

## **Distribution Agreement**

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

---

Chenxi Cai

---

Date

Automatic active space selection for the driven similarity  
renormalization group method

By

Chenxi Cai  
Master of Science  
Chemistry

---

Dr. Francesco A. Evangelista, Ph.D  
Advisor

---

Dr. James T. Kindt, Ph.D  
Committee Member

---

Dr. Joel M. Bowman, Ph.D  
Committee Member

Accepted:

---

Lisa A. Tedesco, Ph.D  
Dean of the James T. Laney School of Graduate Studies

---

Date

Automatic active space selection for the driven similarity  
renormalization group method

By

Chenxi Cai  
B.S., Southern University of Science and Technology, 2016

Advisor: Dr. Francesco A. Evangelista, Ph.D

An abstract of  
A thesis submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science  
in Chemistry  
2018

## Abstract

### Automatic active space selection for the driven similarity renormalization group method

By Chenxi Cai

We proposed a convenient automatic active space selection scheme for the driven similarity renormalization group truncated to second order (DSRG-PT2) [C. Li and F. A. Evangelista, *J. Chem. Theory Comput.* 11, 2097 (2015)]. It is based on state-averaged configuration singles (CIS) natural orbitals and the following DSRG computations become a black-box procedure. The scheme is tested for valence excited states calculations with three DSRG methods: valence CI singles (VCIS) and VCIS with doubles (VCISD) wave functions improved with a second-order perturbation theory (VCIS/VCISD-DSRG-PT2) [C. Li and F. A. Evangelista, *J. Chem. Phys.* 147, 074107 (2017)], and state-averaged multireference driven similarity renormalization group [(SA)-MRDSRG] second-order perturbation theory (SA-DSRG-PT2) [C. Li and F. A. Evangelista, *J. Chem. Phys.* 148, 124106 (2018)]. Results are benchmarked on a set of 24 organic molecules. The scheme is also tested for H<sub>2</sub>O core excitations with the complete-active-space (CAS) DSRG method.

Automatic active space selection for the driven similarity  
renormalization group method

By

Chenxi Cai  
B.S., Southern University of Science and Technology, 2016

Advisor: Dr. Francesco A. Evangelista, Ph.D

A thesis submitted to the Faculty of the  
James T. Laney School of Graduate Studies of Emory University  
in partial fulfillment of the requirements for the degree of  
Master of Science  
in Chemistry  
2018

## Acknowledgement

I would like to thank all the people who helped me to finish this thesis.

- My advisor, Dr. Evangelista, helped me a lot during these two years. He taught me knowledge of quantum chemistry and how to write codes. He also gave me many useful suggestions for my career path.
- I would like to thank Dr. James T. Kindt and Dr. Joel M. Bowman for their time as committee.
- I really thank my group members, Jeffrey Schriber, Tianyan Zhang, Dr. Chenyang (York) Li, Dr. Wallace Derricotte, Shuhe Wang and Nan He for all advice they gave. They helped me solve many problems during my graduate career.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Introduction	1
1.2	Density Matrix and Natural Orbitals	3
1.3	Static and Dynamical Correlation	4
1.4	Configuration Interaction Singles (CIS)	5
1.5	Automatic Active Space Selection	7
1.5.1	Generate State Averaged (SA)-CIS One Body Reduced Density Matrix	7
1.5.2	Active Orbitals Selection Based on NO Occupancy	8
1.5.3	Construct New Coefficient Matrix C	10
<b>2</b>	<b>CIS-NO For Valence Excited States</b>	<b>11</b>
2.1	Introduction	11
2.2	Theory	13
2.2.1	VCIS/VCISD-DSRG-PT2	13
2.2.2	SA-DSRG-PT2	15
2.3	Computational Details	18
2.4	Results	18
2.4.1	Threshold Dependence of Active Space	18
2.4.2	Comparison with Manual Selection	22
2.4.3	Comparison with Established Excited State Methods	24
2.4.4	Comparison with CAS Space	26
2.4.5	Conclusion	27
2.4.6	Appendix	28
<b>3</b>	<b>CIS-NO Work On Core Excitation</b>	<b>34</b>
3.1	Introduction	34
3.2	Computational Details	36
3.3	Results	36
3.3.1	H <sub>2</sub> O Core Excitation Energy Calculations with Different Methods	36
3.3.2	H <sub>2</sub> O Core Excitation Energy Calculations with Different Orbitals	38
3.3.3	Conclusion	43
3.3.4	Appendix	46
<b>4</b>	<b>Summary</b>	<b>49</b>

# List of Figures

1.1	A basic description of static correlation and dynamic correlation. Static correlation is the correlation gained via CASSCF and dynamic correlation describes the correlation between the active orbitals and the external orbitals. . . . .	6
1.2	Average Orbital occupancies of formaldehyde when $n$ $A_2$ excited roots are computed. The $A_2$ excited states are generally vertical excitations: $B_1 \rightarrow B_2$ , $B_2 \rightarrow B_1$ or $A_1 \rightarrow A_2$ . The left part shows occupied orbitals in $B_1$ , $B_2$ and $A_1$ . The right part shows virtual orbitals in $B_1$ , $B_2$ and $A_2$ . When the number of root increases, some occupied orbital occupancies become smaller and some virtual orbital occupancies become larger. . . . .	9
2.1	Orbital spaces used to define the VCIS and VCISD zeroth-order wave functions. . . . .	13
2.2	Number of active orbitals selected by CIS-NO with 0.99, 0.95, 0.90,0.85 thresholds for the set of 24 molecules from Ref. 1. Blue bar shows size with 0.99 threshold, orange bar shows size with 0.95, green bar shows size with 0.9 and grey bar shows size with 0.85. All computations use the def2-TZVP basis set. . . . .	19
2.3	Plots of ethylene orbitals. 1a) Canonical orbitals. b)-d) Natural orbitals with 0.99, 0.95, 0.90 thresholds, respectively. The corresponding occupancies are listed below orbitals. The red number means that this orbital is chosen to be active. $1-B_{3u}$ is the HOMO orbital and $1-B_{2g}$ is the LUMO orbital. All computations use the def2-TZVP basis set. . . . .	20
2.4	Active space size of 6 molecules from Ref. 2. Blue bars represent results of CIS-NO selection with 0.9 threshold; Orange bars represent results of manual selection. All computations use def2-TZVP basis set. . . . .	23
3.1	Example of a X-ray photoabsorption spectrum (XAS). The near edge is located in the low energy region and the extended edge is located in the high energy region of the spectrum. . . . .	34
3.2	Transitions resulting from the excitation of a core electron and the corresponding XAS K, L, M edges. . . . .	35
3.3	XAS spectrum of water in the gas-phase taken from Ref. 3 . . . . .	37
3.4	$H_2O$ TDDFT results. All calculations were done using the B3LYP functional and def2-TZVP basis set. Red lines are experiment peaks, black lines are TDDFT roots in 530-550 eV range. Shift value = 15.2 eV. . . . .	38

3.5	H <sub>2</sub> O CASCI results with canonical orbitals. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI roots in 530-550 eV range. CASCI calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift value is -18.4 eV. . . . .	39
3.6	H <sub>2</sub> O CAS-DSRG-PT2 results with canonical orbitals. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI roots in 530-550 eV range. CASCI calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift value is 3.3 eV. . . . .	39
3.7	H <sub>2</sub> O spectra with NO. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI and CAS-DSRG-PT2 roots in 530-550 eV range. CASCI and CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift values are a) -17 eV and b) 2 eV. . . . .	40
3.8	H <sub>2</sub> O a) <i>a</i> <sub>1</sub> and b) <i>b</i> <sub>2</sub> peaks position with NO and Canonical orbitals. For each calculations, the root is only <i>a</i> <sub>1</sub> or <i>b</i> <sub>2</sub> . Red line is experimental value, blue line represents CIS-NO results and black line represents canonical results. CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set. . . . .	42
3.9	H <sub>2</sub> O <i>a</i> <sub>1</sub> and <i>b</i> <sub>2</sub> peaks positions with a)NO and b)Canonical orbitals. c)Relative values between two peaks. For each calculations, the root is only <i>a</i> <sub>1</sub> or <i>b</i> <sub>2</sub> . CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set. . . . .	44
3.10	H <sub>2</sub> O spectrum with NO. The roots are one <i>a</i> <sub>1</sub> and one <i>b</i> <sub>2</sub> , CIS-NO threshold is 0.99. Red lines are experiment peaks, black lines are CAS-DSRG-PT2 results. CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set. Shift value is -6.1 eV. . . . .	45

# List of Tables

2.1	Singlet excitation energies error statistics (in eV) of different thresholds with respect to CC3 values of 24 organic molecules taken from Ref. 4. All DSRG computations use the def2-TZVP basis set. . . . .	21
2.2	Singlet excitation energy errors statistics (in eV) of different thresholds with respect to CC3 values of 24 organic molecules taken from Ref. 4. All SA-DSRG-PT2 computations use the def2-TZVP basis set. . . . .	22
2.3	Singlet excitation energies error statistics (in eV) of VCIS/VCISD and AUTO-VCIS/VCISD-DSRG-PT2 with respect to CC3 of 6 organic molecules taken from Ref. 4. All computations use def2-TZVP basis set and threshold is 90% for auto selection methods . . . . .	24
2.4	Singlet excitation energies error statistics (in eV) of different methods with respect to CC3 of 24 organic molecules taken from Ref. 4. All DSRG computations use def2-TZVP basis set and other methods use the TZVP basis. . . . .	25
2.5	Singlet excitation energies error statistics (in eV) of VCIS/VCISD/SA and CAS with respect to CC3 values of 24 organic molecules taken from Ref. 4. All computations use def2-TZVP basis set and 0.9 threshold. . . . .	27
2.6	Vertical excitation energies (in eV) for singlet states of the twenty-four molecules computed using VCIS-, VCISD-, SA-, CAS-DSRG-PT2 ( $s = 0.5 E_h^{-2}$ ) with the def2-TZVP basis set and different thresholds. . . . .	29
2.7	Number of active orbitals of six molecules selected with CIS-NO and manually. The threshold of CIS-NO is 0.9 . . . . .	32
2.8	Number of Active orbitals of 24 molecules selected by CIS-NO method with different thresholds. . . . .	33
3.1	$a_1$ and $b_2$ excitation energies (in eV) with different number of NO and canonical active virtual orbitals. All CAS-DSRG-PT2 computations use aug-cc-pCVTZ basis set. . . . .	46
3.2	$a_1$ root: active orbitals with different amount of virtual active orbitals selected by CIS-NO method. . . . .	47
3.3	$b_2$ root: active orbitals with different amount of virtual active orbitals selected by CIS-NO method. . . . .	48

# Chapter 1 Introduction

## 1.1 Introduction

Efficient and reliable theoretical methods for describing electronic wavefunctions and potential energy surfaces are usually crucial for understanding chemical reactivity. Single reference methods can provide accurate results for vertical excitations and energies.<sup>5</sup> However, they cannot describe transition state structures, bond breaking, and non-radiative decay via conical intersection. The traditional Hartree-Fock (HF) reference used in standard single-reference methods is poorly suited for describing orbital near degeneracies. For example, the Restricted Hartree-Fock (RHF) method is usually not applicable to transition states.<sup>6</sup> When electron configurations become exactly or nearly degenerate, a multi-configurational/multi-reference treatment may be necessary. In this case, an appropriate zeroth-order wave function is a complete-active-space self-consistent field reference (CASSCF). It includes all possible electron arrangements in the “active space”, a limited set of orbitals, inspired by the Full Optimized Reaction Space (FORS) concept.<sup>7</sup> The CASSCF method<sup>8,9</sup> is the most frequently used approach for treating the dominant static electron correlation effects. Dynamic electron correlation can be accounted for configuration interaction (CI) or multireference perturbation theory.<sup>10-18</sup>

One important advantage of CASSCF is that the wavefunction is invariant with respect to separate orbital rotations within the active, the doubly occupied, and the virtual spaces. However, the variational optimization of orbitals depend strongly and unpredictably on the choice of the space of active orbitals. It can become costly due to the repeated active space full configuration interaction computations. Even with a relatively modest active space, the cost of CASSCF is higher than that of Hartree-

Fock. The nonlinear nature of the orbital optimization can also lead to convergence difficulties. These difficulties can be eased by using natural orbitals (NOs) instead of canonical molecular orbitals and systematically removing configurations from the full interaction expansion. The natural orbital occupation number of a given orbital indicates an orbital's relative contribution to the wave function, and as such, it can be used to identify a compact basis that is well suited for the efficient expansion of the electronic wave function.

The NOs introduced by Löwdin<sup>19,20</sup> are the eigenvectors of the one-particle density matrix. In 1955, Löwdin and Shull<sup>21</sup> showed that for a two-electron system, a CI expansion based on natural orbitals has the best convergence in the sense that it requires the fewest configurations to achieve a given accuracy in the energy. Since then, NOs have been used extensively in CI expansions<sup>22-27</sup> and advantages over canonical Hartree-Fock orbitals were demonstrated.<sup>28-34</sup> In 1976, Shavitt<sup>35</sup> showed that natural orbitals lead to the fastest configuration interaction convergence within different orbital transformations. Jensen et al<sup>28</sup> showed that the magnitude of the natural orbital occupation numbers of second-order Møller-Plesset perturbation theory (MP2) can guide the selection of the active space for CASSCF calculations. The UNO-CAS method<sup>29,30</sup> is an inexpensive alternative to CASSCF, based on NOs of the unrestricted Hartree-Fock (UHF) wave function. Grev and Schaefer<sup>31</sup> investigated the use of NOs obtained from CI wave functions including all single and double excitations (CISD) for use in multireference CI (MRCI) studies. Gordon<sup>32</sup> showed that natural orbital occupation numbers can be used as a diagnostic for the multiconfigurational character of the wave function. Abrams and Sherrill<sup>34</sup> employed single-reference based NOs in complete active space configuration interaction (CASCI) as an alternative to CASSCF orbitals. Barr and Davidson<sup>36</sup> introduced the frozen natural orbitals (FNOs) and many applications were investigated.<sup>37-39</sup> Zhen and Spiridoula<sup>5</sup> proposed the use of NOs obtained from a single reference correlated calculation de-

scribing a high-multiplicity state (HMNOs), as a substitute for CASSCF molecular orbitals. Recently, Levine and co-workers<sup>40</sup> presented a CASCI expansion built from the state-averaged NOs of configuration interaction singles calculations (CIS-NO). The CIS-NO-CASCI approach provided a size intensive/consistent description of excited states. In these CI applications, different procedures of generating NOs have been investigated. Then NOs can be used to choose an active space for subsequent CI calculations and include all the configurations as specified by the correlation method within this active space. In this respect, various schemes based on NOs for automatic active space selection are exploited.<sup>5,34,41–43</sup> It should be mentioned that Neese and co-workers<sup>43</sup> recently introduced an efficient scheme for the automatic selection of an active space for similarity transformed equations of motion (STEOM) coupled cluster method based on Levine’s work.<sup>40</sup>

In this thesis, we assess the accuracy of the automatic selection of active space scheme introduced by Neese<sup>43</sup> for different versions of the driven similarity renormalization with second-order perturbation theory (DSRG-PT2).<sup>44–46</sup> The state-averaged NOs are constructed by single-reference calculations of electronic excited states: an inexpensive CIS-NO. The active space is selected automatically based on CIS-NO and then a DSRG-PT2 computation is performed using those fixed orbitals to recover dynamic effects.

