second derivatives. *J. Chem. Phys.*, 97:7825, 1992.

[200] Z. Li and W. Liu. Critical assessment of td-dft for excited states of open-shell systems: I. Doublet-doublet transitions. *J. Chem. Theory Comput.*, 12:238, 2016.

[201] M. Isegawa, R. Peverati, and D. G. Truhlar. Performance of recent and high-performance approximate density functionals for time-dependent density functional theory calculations of valence and rydberg electronic transition energies. *J. Chem. Phys.*, 137:244104, 2012.

[202] M. Isegawa and D. G. Truhlar. Valence excitation energies of alkenes, carbonyl compounds, and azabenzenes by time-dependent density functional theory: Linear response of the ground state compared to collinear and noncollinear spin-flip tddft with the Tamm-Dancoff approximation. *J. Chem. Phys.*, 138:134111, 2013.

[203] R. Haunschild, A. Barth, and W. Marx. Evolution of DFT studies in view of a scientometric perspective. *J. Cheminf.*, 8:52, 2016.

[204] C. Angeli, R. Cimiraglia, S. Evangelisti, T. Leininger, and J. P. Malrieu. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.*, 114:10252, 2001.

[205] C. Angeli, R. Cimiraglia, and J. P. Malrieu. n-electron valence state perturbation theory: A spinless formulation and an efficient implementation of the strongly contracted and of the partially contracted variants. *J. Chem. Phys.*, 117:9138, 2002.

[206] H. Nakano. MCSCF reference quasidegenerate perturbation theory with Epstein—Nesbet partitioning. *Chem. Phys. Lett.*, 207:372, 1993.

[207] H. Nakano. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.*, 99:7983, 1993.

[208] C. Angeli, S. Borini, M. Cestari, and R. Cimiraglia. A quasidegenerate formulation of the second order n-electron valence state perturbation theory approach. *J. Chem. Phys.*, 121:4043, 2004.

[209] A. A. Granovsky. Extended multi-configuration quasi-degenerate perturbation theory: the new approach to multi-state multi-reference perturbation theory. *J. Chem. Phys.*, 134:214113, 2011.

[210] V. Veryazov, P. Å. Malmqvist, and B. O. Roos. How to select active space for multiconfigurational quantum chemistry? *Int. J. Quantum Chem.*, 111:3329–3338, 2011.

[211] J. M. Bofill and P. Pulay. The unrestricted natural orbital-complete active space (UNO-CAS) method: An inexpensive alternative to the complete active space-self-consistent-field (CAS-SCF) method. *J. Chem. Phys.*, 90:3637, 1989.

[212] H. Yu and J. D. Goddard. Multiconfiguration self-consistent-field (MCSCF) approaches to organic structures and reactivities. *J. Mol. Struct.: THEOCHEM*, 233:129, 1991.

[213] R. G. A. Bone and P. Pulay. Half-projected Hartree–Fock natural orbitals for defining CAS–SCF active spaces. *Int. J. Quantum Chem.*, 45:133, 1993.

[214] S. Keller, K. Boguslawski, T. Janowski, M. Reiher, and P. Pulay. Selection of active spaces for multiconfigurational wavefunctions. *J. Chem. Phys.*, 142:244104, 2015.

[215] B. S. Fales, Y. Shu, B. G. Levine, and E. G. Hohenstein. Complete active space configuration interaction from state-averaged configuration interaction singles natural orbitals: Analytic first derivatives and derivative coupling vectors. *J. Chem. Phys.*, 147:94104, 2017.

[216] S. R. White. Density matrix formulation for quantum renormalization groups. *Phys. Rev. Lett.*, 69:2863–2866, 1992.

[217] S. R. White. Density-matrix algorithms for quantum renormalization groups. *Phys. Rev. B*, 48:10345–10356, 1993.

[218] C. J. Stein and M. Reiher. Measuring multi-configurational character by orbital entanglement. *Mol. Phys.*, 115:2110, 2017.

[219] C. J. Stein and M. Reiher. Automated selection of active orbital spaces. *J. Chem. Theory Comput.*, 12:1760–1771, 2016.