## 1.2 Density Matrix and Natural Orbitals

The one-electron density matrix  $\gamma(x'_1; x_1)$  in the position representation is defined as

$$\gamma(x'_1; x_1) = N \int dx_2 \cdots dx_N \Psi(x'_1, x_2, \dots, x_N)^* \Psi(x_1, x_2, \dots, x_N), \quad (1.1)$$

where  $\Psi$  is a N-electron wavefunction and the quantities  $x_i$  collect the spatial and spin coordinates of electron  $i$ . The density matrix  $\gamma(x'_i; x_i)$  represented in a basis of

spin orbitals  $\phi_i$  may be written as the expectation value

$$\gamma_{pq} = \langle \Psi | \hat{a}_p^\dagger \hat{a}_q | \Psi \rangle, \quad (1.2)$$

where  $\hat{a}_p^\dagger$  and  $\hat{a}_q$  are creation and annihilation operators for orbitals  $\phi_p$  and  $\phi_q$ , respectively. The NOs are eigenstates of the one-electron density matrix,

$$\gamma \mathbf{U} = \mathbf{N} \mathbf{U} \quad (1.3)$$

where the eigenvalues of the density matrix,  $\mathbb{N}$ , are the occupation numbers of the respective NOs. A correlated density matrix has fractional natural occupation numbers and it is not idempotent, which indicates the partial occupation of the associated orbitals. Larger occupation numbers indicate larger contributions to the total correlation energy. As former work has shown,<sup>19</sup> the NOs can be used to represent a correlated wave function in a more compact form.

### 1.3 Static and Dynamical Correlation

In quantum chemistry, the concept of “correlation” is usually taken to encompass all the deficiencies of the Hartree-Fock single-determinantal approach, like, for example, inability of HF to correctly describe the dissociation of molecules. The correlation energy ( $E_{\text{corr}}$ ) is defined as the difference between the exact (non-relativistic) energy ( $E_{\text{exact}}$ ) and the Hartree-Fock energy ( $E_{\text{HF}}$ ).

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}, \quad (1.4)$$

The correlation energy is usually separated into two kinds: static and dynamical correlation.

Static correlation arises from the mixing of near-degenerate configurations (determinants) with the HF Slater determinant. In certain cases, there are degeneracies or near-degeneracies among orbitals and a single Slater determinant does not provide a good description of a many-electron system’s state. In these cases, it is necessary

include all nearly degenerate electron configurations at zeroth-order, which can be done by forming a linear combination of near-degenerate determinants:

$$|\Psi_0\rangle = \sum_i c_i |\Phi_i\rangle, \quad (1.5)$$

where  $|\Psi_0\rangle$  is a  $N$ -electron reference wavefunction,  $|\Phi_i\rangle$  is Slater determinants and  $c_i$  is its corresponding coefficient. The CASSCF involves all possible determinants (full CI) in the active space, which help to find orbitals that minimize the energies of the mixture of near-degenerate determinants.

Dynamical correlation arises from small contributions of excited configurations and describes the correlation between the active orbitals and the external orbitals. Some external orbitals are necessary for quantitative accuracy of matrix diagonalization. However, the process can be expensive if these external orbitals are included in the active space. A solution to this problem will probably require the inclusion of dynamical correlation by perturbation theory (PT),<sup>11</sup> coupled cluster theory (CC),<sup>47</sup> and other approaches. In this thesis, DSRG-PT2 is used to recover dynamical correlation.

Fig. 1.1 shows the static and dynamical correlations differences. Static correlation can be treated with CASSCF and dynamic correlation is typically gained through more approximate means.

## 1.4 Configuration Interaction Singles (CIS)

The configuration interaction expansion of the exact many electron wavefunction equation is typically written as

$$|\Psi\rangle = c_0 |\Phi_0\rangle + \sum_{ia} c_i^a |\Phi_i^a\rangle + \sum_{i<j,a<b} c_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \sum_{i<j<k,a<b<c} c_{ijk}^{abc} |\Phi_{ijk}^{abc}\rangle + \dots, \quad (1.6)$$

where  $|\Phi_0\rangle$  is the HF single Slater determinant,  $|\Phi_i^a\rangle$  is a singly excited determinant formed by replacing spin-orbital  $\phi_i$  in  $|\Phi_0\rangle$  with spin orbital  $\phi_a$ .  $|\Phi_{ij}^{ab}\rangle$  is a doubly

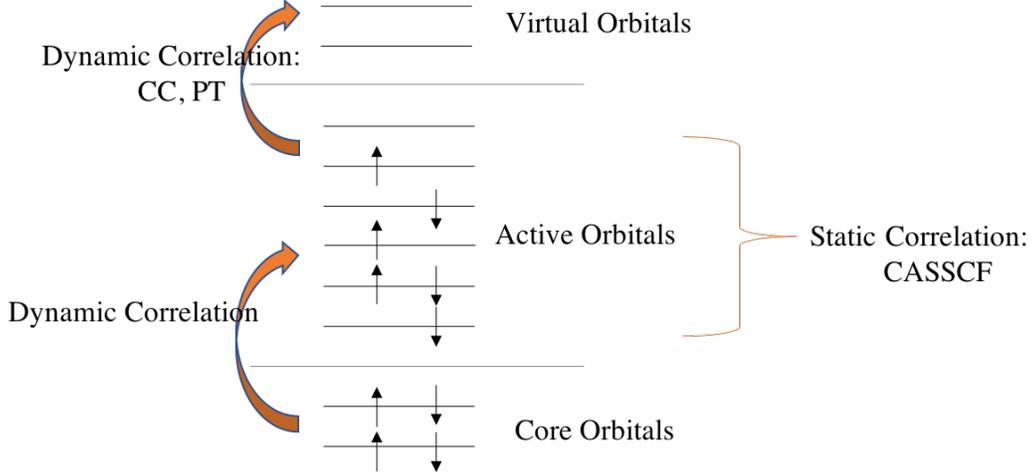


Figure 1.1: A basic description of static correlation and dynamic correlation. Static correlation is the correlation gained via CASSCF and dynamic correlation describes the correlation between the active orbitals and the external orbitals.

excited determinant, etc. In our notations,  $i, j, k$  denote orbitals occupied in the reference and  $a, b, c$  denote orbitals unoccupied (virtual) in the reference. The CIS wavefunction is defined by truncating Eq. 1.6 after single excitations:

$$|\Psi\rangle = c_0|\Phi_0\rangle + \sum_{ia} c_i^a|\Phi_i^a\rangle, \quad (1.7)$$

The CIS energy ( $E_{\text{CIS}}$ ) is given by

$$E_{\text{CIS}} = E_0 + 2 \sum_{ia} c_0 c_i^a F_{ia} + \sum_{iab} c_i^a c_i^b F_{ab} - \sum_{ija} c_i^a c_j^a F_{ij} + \sum_{ijab} c_i^a c_j^b \langle aj || ib \rangle, \quad (1.8)$$

where

$$E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle \quad (1.9)$$

and the Fock matrix element  $F_{pq}$  is defined as

$$F_{pq} = h_{pq} + \sum_{k \in \Phi_0} \langle pk || qk \rangle, \quad (1.10)$$

For a closed-shell SCF reference  $|\Phi_0\rangle$ , off-diagonal terms of the Fock matrix vanish, and the expression for  $E_{\text{CIS}}$  simplifies to

$$E_{\text{CIS}} = E_{\text{SCF}} + \sum_{ia} (c_i^a)^2 (\epsilon_i - \epsilon_j) + \sum_{ijab} c_i^a c_j^b \langle aj || ib \rangle. \quad (1.11)$$

The CI energy and coefficients can be obtained by diagonalizing the CIS Hamiltonian using Davidson's method<sup>48</sup> or the Davidson-Liu Simultaneous Expansion Method.<sup>49</sup>

## 1.5 Automatic Active Space Selection

### 1.5.1 Generate State Averaged (SA)-CIS One Body Reduced Density Matrix

In Levine and co-workers' work,<sup>40</sup> a restricted HF calculation is performed firstly and then a CIS<sup>50</sup> calculation is performed to obtain the natural orbitals. This procedure allows to approximate the lowest  $N$  singlet excited states. By averaging the one body reduced density matrices of low-lying excited (CIS)  $I$ th states ( $\gamma_{pq}^{\text{CIS}(I)}$ ) and ground state ( $\gamma_{\mu\nu}^{\text{HF}}$ ) as shown in the following Equation, the state averaged (SA)-CIS one-body reduced density matrix is

$$D_{pq} = \gamma_{pq}^{\text{SA-CIS}} = \frac{1}{N_\lambda + 1} (\gamma_{\mu\nu}^{\text{HF}} + \sum_{I=1}^{N_\lambda} \gamma_{pq}^{\text{CIS}(I)}), \quad (1.12)$$

where the  $N_\lambda$  is the number of roots of  $I$ th singlet excited states. The CIS one body reduced density matrix is defined

$$\gamma_{pq}^{\text{CIS}(I)} = \langle \Psi_{\text{CIS}(I)} | \hat{a}_p^\dagger \hat{a}_q | \Psi_{\text{CIS}(I)} \rangle, \quad (1.13)$$

where  $|\Psi_{\text{CIS}(I)}\rangle = \sum_i^{\text{occ}} \sum_a^{\text{vir}} C_i^a |\phi_i^a\rangle$ .

The state average CIS density matrix  $D_{pq}$  is block diagonal

$$\mathbf{D} = \begin{pmatrix} \mathbf{D}_o & 0 \\ 0 & \mathbf{D}^v \end{pmatrix} \quad (1.14)$$

where  $\mathbf{D}_o$  and  $\mathbf{D}^v$  are the occupied-occupied and the virtual-virtual blocks and the off-diagonal blocks are zero. Then we separately diagonalize the occupied ( $\mathbf{D}^o$ ) and virtual ( $\mathbf{D}^v$ ) blocks. The corresponding eigenvalues are the occupation numbers  $\Omega$  and unitary matrices  $U$  are obtained as

$$U^{o\dagger} \mathbf{D}^o U^o = \Omega^o, \quad U^{v\dagger} \mathbf{D}^v U^v = \Omega^v, \quad (1.15)$$

The NOs are eigenstates of the one-electron density matrix where  $U^o$  is occupied unitary matrix and  $U^v$  is virtual unitary matrix.

### 1.5.2 Active Orbitals Selection Based on NO Occupancy

Upon diagonalization of the resulting occupied ( $D^o$ ) and virtual ( $D^v$ ) blocks, the occupied and virtual natural orbitals are separately ordered according to their occupation numbers. In order to investigate relationship between occupation numbers and active orbitals, we calculated occupancies of occupied and virtual natural orbitals of formaldehyde, as an example, with increasing number of  $A_2$  excited states. The basis set we used is def2-TZVP.<sup>51</sup> All results are shown in Fig 1.2 with number of  $A_2$  excited states increasing from 0 to 10. The  $C_{2v}$  point group of formaldehyde indicates that for  $A_2$  excited states, the excitations can only be  $B_1 \rightarrow B_2$ ,  $B_2 \rightarrow B_1$ ,  $A_1 \rightarrow A_2$ , or  $A_2 \rightarrow A_1$ . After HF calculations, we knew that all occupied orbitals of formaldehyde are in the  $A_1$ ,  $B_1$  or  $B_2$  irreps. Then the virtual orbitals in  $A_1$  were not considered and shown in Fig 1.2. The figure with zero  $A_2$  excited states illustrates that in group state, occupied orbitals have 1 occupancies while virtual orbitals have zero occupancies. When the number of excited states is set to 1, the  $B_2$  HOMO orbital has occupancy smaller than 1 and the  $B_1$  LUMO orbital has occupancy larger than 0. Then from the following plots with number of roots increasing to 10, we know that the occupied orbitals with smaller occupancies are more likely to loose electrons, and the virtual orbitals with larger occupancies are more likely to receive electrons.

The CIS-NO method with different roots helps us choose active orbitals according to their occupation numbers. Occupied orbitals with smaller occupancies and virtual orbitals with larger occupancies should belong to the active space. To identify the important orbitals, we introduced two cumulative thresholds for the occupied ( $\sigma_o$ ) and virtual spaces ( $\sigma_v$ ) here,

$$\sigma_o \geq \frac{\sum_i^{act} (1 - \Omega_i)}{\sum_j^{No} (1 - \Omega_j)}, \quad \sigma_v \geq \frac{\sum_a^{act} \Omega_a}{\sum_b^{Nv} \Omega_b}, \quad (1.16)$$

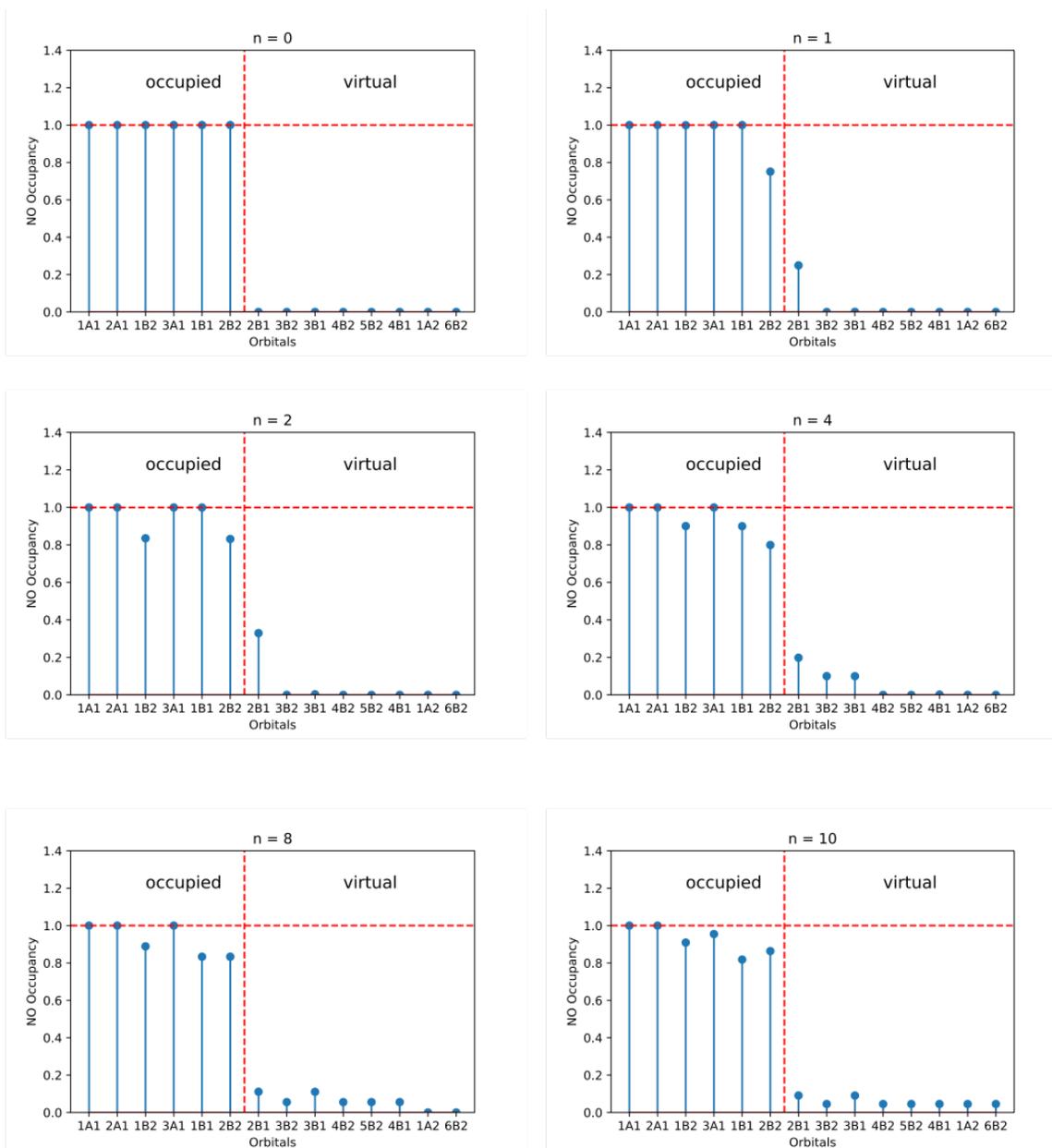


Figure 1.2: Average Orbital occupancies of formaldehyde when  $n$   $A_2$  excited roots are computed. The  $A_2$  excited states are generally vertical excitations:  $B_1 \rightarrow B_2$ ,  $B_2 \rightarrow B_1$  or  $A_1 \rightarrow A_2$ . The left part shows occupied orbitals in  $B_1$ ,  $B_2$  and  $A_1$ . The right part shows virtual orbitals in  $B_1$ ,  $B_2$  and  $A_2$ . When the number of root increases, some occupied orbital occupancies become smaller and some virtual orbital occupancies become larger.

where  $N_o$  and  $N_v$  are the total number of active occupied and active virtual orbitals, respectively.  $\Omega_i$  and  $\Omega_a$  are the occupancies of active occupied and active virtual orbitals, respectively. As these two thresholds indicate, occupied orbitals are chosen

to be the active orbitals in an order of decreasing  $1 - \Omega_i$  until conditions in Equation. (1.16) are satisfied; virtual orbitals are chosen to be the active orbitals in an order of decreasing  $\Omega_a$  until conditions in Equation. (1.16) are satisfied. These cumulative thresholds help select occupied orbitals whose occupancies are small and virtual orbitals whose occupancies are large. They are also more general and convenient than an individual threshold for different molecules. These two thresholds are set to equal in our later work.

### 1.5.3 Construct New Coefficient Matrix $C$

In the last step, we finally transform the MO coefficients to natural orbitals basis with unitary matrix defined as

$$\mathbf{U} = \begin{pmatrix} \mathbf{U}^o & 0 \\ 0 & \mathbf{U}^v \end{pmatrix}, \quad (1.17)$$

in the following way

$$C' = CU. \quad (1.18)$$

The set of CIS-NO orbitals defined by the coefficients  $C'$  are then used in all following computations of the correlation energy.

# Chapter 2 CIS-NO For Valence Excited States

## 2.1 Introduction

As we mentioned in Chapter 1, dynamic electron correlation effects need to be added to achieve highly accurate electron energies. Multireference versions of perturbation theory (MRPT),<sup>11,15,52–60</sup> configuration interaction (MRCI),<sup>1,61–63</sup> coupled cluster theory (MRCC),<sup>47,64–67</sup> and other related approaches<sup>68–70</sup> are generally used to treat dynamic electron correlation. However, these multireference theories are affected by the intruder-state problem,<sup>71,72</sup> which is encountered when determinants that lie within the reference space spanned by the effective Hamiltonian become near-degenerate with determinants that lie outside of it. In perturbative theories, excited configurations that cause the intruder-state problem have small energy denominators. There are some solutions for solving the problem and one is to remove intruders by shifting energy denominators.<sup>73</sup>

Our lab has previously introduced the driven similarity renormalization group (DSRG),<sup>74</sup> a many-body theory inspired by flow renormalization group methods.<sup>75–80</sup> The DSRG performs a continuous unitary transformation of the Hamiltonian controlled by the flow parameter  $s$ :

$$\hat{H} \rightarrow \bar{H}(s) = \hat{U}(s)^\dagger \hat{H} \hat{U}(s), \quad (2.1)$$

where  $\hat{H}$  is the bare Hamiltonian,  $\bar{H}(s)$  is the transformed Hamiltonian, and  $\hat{U}(s)$  is a unitary operator that depends on a time-like parameter  $s$  in range  $\in [0, \infty)$ . When  $s$  goes to infinity, the non-diagonal part  $[\bar{H}(s)]_N$  of  $\bar{H}(s)$  goes to zero. More specifically, for an intermediate given value of  $s$ , states with energy denominators larger than a cutoff energy  $\Lambda = s^{-1/2}$  are folded into an effective Hamiltonian. As

consequence, the excitations that have small energy denominators are avoided by the DSRG transformation.

Li and co-workers have introduced three approaches to electronic excitation energies based on the DSRG: (i) Multireference driven similarity renormalization group (MR-DSRG) derived a second-order perturbation theory (DSRG-MRPT2)<sup>44</sup> (ii) Valence CI singles (VCIS) and VCIS with doubles (VCISD) wave functions improved with a second-order perturbation theory (VCIS/VCISD-DSRG-PT2)<sup>45</sup> (iii) State-averaged multireference driven similarity renormalization group [(SA)-MRDSRG] derived a second-order perturbation theory (SA-DSRG-PT2).<sup>46</sup> Both MR-DSRG and SA-MRDSRG are built by employing the operator algebra of Mukherjee and Kutzelnigg's (MK) generalized normal ordering.<sup>81-83</sup> All these methods avoid the intruder-state problem and may be applied to near-degenerate excited states. They also have other advantages: (1) The DSRG-MRPT2 energy equations are simple and can be solved by a computational approach that requires at most the three-body cumulant of the reference wave function and may be easily combined with density-fitted or Cholesky-decomposed two-electron integrals to reduce the computational cost.<sup>84</sup> (2) VCIS/VCISD-DSRG-PT2 combines an active space truncated CI expansion and second-order multireference perturbation theory and reduce the storage cost for cumulants. (3) In SA-DSRG-PT2 theory, the transformed Hamiltonian can be obtained in a single non-iterative procedure that does not depend on the number of model states.

The present work focus on the implementation of CIS-NO with VCIS/VCISD-DSRG-PT2 and SA-DSRG-PT2 theories. The main advantage of the combination of CIS-NO and different versions of DSRG-PT2 is the automatic selection can lead to an efficient and accurate black-box method for describing electronic wavefunctions of ground and excited states.

## 2.2 Theory

### 2.2.1 VCIS/VCISD-DSRG-PT2

We assume that core (**C**), active hole (**A<sub>H</sub>**, designated by indices  $I, J, \dots$ ), active particle (**A<sub>P</sub>**, designated by the indices  $A, B, \dots$ ), and virtual (**V**) subsets form the full orbital set (**G**). The active orbitals space is defined as the sum of the active hole and partical spaces:  $\mathbf{A} = \mathbf{A}_H + \mathbf{A}_P$ . All these subsets are shown in Fig 2.1.

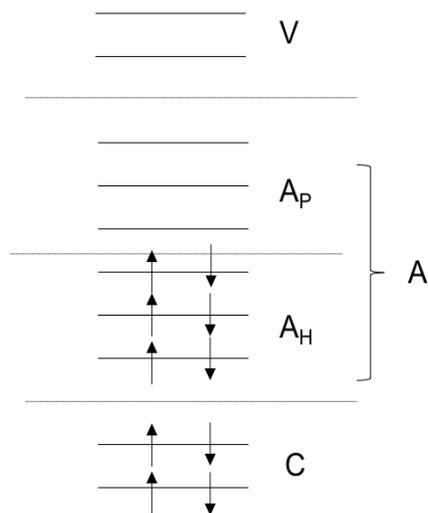


Figure 2.1: Orbital spaces used to define the VCIS and VCISD zeroth-order wave functions.

With this division, we assume that zeroth-order ground state wave functions  $\Psi_0^{(0)}$  for valence excited states can be described by an active space CI with singles (VCIS) or singles and doubles (VCISD) wave functions obtained from a closed-shell reference determinant  $|\Phi_0\rangle$ . The zeroth-order VCIS and VCISD wave functions  $[\Psi_n^{(0)}]$  for the  $n$ th excited state are defined respectively as:

$$|\Psi_n^{(0),S}\rangle = \sum_I^{\mathbf{A}_H} \sum_A^{\mathbf{A}_P} c_I^A |\Phi_I^A\rangle, \quad (2.2)$$

and

$$|\Psi_n^{(0),SD}\rangle = \sum_I^{\mathbf{A}_H} \sum_A^{\mathbf{A}_P} c_I^A |\Phi_I^A\rangle + \frac{1}{4} \sum_{IJ}^{\mathbf{A}_H} \sum_{AB}^{\mathbf{A}_P} c_{IJ}^{AB} |\Phi_{IJ}^{AB}\rangle, \quad (2.3)$$

where  $c_I^A$  and  $c_{IJ}^{AB}$  are the CI coefficients of the singly excited determinant  $|\Phi_I^A\rangle = \hat{a}_A^+ \hat{a}_I |\Phi_0\rangle$ , and doubly excited determinant  $|\Phi_{IJ}^{AB}\rangle = \hat{a}_A^+ \hat{a}_B^+ \hat{a}_J \hat{a}_I |\Phi_0\rangle$ . We improve this zeroth-order description with the driven similarity renormalization group (DSRG-PT2). As Eq.2.1 shows, DSRG performs a continuous unitary transformation of the Hamiltonian. This unitary transformation is controlled by the DSRG flow equation <sup>44,74</sup>

$$[\bar{H}(s)]_N = \hat{R}(s), \quad (2.4)$$

where  $\hat{R}(s)$  is the so-called source operator and it plays the role of driving the DSRG transformation of the Hamiltonian. After solving this flow equation, the electronic energy of a given reference wave function  $\Psi$  (either ground or excited) is given by the expectation value of the DSRG transformed Hamiltonian,

$$E(s) = \langle \Psi | \bar{H}(s) | \Psi \rangle, \quad (2.5)$$

We assume that each electronic state has a different zeroth-order Hamiltonian and the Hamiltonian for the  $n$ th electronic state ( $\hat{H}_n$ ) is partitioned into a zeroth-order component  $[\hat{H}_n^{(0)}]$  plus a first-order perturbation  $[\hat{H}_n^{(1)}]$ . The zeroth-order term  $[\hat{H}_n^{(0)}]$  is defined

$$\hat{H}_n^{(0)} = E_n^{(0)} + \hat{F}_n^{(0)}, \quad (2.6)$$

where  $E_n^{(0)}$  is the reference energy

$$E_n^{(0)} = \langle \Psi_n^{(0)} | \hat{H} | \Psi_n^{(0)} \rangle, \quad (2.7)$$

and  $\hat{F}_n^{(0)}$  is the diagonal apart of the normal-ordered Fock operator

$$\hat{F}_n^{(0)} = \sum_p^{\mathbf{G}} [\mathbf{f}_n^{(0)}]_p^p \{\hat{a}_p^p\}. \quad (2.8)$$

The Fock operator  $[\mathbf{f}_n^{(0)}]$  matrix has the form

$$\mathbf{f}_n^{(0)} = \begin{pmatrix} \mathbf{f}_n^{\mathbf{C}} & 0 & 0 & 0 \\ 0 & \mathbf{f}_n^{\mathbf{A}_H} & 0 & 0 \\ 0 & 0 & \mathbf{f}_n^{\mathbf{A}_P} & 0 \\ 0 & 0 & 0 & \mathbf{f}_n^{\mathbf{V}} \end{pmatrix} \quad (2.9)$$

and it includes the diagonal blocks of the average Fock matrix  $\mathbf{f}_n$  which is defined by the density matrix  $(\gamma_n)$  and the one- and two- electron integrals  $(h_p^q, \langle pr||qs \rangle)$ .

$$(\mathbf{f}_n)_p^q = h_p^q + \sum_m^{\mathbf{C}} \langle pm||qm \rangle + \sum_{\mu\nu}^{\mathbf{A}} \langle p\mu||q\nu \rangle (\gamma_n)_\mu^\nu, \quad (2.10)$$

The DSRG-MRPT2 total energy equation is defined

$$E_n^{[2]}(s) = E_n^{(0)} + \langle \Psi_n^{(0)} | [\tilde{H}_n^{(1)}(s), \hat{T}_n^{(1)}(s)] | \Psi_n^{(0)} \rangle, \quad (2.11)$$

where  $\tilde{H}_n^{(1)}(s)$  is an effective first-order Hamiltonian defined as  $\tilde{H}_n^{(1)}(s) = \hat{H}_n^{(1)} + \hat{R}_n^{(1)}(s)$ . The  $\hat{T}_n^{(1)}(s)$  is the cluster operator that excites electrons from the occupied to the unoccupied orbitals of the Hartree-Fock reference  $(\Psi_0)$ , and is defined by the first-order flow equation

$$\hat{H}_n^{(1)} + [\hat{H}_n^{(0)}, \hat{T}_n^{(1)}(s)]_N = \hat{R}_n^{(1)}(s), \quad (2.12)$$

The excitation energy for the  $n$ th excited state  $[\omega_n(s)]$  is computed by the difference between the DSRG-MRPT2 energy for the ground state  $[E_0^{[2]}(s)]$  and excited state  $[E_n^{[2]}(s)]$ ,

$$\omega_n(s) = E_n^{[2]}(s) - E_0^{[2]}(s). \quad (2.13)$$

### 2.2.2 SA-DSRG-PT2

In SA-MRDSRG, the zeroth-order electronic states  $[\Psi_0^n]$ , which is instead using the active CAS determinant space,

$$|\Psi_0^n\rangle = \sum_{I=1}^{N_{\text{CAS}}} c_I^n |\Psi^I\rangle, \quad (2.14)$$

With different electronic states, they construct a statistical ensemble  $\mathbb{E}_0$ :

$$\mathbb{E}_0 \equiv \{\Psi_0^\alpha, \alpha = 1, 2, \dots, n\}, \quad (2.15)$$

Then by using the theory of generalized normal ordering of Mukherjee and Kutzelnigg (MK-GNO),<sup>82</sup> the bare Hamiltonian  $(\hat{H})$  in a normal-ordered form with respect to

the ensemble is

$$\hat{H} = E_0 + \sum_{pq}^G \bar{f}_p^q \{\hat{a}_q^p\}_\rho + \frac{1}{4} \sum_{pqrs}^G \nu_{pq}^{rs} \{\hat{a}_{rs}^{pq}\}_\rho, \quad (2.16)$$

where  $E_0$  is the SA reference energy and  $\bar{f}_p^q$  is the SA Fock matrix,

$$E_0 = \langle \hat{H} \rangle_\rho = \sum_{\alpha=1}^n \omega_\alpha \langle \Psi_0^\alpha | \hat{H} | \Psi_0^\alpha \rangle, \quad (2.17)$$

$$\bar{f}_p^q = h_p^q + \sum_{ij}^H \nu_{pi}^{qj} \bar{\gamma}_j^i, \quad (2.18)$$

$h_p^q$  is the one-electron integrals,  $\nu_{pi}^{qj} = \langle \phi_p \phi_q | | \phi_r \phi_s \rangle$  is the antisymmetrized two-electron integrals and  $\bar{\gamma}_j^i$  is the SA density matrix.