[220] C. J. Stein and M. Reiher. autocas: A program for fully automated multiconfigurational calculations. *J. Comput. Chem.*, 40:2216–2226, 2019.

[221] J. J. Bao, S. S. Dong, L. Gagliardi, and D. G. Truhlar. Automatic selection of an active space for calculating electronic excitation spectra by MS-CASPT2 or MC-PDFT. *J. Chem. Theory Comput.*, 14:2017–2025, 2018.

[222] F. Krausbeck, D. Mendive-Tapia, A. J. Thom, and M. J. Bearpark. Choosing RASSCF orbital active spaces for multiple electronic states. *Comput. Theor. Chem.*, 1040:14, 2014.

[223] E. R. Sayfutyarova and S. Hammes-Schiffer. Constructing molecular π-orbital active spaces for multireference calculations of conjugated systems. *J. Chem. Theory Comput.*, 15:1679, 2019.

[224] A. Khedkar and M. Roemelt. Active space selection based on natural orbital occupation numbers from n-electron valence perturbation theory. *J. Chem. Theory Comput.*, 15:3522, 2019.

[225] Y. Shao, M. Head-Gordon, and A. I. Krylov. The spin–flip approach within time-dependent density functional theory: Theory and applications to diradicals. *J. Chem. Phys.*, 118:4807, 2003.

[226] K. Yang, R. Peverati, D. G. Truhlar, and R. Valero. Density functional study of multiplicity-changing valence and Rydberg excitations of p-block elements: Delta

self-consistent field, collinear spin-flip time-dependent density functional theory (DFT), and conventional time-dependent DFT. *J. Chem. Phys.*, 135:44118, 2011.

[227] S. V. Levchenko and A. I. Krylov. Equation-of-motion spin-flip coupled-cluster model with single and double substitutions: Theory and application to cyclobutadiene. *J. Chem. Phys.*, 120:175, 2004.

[228] A. I. Krylov. Spin-flip configuration interaction: an electronic structure model that is both variational and size-consistent. *Chem. Phys. Lett.*, 350:522, 2001.

[229] M. W. Schmidt, E. A. Hull, and T. L. Windus. Valence virtual orbitals: An unambiguous ab initio quantification of the LUMO concept. *J. Phys. Chem. A*, 119:10408, 2015.

[230] E. Papajak, J. Zheng, X. Xu, H. R. Leverentz, and D. G. Truhlar. Perspectives on basis sets beautiful: Seasonal plantings of diffuse basis functions. *J. Chem. Theory Comput.*, 7:3027, 2011.

[231] I. Fdez. Galván, M. Vacher, A. Alavi, C. Angeli, F. Aquilante, J. Autschbach, J. J. Bao, S. I. Bokarev, N. A. Bogdanov, R. K. Carlson, L. F. Chibotaru, J. Creutzberg, N. Dattani, M. G. Delcey, S. S. Dong, A. Dreuw, L. Freitag, L. M. Frutos, L. Gagliardi, F. Gendron, A. Giussani, L. González, G. Grell, M. Guo, C. E. Hoyer, M. Johansson, S. Keller, S. Knecht, G. Kovačević, E. Källman, G. Li Manni, M. Lundberg, Y. Ma, S. Mai, J. P. Malhado, P. Å. Malmqvist, P. Marquetand, S. A. Mewes, J. Norell, M. Olivucci, M. Oppel, Q. M. Phung, K. Pierloot, F. Plasser, M. Reiher, A. M. Sand, I. Schapiro, P. Sharma, C. J. Stein, L. K. Sørensen, D. G. Truhlar, M. Ugandi, L. Ungur, A. Valentini, S. Vancoillie, V. Veryazov, O. Weser, T. A. Wesołowski, P. O. Widmark, S. Wouters, A. Zech, J. P. Zobel, and R. Lindh. Openmolcas: From source code to insight, 2019.

[232] M. J. Frisch, G. W. Trucks, G. E. R. M. A. Schlegel, H. B.and Scuseria, J. R. Cheeseman, V. Scalmani, G.and Barone, B. Mennucci, G. A. Petersson, and *et al.* Gaussian16 *Revision* B.01 (Gaussian, Inc., Wallingford CT), 2016.

[233] P. F. Loos, A. Scemama, A. Blondel, Y. Garniron, M. Caffarel, and D. Jacquemin. A mountaineering strategy to excited states: Highly accurate reference energies and benchmarks. *J. Chem. Theory Comput.*, 14:4360, 2018.

[234] K. Ralphs, G. Serna, L. R. Hargreaves, M. A. Khakoo, C. Winstead, and V. McKoy. Excitation of the six lowest electronic transitions in water by 9–20 eV electrons. *J. Phys. B: At., Mol. Opt. Phys.*, 46:125201, 2013.

[235] H. Masuko, Y. Morioka, M. Nakamura, E. Ishiguro, and M. Sasanuma. Absorption spectrum of the H_2S molecule in the vacuum ultraviolet region. *Can. J. Phys.*, 57:745, 1979.

[236] M. J. G. Peach, P. Benfield, T. Helgaker, and D. J. Tozer. Excitation energies in density functional theory: An evaluation and a diagnostic test. *J. Chem. Phys.*, 128:44118, 2008.

[237] A. Skerbele and E. N. Lassettre. Electron-impact spectra. *J. Chem. Phys.*, 42:395, 1965.

[238] R. Dressler and M. Allan. A dissociative electron attachment, electron transmission, and electron energy loss study of the temporary negative ion of acetylene. *J. Chem. Phys.*, 87:4510, 1987.

[239] S. P. McGlynn, J. W. Rabalais, J. R. McDonald, and V. M. Scherr. Electronic spectroscopy of isoelectronic molecules. II. Linear triatomic groupings containing sixteen valence electrons. *Chem. Rev.*, 71:73, 1971.

[240] G. Bélanger and C. Sandorfy. Far-ultraviolet spectra of fluoroethylenes. *J. Chem. Phys.*, 55:2055, 1971.

[241] D. Feller, K. A. Peterson, and E. R. Davidson. A systematic approach to vertically excited states of ethylene using configuration interaction and coupled cluster techniques. *J. Chem. Phys.*, 141:104302, 2014.

[242] M. B. Robin, H. Basch, N. A. Kuebler, K. B. Wiberg, and G. B. Ellison. Optical spectra of small rings. II. the unsaturated three-membered rings. *J. Chem. Phys.*, 51:45, 1969.

[243] I. Sauers, L. A. Grezzo, S. W. Staley, and J. H. Moore. Low energy singlet-triplet and singlet-singlet transitions in cycloalkenes. *J. Am. Chem. Soc.*, 98:4218, 1976.

[244] Y. Shu and D. G. Truhlar. Doubly excited character or static correlation of the reference state in the controversial 2^1A_g state of trans-butadiene? *J. Am. Chem. Soc.*, 139:13770, 2017.

[245] J. P. Doering and R. McDiarmid. Electron impact study of the energy levels of trans-1,3-butadiene: II. Detailed analysis of valence and Rydberg transitions. *J. Chem. Phys.*, 73:3617, 1980.

[246] M. A. Watson and G. K.-L. Chan. Excited states of butadiene to chemical accuracy: Reconciling theory and experiment. *J. Chem. Theory Comput.*, 8:4013, 2012.

[247] P. Sharma, V. Bernales, D. G. Truhlar, and L. Gagliardi. Valence $\pi\pi*$ excitations in benzene studied by multiconfiguration pair-density functional theory. *J. Phys. Chem. Lett.*, 10:75, 2019.

[248] A. Hiraya and K. Shobatake. Direct absorption spectra of jet-cooled benzene in 130–260 nm. *J. Chem. Phys.*, 94:7700, 1991.

[249] M. B. Robin. *Higher Excited States of Polyatomic Molecules*, volume III. 1985.