We introduced state-specific DSRG transformed Hamiltonian in Equation.2.1 Then the state-average transformation of the Hamiltonian is an operator that includes three- and higher-body terms,

$$\bar{H}(s) = \bar{E}_0(s) + \sum_{pq}^G \bar{H}_p^q(s) \{\hat{a}_q^p\}_\rho + \frac{1}{4} \sum_{pqrs}^G \bar{H}_{pq}^{rs}(s) \{\hat{a}_{rs}^{pq}\}_\rho + \dots, \quad (2.19)$$

where  $\bar{E}_0(s) = \langle \bar{H}(s) \rangle_\rho$ ,  $\bar{H}_{pq\dots}^{rs\dots}(s)$  are tensors and the second quantized operators are normal ordered with respect to the ensemble of the states. The state-averaged DSRG applies many-body source operators  $\hat{R}(s)$  normal ordered with respect to the ensemble of states.

$$\hat{R}_k(s) = \frac{1}{(k!)^2} \sum_{ij\dots}^H \sum_{ab\dots}^P r_{ab\dots}^{ij\dots}(s) (\{\hat{a}_{ij\dots}^{ab\dots}\}_\rho + \{\hat{a}_{ab\dots}^{ij\dots}\}_\rho), \quad (2.20)$$

The Equation above shows  $\hat{R}(s)$  with k-body ( $k = 1, 2, \dots, n$ ) components and the  $r_{ab\dots}^{ij\dots}$  is a rank 2k tensor defined below,

$$r_a^i(s) = [f_a^i + \sum_{\mu\nu}^A \Delta_\nu^\mu \bar{\gamma}_\nu^{\mu} t_{a\mu}^{i\nu}(s)] \exp[-s(\Delta_a^i)^2], \quad (2.21)$$

$$r_{ab}^{ij}(s) = \nu_{ab}^{ij} \exp[-s(\Delta_{ab}^{ij})^2], \quad (2.22)$$

The source operator is a nondiagonal operator,  $[\hat{R}(s)]^N = \hat{R}(s)$ . This condition is achieved by applying  $r_{\mu\nu\dots}^{xy\dots}(s) = 0, \forall \mu, \nu, x, y, \dots \in \mathbf{A}$

The DSRG Hamiltonian of SA-DSRG-PT2 is

$$\bar{H}^{[2]}(s) = \bar{H}^{(0)} + \sum_{i=1}^2 \bar{H}^i(s), \quad (2.23)$$

By diagonalizing, we can get the energies. Here we implemented two diagonalization methods, one is contracted and the other is uncontracted. As for contracted scheme, the DSRG Hamiltonian is diagonalized within the space of basis states in the ensemble  $\mathbb{E}_0$ ,

$$\sum_{\beta}^n \langle \Psi_0^{\alpha} | \bar{H}^{[2]}(s) | \Psi_0^{\beta} \rangle C_{\beta}^{\xi} = C_{\alpha}^{\xi} E_{\xi}(s), \xi = 1, \dots, n, \quad (2.24)$$

where  $[E_{\xi}(s)]$  is the energy of state  $\xi$ . The zeroth-order wave function is

$$|\Psi_0^{\xi}\rangle = \sum_{\alpha}^n C_{\alpha}^{\xi} |\Psi_0^{\alpha}\rangle = \sum_{I=1}^{N_{CAS}} \left( \sum_{\alpha}^n C_{\alpha}^{\xi} c_I^{\alpha} \right) |\Phi^I\rangle, \quad (2.25)$$

In the uncontracted scheme, the DSRG Hamiltonian is diagonalized within the space of CAS determinants  $\{\Phi^I, I = 1, 2, \dots, N_{CAS}\}$ ,

$$\sum_J^{N_{CAS}} \langle \Phi_I | \bar{H}^{[2]}(s) | \Phi^J \rangle \tilde{C}_J^{\xi} = \tilde{C}_I^{\xi} \tilde{E}_{\xi}(s), \xi = 1, \dots, n, \quad (2.26)$$

and the wave function is

$$|\tilde{\Psi}_0^{\xi}\rangle = \sum_I^{N_{CAS}} \tilde{C}_I^{\xi} |\Phi^I\rangle, \quad (2.27)$$

The uncontracted scheme is more flexible than the contracted one because each determinant has an corresponding coefficient. These two schemes can avoid the intruder-state problem and both are computational advantageous.

The differences between VCIS/VCISD-DSRG-PT2 and SA-DSRG-PT2 methods are: (1) VCIS/VCISD-DSRG-PT2 are general used for closed-shell computations while SA-DSRG-PT2 is used for open-shell cases. (2) VCIS/VCISD-DSRG-PT2 are specific-state methods while SA-DSRG-PT2 is state-average method, which can solve special case such as conical intersection. Moreover, we have never extensively tested the SA-DSRG-PT2 before. It is important to know the accuracy of it.

## 2.3 Computational Details

The VCIS-, VCISD-, and SA-DSRG-PT2 methods are implemented in Forte, an open-source suite of multireference methods interfaced to the PSI4 quantum chemistry package.<sup>85</sup> We computed 71 singlet states of 24 organic molecules from Thiel’s benchmark set,<sup>4</sup> selecting only molecules with single character %T greater than or equal to 87% in CC3 benchmark, because excited states that are dominated by double excitations have small %T1 values. States with single character %T greater than or equal to 87% are dominated by single excited. Molecular orbitals were optimized at the restricted Hartree-Fock level using the def2-TZVP basis set.<sup>51</sup> Two-electron integrals were approximated by density fitting, where the def2-TZVP-JKFIT<sup>86</sup> basis set was used for Hartree-Fock and DSRG-PT2 computations. The overall calculation procedure is to combine CIS-NO and VCIS/VCISD/SA-DSRG-PT2 (AUTO-VCIS/VCISD/SA-DSRG-PT2). The active space is automatically chosen by CIS-NO for subsequent DSRG computations to evaluate excitation energy. The active orbitals and excitation energies for all single states are provided in Appendix.

## 2.4 Results

### 2.4.1 Threshold Dependence of Active Space

The active space of 24 molecules for their corresponding singlet states were computed with different thresholds: 0.99, 0.95, 0.9 and 0.85, respectively, to investigate threshold effects. Fig. 2.2 shows size of active space of 24 molecules with different thresholds. The number of active orbitals for each molecule are listed in the Appendix. As the threshold decreases from 0.99 to 0.95, the sizes of the active space decrease more than a half for most of the molecules. When threshold decreases to 0.9, most of the active space sizes change little and they are found to be the same as that obtained with the 0.85 threshold.

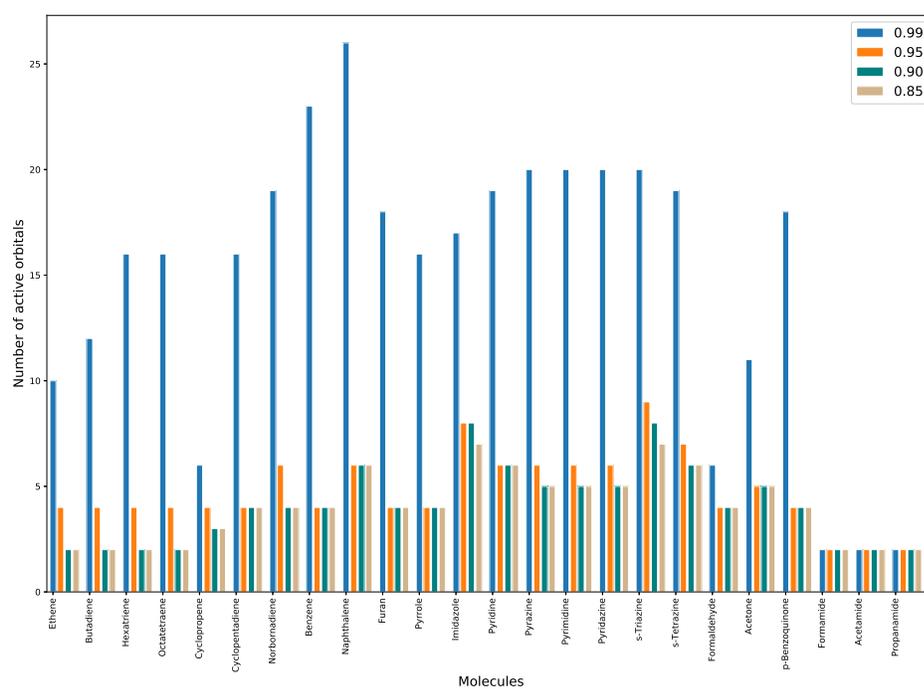


Figure 2.2: Number of active orbitals selected by CIS-NO with 0.99, 0.95, 0.90, 0.85 thresholds for the set of 24 molecules from Ref. 1. Blue bar shows size with 0.99 threshold, orange bar shows size with 0.95, green bar shows size with 0.9 and grey bar shows size with 0.85. All computations use the def2-TZVP basis set.

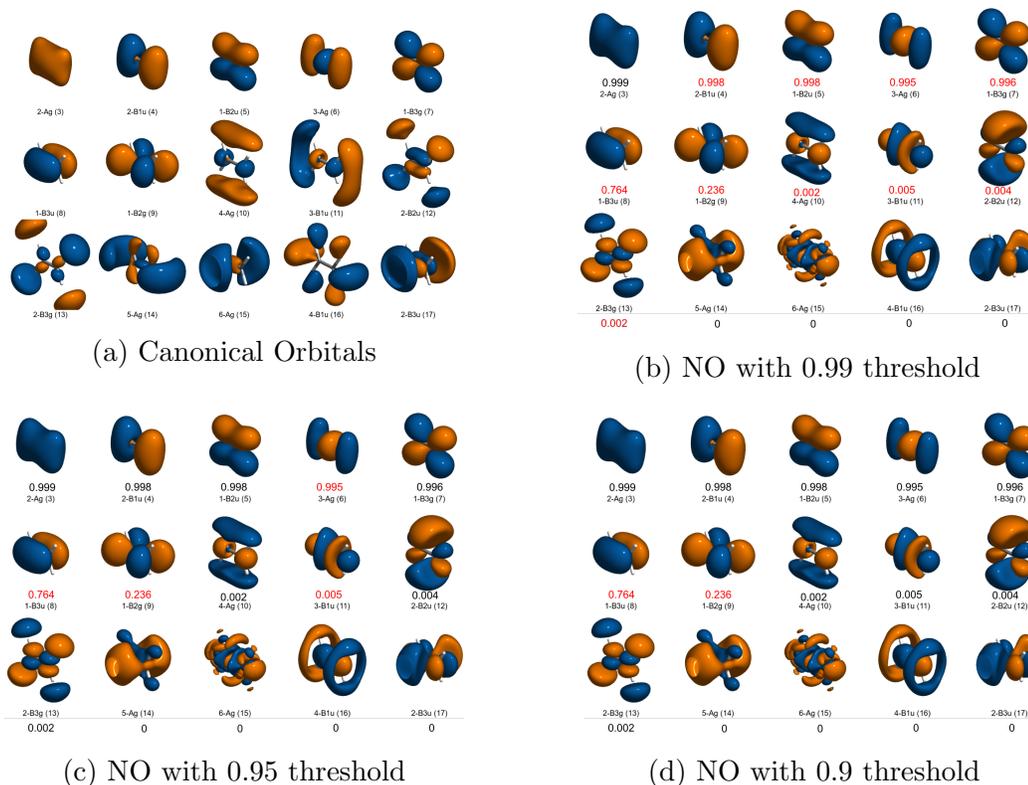


Figure 2.3: Plots of ethylene orbitals. 1a) Canonical orbitals. b)-d) Natural orbitals with 0.99, 0.95, 0.90 thresholds, respectively. The corresponding occupancies are listed below orbitals. The red number means that this orbital is chosen to be active.  $1-B_{3u}$  is the HOMO orbital and  $1-B_{2g}$  is the LUMO orbital. All computations use the def2-TZVP basis set.

For example, Fig. 2.3 shows some important occupied and virtual orbitals of ethylene. The Fig 2.3a shows canonical orbitals and Fig 2.3b, 2.3c, 2.3d show natural orbitals. Orbitals look different from  $4A_g$  orbital because of different types of orbitals. When threshold decreases from 0.99 to 0.95, the number of active orbitals which have red numbers above symmetry is reduced from 10 to 4 and then 2 with 0.90 threshold. As the red number is occupancies of orbitals, Fig 2.3b, 2.3c, 2.3d illustrate that CIS-NO always choose most important orbitals: occupied orbitals with smaller occupancies and virtual orbitals with larger occupancies. When the threshold is smaller, the size of active space also becomes smaller.

Then the excitation energy results with thresholds are investigated. The excitation

Table 2.1: Singlet excitation energies error statistics (in eV) of different thresholds with respect to CC3 values of 24 organic molecules taken from Ref. 4. All DSRG computations use the def2-TZVP basis set.

Parameters	0.99		0.95		0.9	
	VCIS-	VCISD-	VCIS-	VCISD-	VCIS-	VCISD-
ME	-0.05	-0.02	-0.04	-0.08	-0.03	-0.07
SD	0.24	0.28	0.28	0.27	0.28	0.27
MAE	0.20	0.22	0.24	0.24	0.23	0.24
MIN	-0.46	-0.47	-0.52	-0.59	-0.52	-0.59
MAX	0.67	0.92	0.74	0.60	0.74	0.60

energies of 71 singlet states are computed with different threshold and different DSRG methods.

### VCIS/VCISD-DSRG-PT2

The thresholds are set: 0.99, 0.95, and 0.9 for VCIS/VCISD-DSRG-PT2 method and error statistics are with respect to CC3 values. After collecting data in Table 2.1, it can be seen that for AUTO-VCIS-, the mean absolute errors (MAE) difference between 0.99 and 0.95 thresholds is 0.04 eV and that between 0.95 and 0.9 thresholds is -0.01 eV. As for AUTO-VCISD-, the MAE differences are 0.02 eV (between 0.99 and 0.95) and 0 eV (between 0.95 and 0.9). The standard deviation (SD) differences between two thresholds are 0.04 eV (between 0.99 and 0.95) and 0 eV (between 0.95 and 0.9) for AUTO-VCIS-, while -0.01 eV (between 0.99 and 0.95) and 0 eV (between 0.95 and 0.9) for AUTO-VCISD-. The higher threshold is, the more accurate the excited energies. However, all these differences are small, which means the large decrease of active space size leads to slight changes in accuracy. We can truncate active space by decreasing the threshold value and the calculation results still have good accuracy. which improves AUTO-DSRG calculations to be more inexpensive and faster.

## SA-DSRG-PT2

As for SA-DSRG-PT2, the excitation energies of 71 singlet states are computed with different thresholds: 0.95, 0.9 and 0.85, respectively, to investigate threshold effects. Threshold 0.99 is not considered in this part because some of the active spaces are too large to be truncated with CAS. After collecting data in Table 2.2, it can be seen that MAE values are same of three thresholds. The SD differences between two thresholds are 0.01 eV (between 0.95 and 0.90) and 0 eV (between 0.90 and 0.85) for AUTO-SA-DSRG-PT2. From Fig. 2.2 we know that all the molecules have small active space size with thresholds from 0.95 to 0.85. The reason of the little changes of MAE and SD is that as the threshold decreases from 0.95 to 0.85, the amount of active orbitals decrease a little. The active space size keeps unchanged when threshold decreases to 0.85 for most molecules, except imidazole and s-triazine. We still have the conclusion that active space can be truncated by decreasing threshold value and the calculation results still have good accuracy.

Table 2.2: Singlet excitation energy errors statistics (in eV) of different thresholds with respect to CC3 values of 24 organic molecules taken from Ref. 4. All SA-DSRG-PT2 computations use the def2-TZVP basis set.

Parameters	0.95	0.9	0.85
ME	0.13	0.17	0.16
SD	0.24	0.23	0.23
MAE	0.21	0.21	0.21
MIN	-0.26	-0.26	-0.26
MAX	0.75	0.80	0.80

### 2.4.2 Comparison with Manual Selection

We then compared size of truncated active space in AUTO-VCIS/VCISD-DSRG methods with that in Ref. 45, in which active space is selected manually. Because including different singlet states of a molecule in calculation means different set of active space, molecules chosen in comparison need to have same included singlet states.

Ethene, cyclopropene, norbornadiene, imidazole, formaldehyde, and acetone satisfy this requirement. All included singlet states in calculations of these six molecules are the same as that in Ref. 45. Fig. 2.4 shows the number of active orbitals of CIS-NO selection and manual selection. All calculations use the def2-TZVP basis set and CIS-NO threshold is 0.9. The active space data are in the Appendix.

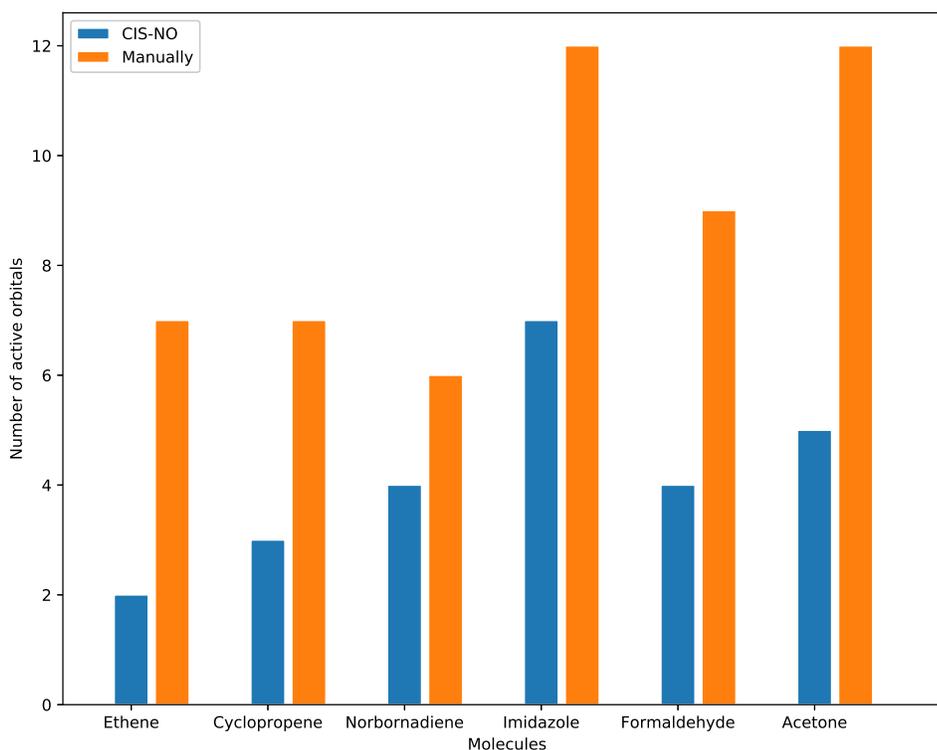


Figure 2.4: Active space size of 6 molecules from Ref. 2. Blue bars represent results of CIS-NO selection with 0.9 threshold; Orange bars represent results of manual selection. All computations use def2-TZVP basis set.

The size of truncated active space with CIS-NO methods is smaller than half of that with manual selection in ethene, cyclopropene, formaldehyde and acetone molecules and larger than a half in norbornadiene and imidazole. Table 2.3 shows MAE and SD of these 6 molecules with respect to CC3 values and both AUTO-VCIS and AUTO-VCISD methods have smaller errors than those of DSRG methods with

manual selection. The size of active space can be truncated well with CIS-NO and active orbitals are more optimal, which means faster, lower-cost and more accurate calculations.

Table 2.3: Singlet excitation energies error statistics (in eV) of VCIS/VCISD and AUTO-VCIS/VCISD-DSRG-PT2 with respect to CC3 of 6 organic molecules taken from Ref. 4. All computations use def2-TZVP basis set and threshold is 90% for auto selection methods

	VCIS-		VCISD-	
	AUTO-	Manually	AUTO-	Manually
SD	0.21	0.24	0.15	0.19
MAE	0.18	0.21	0.14	0.16

### 2.4.3 Comparison with Established Excited State Methods

Here we considering established excited state methods. Table 2.4 shows these 71 states excitation energy error statistics of AUTO-VCIS/VCISD/SA-DSRG-PT2 compared with those of other methods. It includes TDDFT (B3LYP functional),<sup>90–92</sup> DFT combined with multireference CI (DFT/MRCI),<sup>87</sup> single-reference CC methods (CC2 and CCSD),<sup>2</sup> CASPT2 [multi-state (MS)<sup>88</sup> and state-specific (SS)<sup>89</sup>] schemes, and VCIS/VCISD-DSRG-PT2 with manually selected active space.<sup>45</sup> All DSRG computations use def2-TZVP basis set and other methods use the TZVP basis. The CIS-NO threshold was set to 0.99 for VCIS/VCISD-DSRG-PT2 computations and 0.95 for SA-DSRG-PT2 method.

As shown in Table 2.4, both MAE and SD of AUTO-VCIS-DSRG-PT2 (0.20 and 0.24 eV) are smaller than those of previously computed VCIS-DSRG-PT2 (0.23 and 0.28 eV). The MAE of AUTO-VCISD-DSRG-PT2 is the same as VCISD-DSRG-PT2 and the corresponding SD is larger by 0.02 eV. This means that the AUTO-VCIS/VCISD-DSRG-PT2 methods improves upon VCIS/VCISD-DSRG-PT2 with manually selected active space. The SD and MAE differences between VCIS-DSRG-PT2 and VCISD-DSRG-PT2 are 0.02 eV and 0.01 eV, respectively. However, those

Table 2.4: Singlet excitation energies error statistics (in eV) of different methods with respect to CC3 of 24 organic molecules taken from Ref. 4. All DSRG computations use def2-TZVP basis set and other methods use the TZVP basis.

Method	ME	SD	MAE	MIN	MAX
B3LYP <sup>a</sup>	-0.42	0.24	0.43	-0.97	0.08
DFT/MRCI <sup>a</sup>	-0.34	0.17	0.34	-1.12	0.03
CC2 <sup>b</sup>	0.04	0.11	0.09	-0.22	0.27
CCSD <sup>b</sup>	0.19	0.08	0.19	0	0.39
MS-CASPT2 <sup>c</sup>	-0.22	0.17	0.23	-0.62	0.17
SS-CASPT2 <sup>c</sup>	-0.17	0.20	0.21	-0.59	0.41
VCIS-DSRG-PT2 <sup>d</sup>	0.04	0.28	0.23	-0.41	1.02
VCISD-DSRG-PT2 <sup>d</sup>	-0.07	0.26	0.22	-0.58	0.57
VCISD <sub>+HF</sub> -DSRG <sup>d</sup>	-0.06	0.26	0.22	-0.75	0.57
AUTO-VCIS-DSRG-PT2 <sup>e</sup>	-0.05	0.24	0.20	-0.46	0.67
AUTO-VCISD-DSRG-PT2 <sup>e</sup>	-0.02	0.28	0.22	-0.47	0.92
AUTO-SA-DSRG-PT2 <sup>f</sup>	0.13	0.21	0.20	-0.26	0.75

<sup>a</sup> Data taken from Ref. 87

<sup>b</sup> Data taken from Ref. 2

<sup>c</sup> Multi-state (MS) CASPT2 (with IPEA shifts) data taken from Ref. 88 State-specific (SS) CASPT2 data taken from Ref. 89 using the same active spaces and IPEA shifts as Ref. 88

<sup>d</sup> Data taken from Ref. 45. Active space is selected manually.

<sup>e</sup> Threshold is 0.99.

<sup>f</sup> Threshold is 0.95.

differences of AUTO-VCIS-DSRG-PT2 and VCISD-DSRG-PT2 are 0.04 eV and 0.02 eV, respectively. The reason of the slightly larger differences of SD and MAE between AUTO selection than manual selection may be the lack of double excited determinants in CIS-NO part. CIS-NO selection only considers single excited determinants but double excited determinants are necessary for describing VCISD-DSRG-PT2 method. The MAE value order of all these methods are: CC2 < CCSD < AUTO-VCIS-DSRG-PT2  $\sim$  AUTO-SA-DSRG-PT2 < SS-CASPT2 < VCISD-DSRG-PT2  $\sim$  AUTO-VCISD-DSRG-PT2 < VCIS-DSRG-PT2  $\sim$  MS-CASPT2 < DFT/MRCI < B3LYP, which shows that the AUTO-DSRG methods reproduce high-quality CC3 results well. CIS-NO is a convenient selection method with accuracy that can select more optimal active space for excitation energy calculations with VCIS/VCISD/SA-DSRG-PT2 methods.

#### 2.4.4 Comparison with CAS Space

The truncation is also helpful for excitation energies calculations using complete active space (CAS)-based references wave functions improved by DSRG-PT2 (CAS-DSRG-PT2) method. Calculations based on CAS wave function are expensive with large active space because of too many determinants in active space. Here, we can set threshold to 0.9 so that the size of active space selected by CIS-NO is small, which can be seen in Fig. 2.2. Error statistics of AUTO-VCIS/VCISD/SA-DSRG-PT2 and AUTO-CAS-DSRG-PT2 are listed in Table 2.5.

It shows that MAE and SD values of these four methods have slight differences, which means (i) CIS-NO method automatically selects optimal active orbitals for these methods (ii) CIS-NO helps to decrease active space for CAS and the results have good agreement of CC3 reference values (iii) VCIS/VCISD/SA methods work well for singlet excitation energies calculation compared with CAS method. (iv) SA-DSRG-PT2 method works the best for singlet excitation energies calculation compared with

Table 2.5: Singlet excitation energies error statistics (in eV) of VCIS/VCISD/SA and CAS with respect to CC3 values of 24 organic molecules taken from Ref. 4. All computations use def2-TZVP basis set and 0.9 threshold.

	AUTO-SA-	AUTO-VCIS-	AUTO-VCISD-	AUTO-CAS-
ME	0.17	-0.03	-0.07	0.02
SD	0.23	0.28	0.27	0.28
MAE	0.21	0.23	0.24	0.22
MIN	-0.26	-0.52	-0.59	-0.38
MAX	0.80	0.74	0.60	0.87

other methods.

### 2.4.5 Conclusion

We have proposed and implemented a method to automatically select active spaces for multireference computations of excited states based on state average CIS natural orbitals. This method selects important orbitals into active space for excitation energy calculations. The size of active space can be truncated by decreasing the cumulative NO occupancy. The size is truncated significantly when threshold is changed from 0.99 to 0.85 and keep same when threshold is smaller than 0.85. Although the number of active orbitals with 0.9 threshold is smaller than those with 0.99 and 0.95 threshold, the accuracy remains comparable and satisfactory. The three DSRG methods, AUTO-VCIS, AUTO-VCISD and AUTO-SA-DSRG-PT2 have good results compared with established excited state methods such as DFT/MRCI. What’s more, AUTO-SA- is the best one. The truncated active space is also helpful for CAS-based DSRG calculations. The small active space selected helps CAS procedure to be faster and inexpensive. By comparing six selected molecule from Ref. 45, the truncated active space selected by CIS-NO with 90% cumulative occupancy is nearly half the size of manually selected active spaces. In conclusion, the advantages of CIS-NO method are: (i) Active space selection is automatic and results have good accuracy respect with CC3 values. (ii) We can truncate active space with lower threshold,

which help calculations have lower cost. (iii) The CIS-NO method is helpful to make VCIS/VICSD/SA-DSRG-PT2 a black-box methods for excited states.

## 2.4.6 Appendix

### Vertical Excitation Energies

The vertical excitation energies of all 71 singlet transitions from the twenty-four molecules are shown in Table 2.6. We list the excitation energies of CC3 computed with the TZVP basis set and the AUTO-VCIS/VICSD/SA-DSRG-PT2/def2-TZVP results with different thresholds. The CC3 results are taken from Ref. 4

Table 2.6: Vertical excitation energies (in eV) for singlet states of the twenty-four molecules computed using VCIS-, VCISD-, SA-, CAS-DSRG-PT2 ( $s = 0.5 E_h^{-2}$ ) with the def2-TZVP basis set and different thresholds.

Molecule	State	DSRG-PT2										
		CC3	VCIS			VCISD			SA		CAS	
			0.99	0.95	0.9	0.99	0.95	0.9	0.95	0.9	0.85	0.9
Ethene	1 $^1B_{1u}(\pi \rightarrow \pi^*)$	8.37	8.08	8.14	8.11	8.08	8.14	8.11	8.30	8.34	8.34	8.20
E-Butadiene	1 $^1B_u(\pi \rightarrow \pi^*)$	6.58	6.29	6.29	6.28	6.22	6.21	6.28	6.56	6.58	6.58	6.39
all-E-Hexatriene	1 $^1B_u(\pi \rightarrow \pi^*)$	5.58	5.34	5.22	5.26	5.22	5.14	5.26	5.61	5.62	5.62	5.37
all-E-Octatetraene	1 $^1B_u(\pi \rightarrow \pi^*)$	4.94	4.74	4.79	4.85	4.62	4.73	4.85	5.09	5.04	5.04	4.96
Cyclopropene	1 $^1B_1(\sigma \rightarrow \pi^*)$	6.9	6.73	6.71	6.72	6.75	6.67	6.72	6.86	6.92	6.92	6.77
	1 $^1B_2(\pi \rightarrow \pi^*)$	7.1	6.85	7.00	6.99	6.84	7.00	6.99	7.24	7.22	7.22	7.04
Cyclopentadiene	1 $^1B_2(\pi \rightarrow \pi^*)$	5.73	5.42	5.38	5.38	5.46	5.39	5.39	5.70	5.70	5.70	5.51
	3 $^1A_1(\pi \rightarrow \pi^*)$	8.69	8.60	8.73	8.73	8.60	9.06	9.06	8.86	8.86	8.86	8.69
Norbornadiene	1 $^1A_2(\pi \rightarrow \pi^*)$	5.64	5.45	5.43	5.42	5.39	5.42	5.42	5.66	5.70	5.67	5.56
	1 $^1B_2(\pi \rightarrow \pi^*)$	6.49	6.39	6.40	6.40	6.27	6.40	6.40	6.65	6.68	6.65	6.52
	2 $^1B_2(\pi \rightarrow \pi^*)$	7.64	7.65	7.79	7.78	7.62	7.79	7.78	7.95	8.02	7.99	7.86
Benzene	2 $^1A_2(\pi \rightarrow \pi^*)$	7.71	7.65	7.66	7.71	7.59	7.66	7.71	7.87	7.97	7.94	7.82
	1 $^1B_{1u}(\pi \rightarrow \pi^*)$	6.68	6.27	6.23	6.39	6.28	6.23	6.23	6.60	6.60	6.60	6.40
	1 $^1E_{1u}(\pi \rightarrow \pi^*)$	7.45	7.43	7.59	7.73	7.36	7.59	7.59	7.72	7.72	7.72	7.63
Naphthalene	1 $^1B_{2u}(\pi \rightarrow \pi^*)$	5.03	5.11	5.25	5.25	4.83	5.02	5.02	5.35	5.35	5.35	5.29
	2 $^1B_{3u}(\pi \rightarrow \pi^*)$	6.33	6.70	6.98	6.98	6.46	6.93	6.93	7.08	7.08	7.08	7.02
	2 $^1B_{1g}(\pi \rightarrow \pi^*)$	6.79	6.68	6.80	6.80	6.79	7.10	7.10	7.38	7.38	7.38	7.29
	2 $^1B_{2u}(\pi \rightarrow \pi^*)$	6.57	6.86	7.11	7.11	6.61	7.08	7.08	7.22	7.22	7.22	7.19
	3 $^1B_{2u}(\pi \rightarrow \pi^*)$	8.44	9.11	9.18	9.18	8.57	8.86	8.86	9.17	9.17	9.17	9.05
Furan	2 $^1A_1(\pi \rightarrow \pi^*)$	6.62	7.14	7.13	7.13	6.60	6.41	6.41	6.70	6.70	6.70	6.53
	3 $^1A_1(\pi \rightarrow \pi^*)$	8.53	8.61	8.67	8.67	8.51	8.46	8.46	8.65	8.65	8.65	8.42
Pyrrole	1 $^1B_2(\pi \rightarrow \pi^*)$	6.71	6.70	6.68	6.68	6.59	6.62	6.62	6.88	6.88	6.88	6.67
	3 $^1A_1(\pi \rightarrow \pi^*)$	8.17	8.35	8.42	8.42	8.26	8.23	8.23	8.37	8.37	8.37	8.21
Imidazole	2 $^1A''(n \rightarrow \pi^*)$	6.83	7.04	7.03	7.03	7.02	6.74	6.74	6.92	6.95	6.95	6.78
	2 $^1A'(\pi \rightarrow \pi^*)$	6.58	6.89	6.78	6.78	6.63	6.48	6.48	6.64	6.64	6.64	6.53