[250] K. N. Walzl, C. F. Koerting, and A. Kuppermann. Electron-impact spectroscopy of acetaldehyde. *J. Chem. Phys.*, 87:3796, 1987.

[251] J. Oddershede, N. E. Gruner, and G. H. Diercksen. Comparison between equation of motion and polarization propagator calculations. *Chem. Phys.*, 97:303, 1985.

[252] E. S. Nielsen, P. Jorgensen, and J. Oddershede. Transition moments and dynamic polarizabilities in a second order polarization propagator approach. *J. Chem. Phys.*, 73:6238, 1980.

[253] D. Truhlar and Y. Zhao. The M06 suite of density functionals for main group thermochemistry, thermochemical kinetics, noncovalent interactions, excited states, and transition elements: two new functionals and systematic testing of four M06-class functionals and 12 other functionals. *Theor. Chem. Acc.*, 120:215–241, 2008.

[254] P. Verma and D. G. Truhlar. HLE16: A local Kohn–Sham gradient approximation with good performance for semiconductor band gaps and molecular excitation energies. *J. Chem. Phys. Lett.*, 8:380–387, 2017.

[255] T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. the atoms boron through neon and hydrogen. *J. Chem. Phys.*, 90:1007–1023, 1989.

[256] J. L. Whitten and M. Hackmeyer. Configuration interaction studies of ground and excited states of polyatomic molecules. I. The ci formulation and studies of formaldehyde. *J. Chem. Phys.*, 51:5584, 1969.

[257] A. D. McLean and B. Liu. Classification of configurations and the determination of interacting and noninteracting spaces in configuration interaction. *J. Chem. Phys.*, 58:1066, 1973.

[258] R. J. Buenker, S. D. Peyerimhoff, and W. Butscher. Applicability of the multi-reference double-excitation CI (MRD-CI) method to the calculation of electronic wavefunctions and comparison with related techniques. *Mol. Phys.*, 35:771, 1978.

[259] H.-J. Werner. Problem decompositon in quantum chemistry. In D. E. Keyes, Y. Saad, and D. G. Truhlar, editors, *Domain-Based Paralleism and Problem Decompostion Methods in Computational Science and Engineering*, page 239– 261. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1995.

[260] J. Čížek. On the use of the cluster expansion and the technique of diagrams in calculations of correlation effects in atoms and molecules. *Adv. Chem. Phys.*, 14:35, 2007.

[261] J. Paldus, J. Čížek, and I. Shavitt. Correlation problems in atomic and molecular systems. IV. extended coupled-pair many-electron theory and its application to the BH_3 molecule. *Phys. Rev. A*, 5:50, 1972.

[262] R. J. Bartlett and G. D. Purvis. Molecular applications of coupled cluster and many-body perturbation methods. *Phys. Scr.*, 21:255, 1980.

[263] P. Piecuch. Active-space coupled-cluster methods. *Mol. Phys.*, 108:2987, 2010.

[264] K. Kowalski, S. Krishnamoorthy, O. Villa, J. R. Hammond, and N. Govind. Active-space completely-renormalized equation-of-motion coupled-cluster formalism: Excited-state studies of green fluorescent protein, free-base porphyrin, and oligoporphyrin dimer. *J. Chem. Phys.*, 132:154103, 2010.

[265] A. Melnichuk and R. J. Bartlett. Relaxed active space: Fixing tailored-CC with high order coupled cluster. *J. Chem. Phys.*, 140:64113, 2014.

[266] A. K. Dutta, M. Nooijen, F. Neese, and R. Izsák. Automatic active space selection for the similarity transformed equations of motion coupled cluster method. *J. Chem. Phys.*, 146:74103, 2017.

[267] L. M. J. Huntington, M. Krupička, F. Neese, and R. Izsák. Similarity transformed equation of motion coupled-cluster theory based on an unrestricted Hartree-Fock reference for applications to high-spin open-shell systems. *J. Chem. Phys.*, 147:174104, 2017.