Table 2.6 (*Continued.*)

Molecule	State	CC3	DSRG-PT2										
			VCIS			VCISD			SA		CAS		
			0.99	0.95	0.9	0.99	0.95	0.9	0.95	0.9	0.85	0.9	
Pyridine	3 $^1A'(\pi \rightarrow \pi^*)$	7.10	7.14	7.12	7.12	6.93	6.97	6.97	7.21	7.21	7.21	7.01	
	4 $^1A''(n \rightarrow \pi^*)$	7.94	8.49	8.45	8.45	8.40	8.14	8.14	8.14	8.20	8.20	8.13	
	5 $^1A'(\pi \rightarrow \pi^*)$	8.45	8.59	8.55	8.55	8.63	8.66	8.66	8.66	8.66	8.66	8.56	
	1 $^1B_1(n \rightarrow \pi^*)$	5.06	5.18	5.02	5.02	5.08	4.84	4.84	4.97	4.97	4.97	4.95	
	2 $^1A_2(n \rightarrow \pi^*)$	5.51	5.50	5.40	5.40	5.59	5.39	5.39	5.50	5.50	5.50	5.45	
	2 $^1A_1(\pi \rightarrow \pi^*)$	6.85	6.46	6.43	6.43	6.44	6.35	6.35	6.72	6.72	6.72	6.52	
Pyrazine	3 $^1A_1(\pi \rightarrow \pi^*)$	7.70	7.68	7.82	7.824	8.15	7.92	7.92	7.95	7.95	7.95	7.85	
	2 $^1B_2(\pi \rightarrow \pi^*)$	7.59	7.66	7.76	7.76	7.53	7.74	7.74	7.88	7.88	7.88	7.79	
	1 $^1B_{3u}(n \rightarrow \pi^*)$	4.25	4.14	4.01	4.16	4.28	4.06	4.02 7	4.32	4.34	4.34	4.25	
	1 $^1A_u(n \rightarrow \pi^*)$	5.05	5.07	4.93	4.87	5.23	4.84	4.93	5.11	5.22	5.22	4.97	
	1 $^1B_{1u}(\pi \rightarrow \pi^*)$	7.07	6.68	6.59	6.77	6.75	6.58	6.59	6.96	7.01	7.01	6.93	
	2 $^1B_{1u}(\pi \rightarrow \pi^*)$	8.06	8.04	8.17	8.39	7.89	8.20	8.17	8.37	8.39	8.39	8.50	
Pyrimidine	2 $^1B_{2u}(\pi \rightarrow \pi^*)$	8.05	8.25	8.29	8.33	8.12	8.21	8.28	8.44	8.47	8.47	8.44	
	1 $^1B_1(n \rightarrow \pi^*)$	4.51	4.52	4.32	4.34	4.65	4.25	4.26	4.41	4.49	4.49	4.36	
	1 $^1A_2(n \rightarrow \pi^*)$	4.93	4.97	4.72	4.71	5.11	4.73	4.82	4.88	5.01	5.01	4.88	
	2 $^1A_1(\pi \rightarrow \pi^*)$	7.06	6.77	6.69	7.83	6.75	6.55	7.86	6.96	7.01	7.01	7.93	
	2 $^1B_2(\pi \rightarrow \pi^*)$	8.01	7.97	8.14	8.13	8.19	8.29	8.17	8.30	8.36	8.36	8.25	
	3 $^1A_1(\pi \rightarrow \pi^*)$	7.74	7.69	7.84	7.83	7.61	7.87	7.86	8.02	8.05	8.05	7.93	
Pyridazine	1 $^1B_1(n \rightarrow \pi^*)$	3.93	3.86	3.71	3.70	4.05	3.56	3.77	3.76	3.90	3.90	3.70	
	1 $^1B_2(\pi \rightarrow \pi^*)$	6.93	6.62	6.54	6.54	6.59	6.423	6.65	6.83	6.87	6.87	6.59	
	2 $^1B_2(\pi \rightarrow \pi^*)$	7.55	7.45	7.70	7.67	7.52	7.63	7.84	7.85	7.88	7.88	7.71	
	3 $^1A_1(\pi \rightarrow \pi^*)$	7.82	7.91	8.08	8.03	7.83	8.00	8.18	8.15	8.18	8.18	8.06	
	<i>s</i> -Triazine	1 $^1A_2''(n \rightarrow \pi^*)$	4.76	4.61	4.62	4.63	4.64	4.39	4.39	4.56	4.61	4.61	4.46
		1 $^1A_1''(n \rightarrow \pi^*)$	4.78	4.34	4.29	4.31	4.75	4.50	4.47	4.60	4.74	4.7	4.49
1 $^1E''(n \rightarrow \pi^*)$		4.82	4.58	4.56	4.55	4.75	4.50	4.48	4.64	4.73	4.73	4.54	
2 $^1A_1'(\pi \rightarrow \pi^*)$		7.41	7.06	7.04	7.04	7.07	6.82	6.82	7.29	7.32	7.32	7.03	
2 $^1E''(n \rightarrow \pi^*)$		7.82	8.10	8.06	7.98	7.68	7.53	7.54	7.67	7.84	7.84	7.60	

Table 2.6 (*Continued.*)

Molecule	State	CC3	DSRG-PT2									
			VCIS			VCISD			SA		CAS	
			0.99	0.95	0.9	0.99	0.95	0.9	0.95	0.9	0.85	0.9
<i>s</i> -Tetrazine	1 $^1E'(\pi \rightarrow \pi^*)$	8.04	7.83	8.04	8.04	8.00	8.22	8.22	8.32	8.37	8.37	8.26
	1 $^1B_{3u}(n \rightarrow \pi^*)$	2.54	2.27	2.09	2.11	2.74	2.16	2.11	2.51	2.63	2.63	2.19
	1 $^1A_u(\pi \rightarrow \pi^*)$	3.80	4.03	3.63	3.63	4.24	3.60	3.63	3.87	4.05	4.05	3.70
	2 $^1A_u(n \rightarrow \pi^*)$	5.46	5.10	5.18	5.20	5.51	5.27	5.20	5.63	5.74	5.74	5.32
	1 $^1B_{1u}(\pi \rightarrow \pi^*)$	7.45	6.99	6.93	6.93	7.12	6.95	6.93	7.40	7.54	7.54	7.09
	2 $^1B_{1u}(\pi \rightarrow \pi^*)$	7.79	7.60	7.81	7.82	7.81	7.86	7.82	8.20	8.20	8.20	7.91
Formaldehyde	2 $^1B_{2u}(\pi \rightarrow \pi^*)$	8.51	8.69	8.78	8.78	8.70	8.80	8.78	9.04	9.12	9.12	8.91
	1 $^1A_2(n \rightarrow \pi^*)$	3.95	3.70	3.70	3.70	4.51	3.81	3.81	4.08	4.08	4.08	3.73
	1 $^1B_1(\sigma \rightarrow \pi^*)$	9.19	9.12	9.04	9.04	9.45	9.12	9.120	9.16	9.16	9.16	9.07
Acetone	2 $^1A_1(\pi \rightarrow \pi^*)$	9.53	9.70	9.74	9.74	9.71	9.65	9.65	10.53	10.53	10.53	9.63
	1 $^1A_2(n \rightarrow \pi^*)$	4.40	4.12	4.15	4.15	5.09	4.25	4.25	4.55	4.55	4.55	4.23
<i>p</i> -Benzoquinone	1 $^1B_1(\sigma \rightarrow \pi^*)$	9.17	9.03	8.95	8.95	9.87	9.02	9.02	9.13	9.13	9.13	8.99
	3 $^1A_1(\pi \rightarrow \pi^*)$	9.65	9.72	9.56	9.56	10.57	9.76	9.76	9.39	9.39	9.39	9.66
	1 $^1B_{3g}(\pi \rightarrow \pi^*)$	4.59	4.72	4.83	4.83	4.12	4.83	4.83	5.32	5.39	5.39	5.09
	1 $^1B_{1u}(\pi \rightarrow \pi^*)$	5.62	5.63	5.76	5.76	5.48	5.76	5.76	6.10	6.14	6.14	5.97
Formamide	2 $^1B_{3g}(\pi \rightarrow \pi^*)$	7.28	7.43	7.48	7.48	7.03	7.48	7.48	7.76	7.81	7.81	7.72
	1 $^1A''(n \rightarrow \pi^*)$	5.66	5.36	5.32	5.32	5.36	5.36	5.36	5.43	5.43	5.43	5.36
Acetamide	1 $^1A''(n \rightarrow \pi^*)$	5.70	5.37	5.37	5.37	5.37	5.37	5.37	5.63	5.63	5.63	5.37
Propanamide	1 $^1A''(n \rightarrow \pi^*)$	5.72	5.38	5.38	5.378	5.38	5.38	5.38	5.64	5.64	5.64	5.38

## Active Space Information

Table 2.7 lists the number of active orbitals of ethene, cyclopropene, norbornadiene, imidazole, formaldehyde and acetone six molecules with CIS-NO selection and manual selection. The number of active orbitals selected by CIS-NO with different thresholds are listed in Table 2.8.

Table 2.7: Number of active orbitals of six molecules selected with CIS-NO and manually. The threshold of CIS-NO is 0.9

Molecule	CIS-NO	Manually
Ethene	2	7
Cyclopropene	3	7
Norbornadiene	4	6
Imidazole	7	12
Formaldehyde	4	9
Acetone	5	12

Table 2.8: Number of Active orbitals of 24 molecules selected by CIS-NO method with different thresholds.

Molecule	0.99	0.95	0.9	0.85
Ethene	10	4	2	2
Butadiene	12	4	2	2
E-Hexatriene	16	4	2	2
E-Octatetraene	16	4	2	2
Cyclopropene	6	4	3	3
Cyclopentadiene	16	4	4	4
Norbornadiene	19	6	4	4
Benzene	23	4	4	4
Naphthalene	26	6	6	6
Furan	18	4	4	4
Pyrrole	16	4	4	4
Imidazole	17	8	8	7
Pyridine	19	6	6	6
Pyrazine	20	6	5	5
Pyrimidine	20	6	5	5
Pyridazine	20	6	5	5
<i>s</i> -Triazine	20	9	8	7
Formaldehyde	6	4	4	4
Acetone	11	5	5	5
<i>p</i> -Benzoquinone	18	4	4	4
Formamide	2	2	2	2
Acetamide	2	2	2	2
Propanamide	2	2	2	2

# Chapter 3 CIS-NO Work On Core Excitation

## 3.1 Introduction

X-ray Absorption Spectroscopy (XAS) is a widely used technique for investigating local geometries and electron structures. There are two regions of XAS: X-ray Absorption Near Edge Structure (XANES), and Extended X-ray Absorption Fine Structure (EXAFS).<sup>93</sup> XANES gives information about the valance, bond angles and energy bandwidth, while EXAFS gives information about the lattice dynamics, interatomic distances, and near neighbor coordination numbers. As shown in Fig. 3.1 taken from Ref. 94, when a core electron absorbs energy less or greater than its binding energy (shown as the black line), there are edges of XAS, near edge and extended edge.

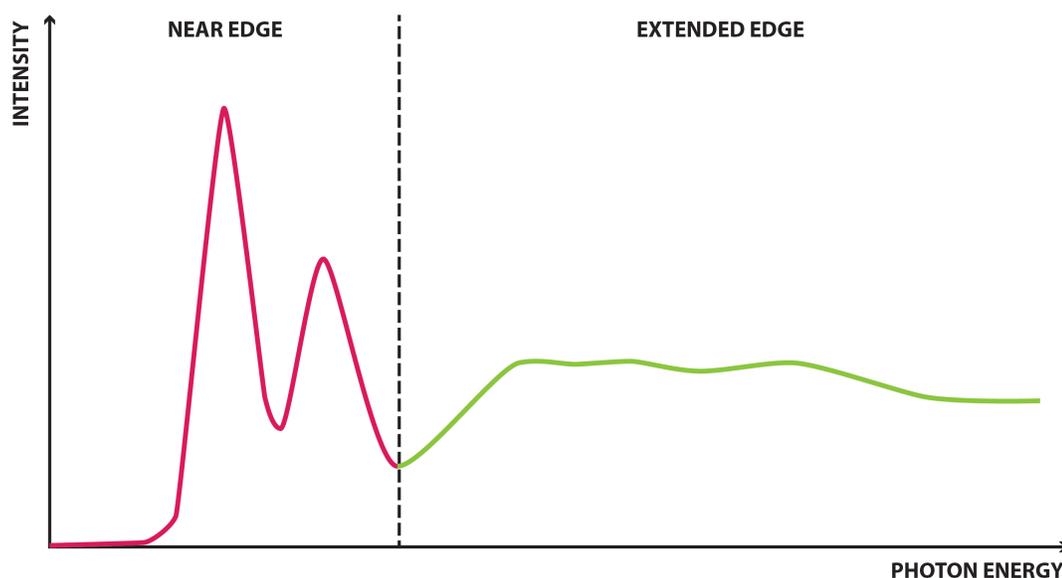


Figure 3.1: Example of a X-ray photoabsorption spectrum (XAS). The near edge is located in the low energy region and the extended edge is located in the high energy region of the spectrum.

The XAS edges are labeled according to the core electron excitation original shell. As shown in Fig. 3.2, labels K,L,M are associated with excitation from the  $n = 1, 2, 3$  atomic shells. For first-row transition metals, metal 1s excitations into partially occupied or empty 3d orbitals from the K-edge, whose analysis is comparatively simple due to low spin-orbit coupling and weak multiconfigurational character.

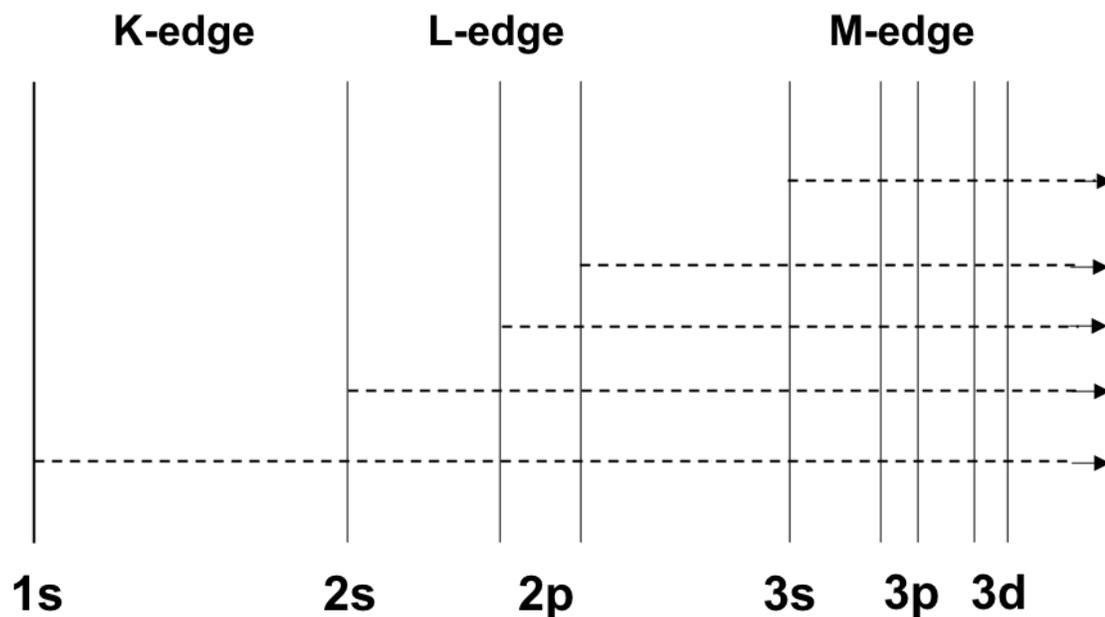


Figure 3.2: Transitions resulting from the excitation of a core electron and the corresponding XAS K, L, M edges.

The traditional method to calculate excitation state is time-dependent density functional theory (TDDFT), which is an extension of DFT. However, core excitation energies from TDDFT are lower than experimental ones due to exchange-correlation functionals. For example, TDDFT yields over 10 eV errors compared with experimental K-edge excitation energies of second-row main-group molecules like CO and N<sub>2</sub>O.<sup>96</sup> Therefore, in order to improve XANES calculation, especially in L-edge region, where states are heavily mixed, multireference methods are needed. Here we consider the DSRG method with second-order multireference perturbation theory (DSRG-MRPT2) for XAS core excitation energy calculations, using CIS-NO to auto-

matically select active orbitals. In this thesis, we tested H<sub>2</sub>O to investigate the effectiveness of active space chosen by CIS-NO used in the multireference DSRG algorithm on the prediction of core-excitations which generate X-ray near-edge absorption.

## 3.2 Computational Details

The DSRG-PT2 method is implemented in Forte, an open-source suite of multireference methods interfaced to the PSI4 quantum chemistry package.<sup>85</sup> We added an option in CIS-NO codes part which can help to choose the amount of active occupied and virtual orbitals. This option can be more convenient for core excitation calculations in the later DSRG part. We computed H<sub>2</sub>O core excitations optimized at the restricted Hartree-Fock level using the correlation-consistent polarized core-valence basis set, aug-cc-pCVTZ basis set.<sup>97</sup> The overall calculation starts with the automatic selection of active space using CIS-NO and it is followed by a DSRG-PT2 computations based on a CASCI reference to evaluate excitation energy.

The TDDFT method excitation energies were computed using B3LYP<sup>98-101</sup> function with def-TZVP basis set.

## 3.3 Results

### 3.3.1 H<sub>2</sub>O Core Excitation Energy Calculations with Different Methods

The experimental K-edge spectrum of H<sub>2</sub>O is shown in Fig. 3.3 taken from Ref. 3. From this figure, the photo energies of 4a<sub>1</sub>, 2b<sub>2</sub> peaks are about 534 and 536 eV, separately. The Rydberg peak is largely composed of a mixture of diffuse Rydberg states within the energy interval of 537 to 538 eV. We mainly focus on 4a<sub>1</sub>, 2b<sub>2</sub> peaks in our later work.

We then performed TDDFT core excitation calculations of H<sub>2</sub>O, from 1s orbital to 50 virtual orbitals. As results shown in Fig. 3.4, the three red lines are experimental

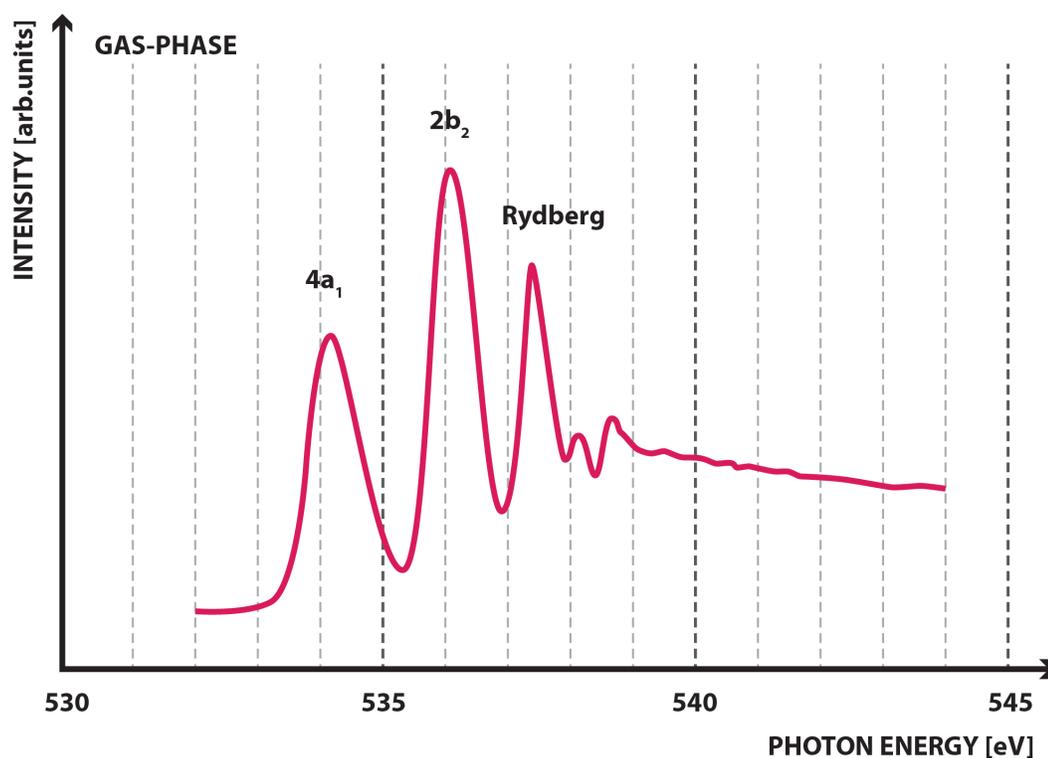


Figure 3.3: XAS spectrum of water in the gas-phase taken from Ref. 3

peaks and black lines are TDDFT results with roots in the 530-550 eV area. The shift value of TDDFT peaks is 15.2 eV. The first two black peaks are overlapped with experimental  $4a_1$ ,  $2b_2$  peaks with similar intensities. However, the Rydberg part with energy interval of 540 to 541.2 eV is far from the experimental value. This figure indicates that TDDFT has converged results of first two excitations, while the Rydberg peak is poorly described.

Next, we computed the K-edge spectrum with various wave function methods. Firstly, we compare results between TDDFT, CASCI, and CAS-DSRF-PT2. After HF computations, we selected the 20 most important virtual orbitals into active space according to their energies. We also set the core excitation root to 20 because only 1s is considered as active occupied orbital and all the core excitations are from 1s to these 20 virtual orbitals. We then did CASCI and CASCI-DSRG-PT2 calculations. It

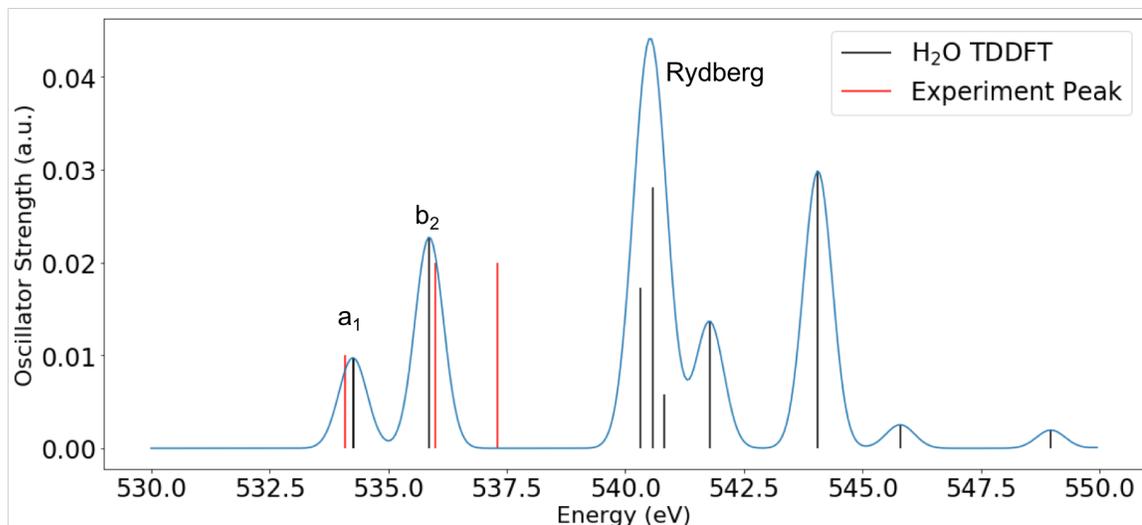


Figure 3.4: H<sub>2</sub>O TDDFT results. All calculations were done using the B3LYP functional and def2-TZVP basis set. Red lines are experiment peaks, black lines are TDDFT roots in 530-550 eV range. Shift value = 15.2 eV.

should be mentioned that, all computations use canonical orbitals. These two results are shown in Fig. 3.5 and Fig. 3.6. By comparing these two figures with TDDFT, both CAS calculations shows the  $4a_1, 2b_2$  peaks closer to experimental ones. The Rydberg peaks in these two methods are also more similar to experimental peak than that of TDDFT. The positions of  $4a_1, 2b_2$  peaks in two CAS methods indicated that multireference methods improve results. Moreover, the CAS-DSRG-PT2 shows the most accurate spectrum, which is nearly identical to the experimental one, and requires a much smaller shift (3.3eV). In conclusion, the DSRG-PT2 is a promising method to compute energies.

### 3.3.2 H<sub>2</sub>O Core Excitation Energy Calculations with Different Orbitals

In this part, we compare H<sub>2</sub>O core excitation energy calculations based on two different orbitals: canonical orbitals and natural orbitals. We already show canonical orbitals results with CASCI and CAS-DSRG-PT2 in the former section. The NO calculations are based on the same basis set, active space, roots and methods, but differ in the use of natural orbitals. Fig. 3.7a shows the K-edge spectrum computed

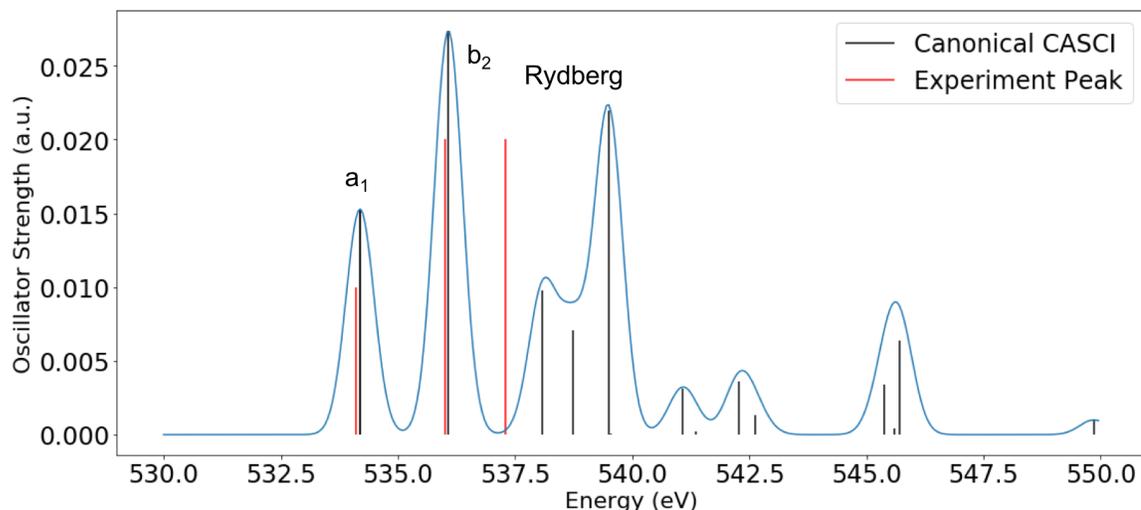


Figure 3.5: H<sub>2</sub>O CASCI results with canonical orbitals. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI roots in 530-550 eV range. CASCI calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift value is -18.4 eV.

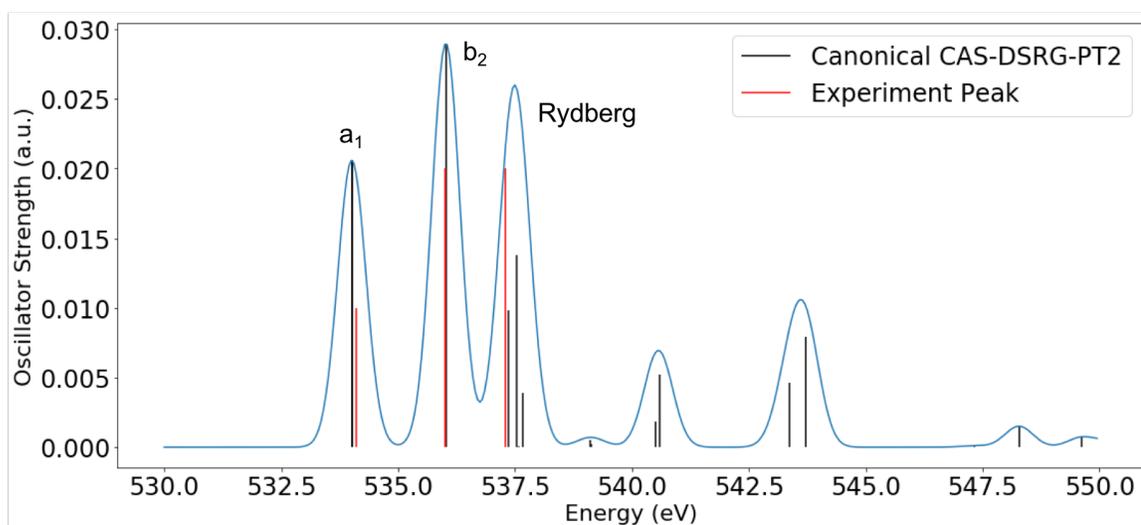


Figure 3.6: H<sub>2</sub>O CAS-DSRG-PT2 results with canonical orbitals. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI roots in 530-550 eV range. CASCI calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift value is 3.3 eV.

with the CASCI method and Fig. 3.7b shows the results of the CAS-DSRG-PT2 computations. The shapes and energies of peaks in these two figures have little overlapped with red experimental peaks. In the figure b of CAS-DSRG-PT2, the first peak is  $a_1$  and the second peak is the mixed result by  $b_2$  and Rydberg. Although the

shift value of CAS-DSRG-PT2 is smaller than that of TDDFT, the messed results with NO are even worse than TDDFT results. We then investigate the reason and improve it.