[268] J. J. Bao, C. Zhou, Z. Varga, S. Kanchanakungwankul, L. Gagliardi, and D. G. Truhlar. Multi-state pair-density functional theory. *Faraday Discuss.*, 224:348–372, 2020.

[269] H. S. Yu, S. L. Li, and D. G. Truhlar. Perspective: Kohn-Sham density functional theory descending a staircase. *J. Chem. Phys.*, 145, 2016.

[270] W. Domcke, H. Köppel, and L. S. Cederbaum. Spectroscopic effects of conical intersections of molecular potential energy surfaces. *Mol. Phys.*, 43:851–875, 1981.

[271] A. W. Jasper, B. K. Kendrick, C. A. Mead, and D. G. Trular. Non-born–oppenheimer chemistry: Potential surfaces, couplings, and dynamics. In X. Yang and K. Liu, editors, *Modern Trends in Chemical Reaction Dynamics: Experiment and Theory (Part 1)*, page 329–391. World Scientific, Singapore, 2004.

[272] R. Maurice, P. Verma, J. M. Zadrozny, S. Luo, J. Borycz, J. R. Long, D. G. Truhlar, and L. Gagliardi. Single-ion magnetic anisotropy and isotropic magnetic couplings in the metal-organic framework Fe_2(dobdc). *Inorg. Chem.*, 52:9379–9389, 2013.

[273] M. Atanasov, J. M. Zadrozny, J. R. Long, and F. Neese. A theoretical analysis of chemical bonding, vibronic coupling, and magnetic anisotropy in linear iron (II) complexes with single-molecule magnet behavior. *Chem. Sci.*, 4:139–156, 2013.

[274] K. Hirao. Multireference Møller-Plesset method. *Chem. Phys. Lett.*, 190:374–380, 1992.

[275] A. M. Sand, C. E. Hoyer, D. G. Truhlar, and L. Gagliardi. State-interaction pair-density functional theory. *J. Chem. Phys.*, 149:024106, 2018.

[276] S. S. Dong, K. B. Huang, L. Gagliardi, and D. G. Truhlar. State-interaction pair-density functional theory can accurately describe a spiro mixed valence compound. *J. Chem. Phys. A*, 123:2100–2106, 2019.

[277] C. Zhou, L. Gagliardi, and D. G. Truhlar. State-interaction pair density functional theory for locally avoided crossings of potential energy surfaces in methylamine. *Phys. Chem. Chem. Phys.*, 21:13486–13493, 2019.

[278] T. Shiozaki, W. Gyroffy, P. Celani, and H. J. Werner. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.*, 135, 2011.

[279] M. B. Robin and P. Day. Mixed valence chemistry-a survey and classification. In H. J. Emeléus and A. G. Sharpe, editors, *Advances in Inorgorganic Chemistry and Radiochemistry*, pages 247–422. Elsevier, 1968.

[280] L. R. Kahn, P. J. Hay, and I. Shavitt. Theoretical study of curve crossing: Ab initio calculations on the four lowest $1\Sigma^+$ states of LiF. *J. Chem. Phys.*, 61:3530–3546, 1974.

[281] B. J. Botter, J. A. Kooter, and J. J. Mulder. Ab-initio calculations of the covalent-ionic curve crossing in LiF. *Chem. Phys. Lett.*, 33:532–534, 1975.

[282] H. J. Werner and W. Meyer. MCSCF study of the avoided curve crossing of the two lowest $1\Sigma^+$ states of LiF. *J. Chem. Phys.*, 74:5802–5807, 1981.

[283] J. P. Finley and H. A. Witek. Diagrammatic complete active space perturbation theory: Calculations on benzene, N_2, and LiF. *J. Chem. Phys.*, 112:3958–3963, 2000.

[284] J. Meller, J. Malrieu, and J. Heully. Size-consistent multireference configuration interaction method through the dressing of the norm of determinants. *Mol. Phys.*, 101:2029–2041, 2003.

[285] Legeza, J. RÖDER, and B. A. Hess. QC-DMRG study of the ionic-neutral curve crossing of LiF. *Mol. Phys.*, 101:2019–2028, 2003, 0208187.

[286] C. Angeli, R. Cimiraglia, and J.-P. Malrieu. A simple approximate perturbation approach to quasi-degenerate systems. *Theor. Chem. Acc.*, 116:434–439, 2006.

[287] M. Hanrath. Multi-reference coupled-cluster study of the ionic-neutral curve crossing LiF. *Mol. Phys.*, 106:1949–1957, 2008.

[288] F. Sasaki and M. Yoshimine. Configuration-interaction study of atoms. II. Electron affinities of B, C, N, O, and F. *Phys. Rev. A*, 9:26–34, 1974.

[289] B. O. Roos. The complete active space self-consistent field method and its applications in electronic structure calculations. *Adv. Chem. Phys.*, 69:399, 1987.

[290] K. Ruedenberg, L. M. Cheung, and S. T. Elbert. MCSCF optimization through combined use of natural orbitals and the Brillouin-Levy-Berthier theorem. *Int. J. Quantum Chem.*, 16:1069, 1979.

[291] M. Garavelli. Computational organic photochemistry: Strategy, achievements and perspectives. *Theor. Chem. Acc.*, 116:87, 2006.

[292] P. E. M. Siegbahn, J. Almlöf, A. Heiberg, and B. O. Roos. The complete active space SCF (CASSCF) method in a Newton–Raphson formulation with application to the HNO molecule. *J. Chem. Phys.*, 74:2384, 1981.

[293] C. Edmiston and K. Ruedenberg. Localized atomic and molecular orbitals. *Rev. Mod. Phys.*, 35:457, 1963.

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

[295] J. L. Bao, Y. Wang, X. He, L. Gagliardi, and D. G. Truhlar. Multiconfiguration pair-density functional theory is free from delocalization error. *Phys. Chem. Lett.*, 8:5616, 2017.

[296] D. Feller. The role of databases in support of computational chemistry calculations. *J. Comput. Chem.*, 17:1571–1586, 1996.

[297] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus. Basis set exchange: A community database for computational sciences. *J. Chem. Inf. Model*, 47:1045–1052, 2007.

[298] P. C. Hariharan and J. A. Pople. The influence of polarization functions on molecular orbital hydrogenation energies. *Theor. Chim. Acta.*, 28:213–222, 1973.

[299] C. Li, R. Lindh, and F. A. Evangelista. Dynamically weighted multireference perturbation theory: Combining the advantages of multi-state and state-averaged methods. *J. Chem. Phys.*, 150:144107, 2019.

[300] R. J. Fallon, J. T. Vanderslice, and E. A. Mason. Potential energy curves for lithium hydride. *J. Chem. Phys.*, 32:1453, 1960.

[301] K. C. Li and W. C. Stwalley. The $A^1\Sigma^+ \rightarrow X^1\Sigma^+$ bands of the isotopic lithium hydrides. *J. Mol. Spectrosc.*, 69:294, 1978.

[302] A. Pardo, J. J. Camacho, and J. M. L. Poyato. The Padé-approximant method and its applications in the construction of potential-energy curves for the lithium hydride molecule. *Chem. Phys. Lett.*, 131:490, 1986.

[303] W. C. Stwalley and W. T. Zemke. Spectroscopy and structure of the lithium hydride diatomic molecules and ions. *J. Phys. Chem. Ref. Data*, 22:87, 1993.

[304] W.-C. Tung, M. Pavanello, and L. Adamowicz. Very accurate potential energy curve of the LiH molecule. *J. Chem. Phys.*, 134:64117, 2011.

[305] D. P. Taylor, C. F. Dion, and E. R. Bernstein. On the electronic states and photochemistry of simple alkyl amines. *J. Chem. Phys.*, 106:3512, 1997.

[306] J. V. Michael and W. A. Noyes. The photochemistry of methylamine. *J. Am. Chem. Soc.*, 85:1228–1233, 1963.

[307] E. Kassab, E. M. Evleth, and J. T. Gleghorn. Theoretical aspects of the photochemistry of methanol, methylamine, and related materials. *J. Am. Chem. Soc.*, 105:1746–1753, 1983.

[308] G. C. G. Waschewsky, D. C. Kitchen, P. W. Browning, and L. J. Butler. Competing bond fission and molecular elimination channels in the photodissociation of CH_3NH_2 at 222 nm. *J. Chem. Phys.*, 99:2635–2645, 1995.

[309] C. L. Reed, M. Kono, and M. N. Ashfold. Near-UV photolysis of methylamine studied by H-atom photofragment translational spectroscopy. *J. Chem. Soc. Faraday Trans.*, 92:4897–4904, 1996.

[310] K. M. Dunn and K. Morokuma. Ab initio study of the photochemical dissociation of methylamine. *J. Chem. Phys.*, 100:123–129, 1996.

[311] S. J. Baek, K. W. Choi, Y. S. Choi, and S. K. Kim. Spectroscopy and dynamics of methylamine. I. Rotational and vibrational structures of CH_3NH_2 and CH_3ND_2 in Ã states. *J. Chem. Phys.*, 118:11026–11039, 2003.

[312] M. H. Park, K. W. Choi, S. Choi, S. K. Kim, and Y. S. Choi. Vibrational structures of methylamine isotopomers in the predissociative Ã states: CH_3NHD, CD_3NH_2, CD_3NHD, and CD_3ND_2. *J. Chem. Phys.*, 125, 2006.

[313] A. Golan, S. Rosenwaks, and I. Bar. Mode-dependent enhancement of photodissociation and photoionization in a seven atom molecule. *J. Chem. Phys.*, 125:151103, 2006.

[314] D. S. Ahn, J. Lee, J. M. Choi, K. S. Lee, S. J. Baek, K. Lee, K. K. Baeck, and S. K. Kim. State-selective predissociation dynamics of methylamines: The vibronic and HD effects on the conical intersection dynamics. *J. Chem. Phys.*, 128, 2008.

[315] C. Levi, R. Kosloff, Y. Zeiri, and I. Bar. Time-dependent quantum wave-packet description of H and D atom tunneling in N-H and N-D photodissociation of methylamine and methylamine-d_2. *J. Chem. Phys.*, 131, 2009.

[316] R. Marom, C. Levi, T. Weiss, S. Rosenwaks, Y. Zeiri, R. Kosloff, and I. Bar. Quantum tunneling of hydrogen atom in dissociation of photoexcited methylamine. *J. Chem. Phys. A*, 114:9623–9627, 2010.

[317] D. S. Ahn, J. Lee, Y. Choon Park, Y. Sup Lee, and S. Kyu Kim. Nuclear motion captured by the slow electron velocity imaging technique in the tunnelling predissociation of the S_1 methylamine. *J. Chem. Phys.*, 136, 2012.

[318] J. O. Thomas, K. E. Lower, and C. Murray. Observation of NH $X^3\Sigma^-$ as a primary product of methylamine photodissociation: Evidence of roaming-mediated intersystem crossing? *J. Chem. Phys. Lett.*, 3:1341–1345, 2012.

[319] H. Xiao, S. Maeda, and K. Morokuma. Theoretical study on the photodissociation of methylamine involving S_1, T_1, and S_0 states. *J. Chem. Phys. A*, 117:5757–5764, 2013.

[320] M. Epshtein, Y. Yifrach, A. Portnov, and I. Bar. Control of nonadiabatic passage through a conical intersection by a dynamic resonance. *J. Chem. Phys. Lett.*, 7:1717–1724, 2016.

[321] Y.-K. Choe, H. A. Witek, J. P. Finley, and K. Hirao. Identifying and removing intruder states in multireference Møller–Plesset perturbation theory. *J. Chem. Phys.*, 114:3913, 2001.

[322] C.-M. Tseng, Y. T. Lee, and C.-K. Ni. H atom elimination from the $\pi\sigma*$ state in the photodissociation of phenol. *J. Chem. Phys.*, 121:2459, 2004.

[323] C.-M. Tseng, Y. T. Lee, M.-F. Lin, C.-K. Ni, S.-Y. Liu, Y.-P. Lee, Z. F. Xu, and M. C. Lin. Photodissociation dynamics of phenol. *J. Phys. Chem. A*, 111:9463, 2007.

[324] M. L. Hause, Y. Heidi Yoon, A. S. Case, and F. F. Crim. Dynamics at conical intersections: The influence of O–H stretching vibrations on the photodissociation of phenol. *J. Chem. Phys.*, 128:104307, 2008.

[325] R. N. Dixon, T. A. A. Oliver, and M. N. R. Ashfold. Tunneling under a conical intersection: Application to the product vibrational state distributions in the UV photodissociation of phenols. *J. Chem. Phys.*, 134:194303, 2011.

[326] X. Xu, K. R. Yang, D. G. Truhlar, and D. G. Truhlar. Diabatic molecular orbitals, potential energies, and potential energy surface couplings by the fourfold way for photodissociation of phenol. *J. Chem. Theory Comput.*, 9:3612, 2013.

[327] X. Xu, J. Zheng, K. R. Yang, and D. G. Truhlar. Photodissociation dynamics of phenol: Multistate trajectory simulations including tunneling. *J. Am. Chem. Soc.*, 136:16378, 2014.

[328] C. Xie, J. Ma, X. Zhu, D. R. Yarkony, D. Xie, and H. Guo. Nonadiabatic tunneling in photodissociation of phenol. *J. Am. Chem. Soc.*, 138:7828, 2016.

[329] O. P. Vieuxmaire, Z. Lan, A. L. Sobolewski, and W. Domcke. Ab initio characterization of the conical intersections involved in the photochemistry of phenol. *J. Chem. Phys.*, 129:224307, 2008.

[330] A. L. Sobolewski, W. Domcke, C. Dedonder-Lardeux, and C. Jouvet. Excited-state hydrogen detachment and hydrogen transfer driven by repulsive $1\pi\sigma^*$ states: A new paradigm for nonradiative decay in aromatic molecules. *Phys. Chem. Chem. Phys.*, 4:1093, 2002.

[331] M. N. R. Ashfold, B. Cronin, A. L. Devine, R. N. Dixon, and M. G. D. Nix. The role of $\pi\sigma^*$ states in the heteroaromatic molecules. *Science*, 312:1637, 2006.

[332] A. L. Devine, M. G. D. Nix, R. N. Dixon, and M. N. R. Ashfold. Near-ultraviolet photodissociation of thiophenol. *J. Phys. Chem. A*, 112:9563, 2008.

[333] M. N. R. Ashfold, A. L. Devine, R. N. Dixon, G. A. King, and T. A. A. Oliver. Exploring nuclear motion through conical intersections in the UV photodissociation of phenols and thiophenol. *Proc. Natl. Acad. Sci. U.S.A.*, 105:12701, 2008.

[334] J. S. Lim, H. Choi, I. S. Lim, S. B. Park, Y. S. Lee, and S. K. Kim. Photodissociation dynamics of thiophenol-d_1: the nature of excited electronic states along the S–D bond dissociation coordinate. *J. Phys. Chem. A*, 113:10410, 2009.

[335] L. Zhang, D. G. Truhlar, and S. Sun. Electronic spectrum and characterization of diabatic potential energy surfaces for thiophenol. *Phys. Chem. Chem. Phys.*, 20:28144, 2018.

[336] L. Zhang, D. G. Truhlar, and S. Sun. Full-dimensional three-state potential energy surfaces and state couplings for photodissociation of thiophenol. *J. Chem. Phys.*, 151:154306, 2019.

[337] T. A. Halgren and W. N. Lipscomb. The synchronous-transit method for determining reaction pathways and locating molecular transition states. *Chem. Phys. Lett.*, 49:225–232, 1977.

[338] J. Olsen, B. O. Roos, P. Jörgensen, and H. J. A. Jensen. Determinant-based configuration interaction algorithms for complete and restricted configuration interaction spaces. *J. Chem. Phys.*, 89:2185, 1988.