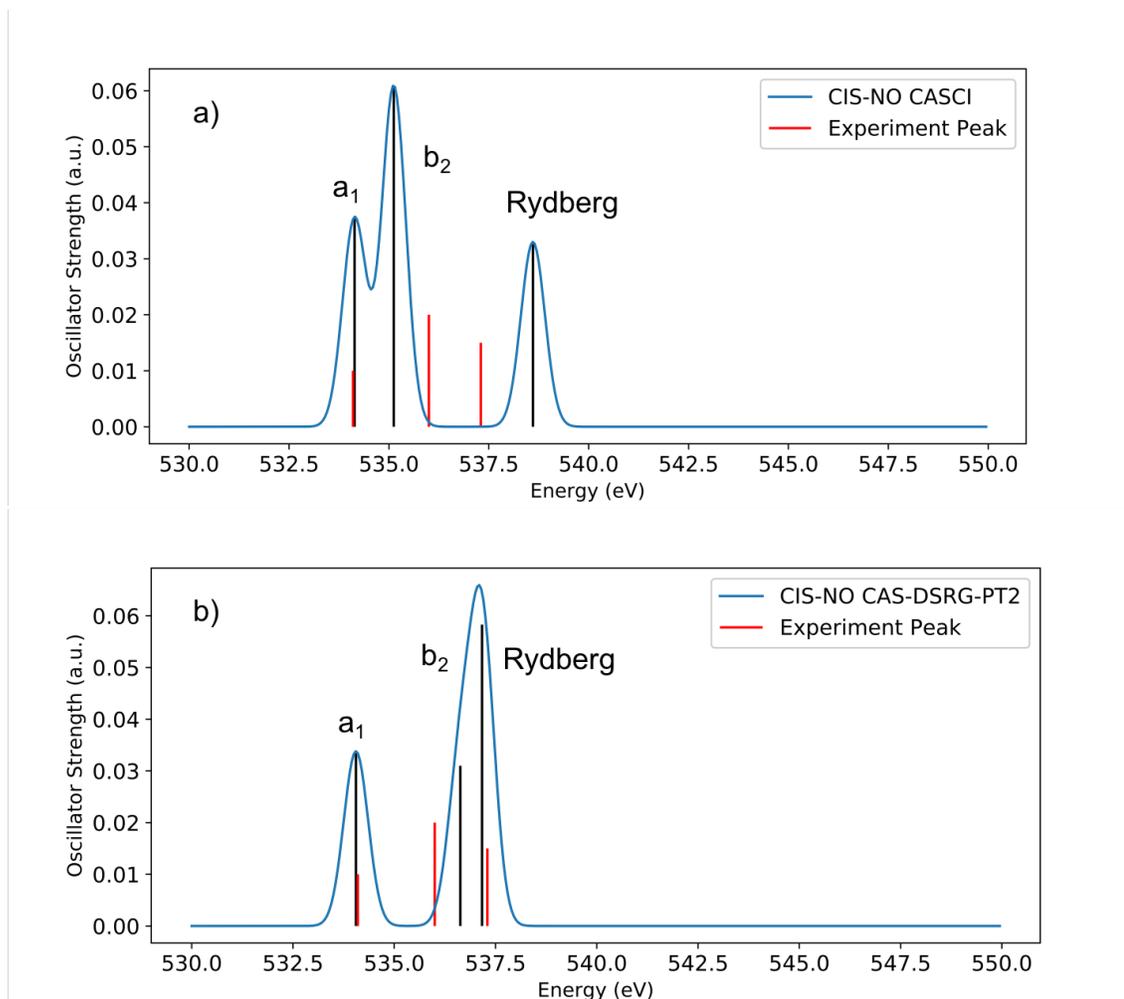


Figure 3.7:  $\text{H}_2\text{O}$  spectra with NO. 20 virtual orbitals are included in active space. Red lines are experiment peaks, black lines are CASCI and CAS-DSRG-PT2 roots in 530-550 eV range. CASCI and CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set and total roots are 20. Shift values are a) -17 eV and b) 2 eV.

To examine the effects of NO compared to canonical orbitals, we only compute one excited state and increase the number of active virtual orbitals by 1 each time and see how the energies change. Here we investigate only one  $a_1$  and one  $b_2$  state, separately. We still use aug-cc-pCVTZ basis set and the procedure is: 1) Select an active space

that has specific number of active virtual orbitals through the CIS-NO procedure. 2) Perform CAS-DSRG-PT2 calculations based on canonical and natural orbitals. All the active orbitals and energies for  $a_1$  and  $b_2$  are provided in the Appendix.

The Fig. 3.8 shows energies of  $a_1$  and  $b_2$  peaks with two types of orbitals, separately. As illustrated in Fig. 3.8a, in the case of the  $a_1$  peak, energies with CIS-NO and canonical orbitals are all smaller than experimental value. The energies with CIS-NO fluctuate in a small range and does not have a tendency: the energies have small changes when the active space is small, while have larger changes when the active space is larger. The energies with canonical orbitals are slowly increasing and show a tendency to be stabilized. The plot is flat when the active space is large and the corresponding energy is most accurate. However, this largest energy with canonical orbitals is still less accurate than the energies with CIS-NO. CIS-NO energies fluctuate little around a more accurate value and are nearly keep unchanged when active space is very small. The same situation happens in the case of the  $b_2$  orbital shown in Fig. 3.8b, which indicates that CIS-NO with much smaller active space has better results than canonical orbitals with large active space.

We also perform vertical comparison between  $a_1$  and  $b_2$  peaks for each kind of orbitals. As shown in Fig. 3.9a, which indicates energies of  $a_1$  and  $b_2$  peaks with CIS-NO, the two plots have nearly the same relative position when the number of active virtual orbitals is small. However, these plots have larger energy difference when active space is larger. The Fig. 3.9b shows those of canonical orbitals. These two plots have similar tendencies to achieve the experimental values and always have similar relative position. The energy differences can be seen more directly in Fig. 3.9c, which shows energy differences between  $a_1$  and  $b_2$  peaks. The experimental value difference, 2 eV, are set as a reference line. The figure c illustrated that canonical orbitals represented two peaks with relative energy difference ranging from 1.3 eV to 4 eV; CIS-NO represented two peaks with relative energy difference ranging from 1.3

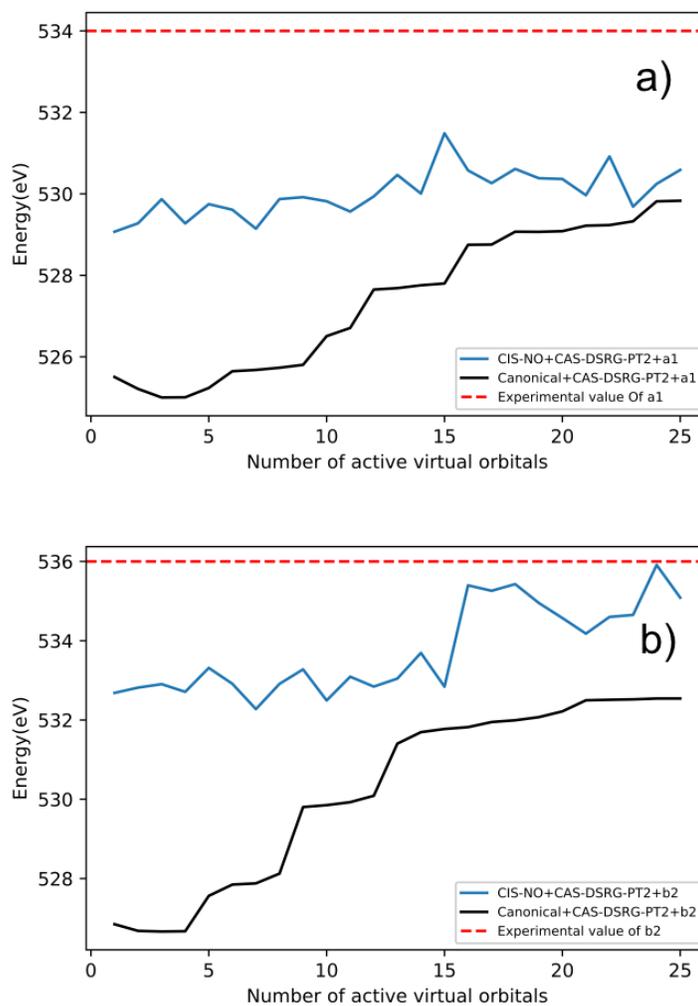


Figure 3.8: H<sub>2</sub>O a) $a_1$  and b) $b_2$  peaks position with NO and Canonical orbitals. For each calculations, the root is only  $a_1$  or  $b_2$ . Red line is experimental value, blue line represents CIS-NO results and black line represents canonical results. CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set.

eV to 5.4 eV. However, CIS-NO represented two peaks with relative energy difference ranging around 3 eV when the number of active virtual orbital is small, which indicates that CIS-NO has more stable results when the active space is small.

From the discussions above, we found that as for simple states, the energies with CIS-NO are more accurate when the active space is small than that of canonical orbitals with large active space. The relative energies of two peaks are also stable with small active space in the case of CIS-NO. The reason may be that the virtual orbitals of H<sub>2</sub>O is far between each other. After CIS-NO, the most important canonical virtual orbitals are transformed and combined to be the first few virtual NOs with large occupation numbers. The far virtual NOs have less possibilities to receive the electrons excited from core orbitals and their occupation numbers are nearly zero. Then when more far virtual NOs are included into the active space, they reduce the accuracy. The Fig 3.10 verifies the conclusion. We only calculated a<sub>1</sub> and b<sub>2</sub> core excitations and set threshold of CIS-NO to 0.99, which then get the smallest active space that including one occupied 1s orbital and two most important virtual orbital in A<sub>1</sub> and B<sub>2</sub>. The results are similar to experimental peaks and the computations are inexpensive because of small active space. Moreover, the shift value is -6.1 eV, which is smaller than that of TDDFT. This conclusion also explains the worse results shown in Fig 3.7. Large active space including less important virtual orbitals will increase errors.

### 3.3.3 Conclusion

We implemented TDDFT, CASCI, and CAS-DSRG-PT2 methods to calculate core excitation energies of H<sub>2</sub>O. Two CAS methods with canonical orbitals represented better than TDDFT, which indicates multireference methods are necessary for core excitation energy calculations. We then investigated influences of different orbitals: NO and canonical orbitals. Results illustrate that, as for simple states 4a<sub>1</sub>

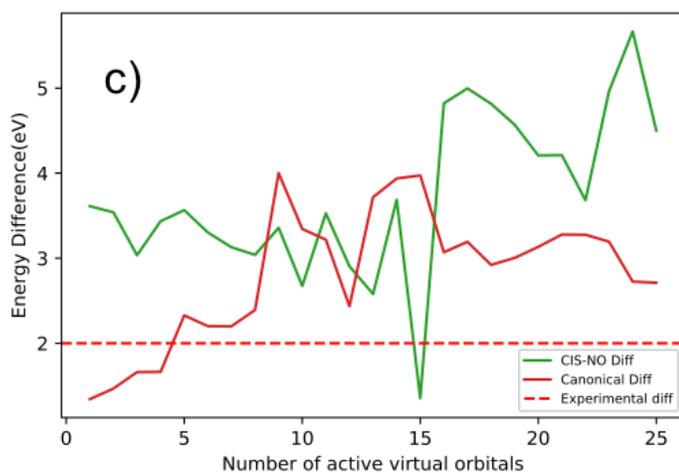
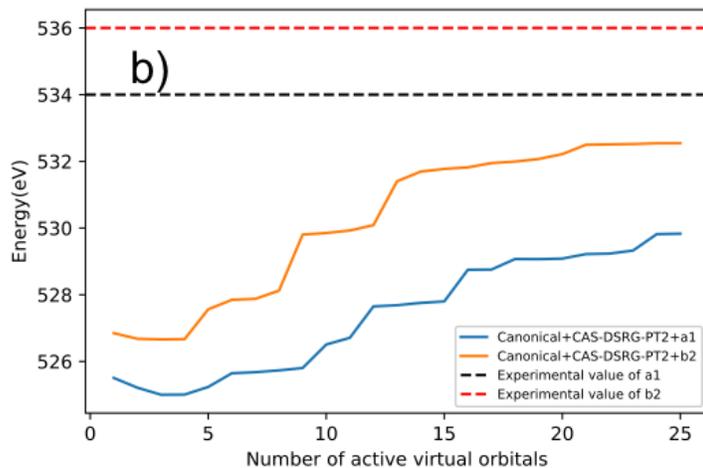
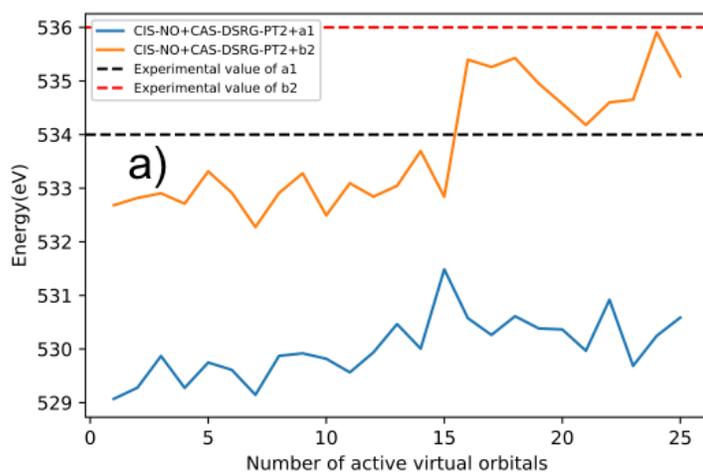


Figure 3.9:  $\text{H}_2\text{O}$   $a_1$  and  $b_2$  peaks positions with a) NO and b) Canonical orbitals. c) Relative values between two peaks. For each calculations, the root is only  $a_1$  or  $b_2$ . CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set.

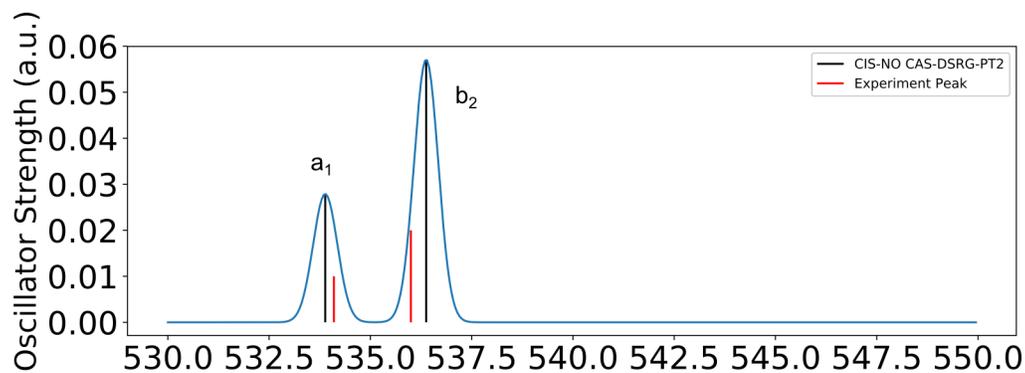


Figure 3.10:  $\text{H}_2\text{O}$  spectrum with  $\text{NO}$ . The roots are one  $a_1$  and one  $b_2$ , CIS-NO threshold is 0.99. Red lines are experiment peaks, black lines are CAS-DSRG-PT2 results. CAS-DSRG-PT2 calculations were done using aug-cc-pCVTZ basis set. Shift value is -6.1 eV.

and  $2b_2$ , CIS-NO represents better results with smaller active space than canonical orbitals. The CIS-NO helps to select most important virtual orbitals into active space and the corresponding active space is small. The computations are cheaper and more convenient. The shift value is also smaller than that of TDDFT. The Rydberg peak with NO needs more work in the future because of its complexity.

### 3.3.4 Appendix

Table 3.1 lists the  $a_1$  and  $b_2$  energy calculations based on NO and canonical orbitals with number of active virtual orbitals changing from 1 to 25. All active orbitals in CAS-DSRG-PT2 calculations are selected by CIS-NO.

Table 3.1:  $a_1$  and  $b_2$  excitation energies (in eV) with different number of NO and canonical active virtual orbitals. All CAS-DSRG-PT2 computations use aug-cc-pCVTZ basis set.

Virtual orbitals	CIS-NO		Canonical Orbitals	
	$a_1$	$b_2$	$a_1$	$b_2$
1	529.070	532.683	525.505	526.847
2	529.278	532.818	525.212	526.679
3	529.869	532.905	525.001	526.663
4	529.275	532.710	525.006	526.670
5	529.748	533.314	525.235	527.563
6	529.611	532.914	525.646	527.847
7	529.145	532.274	525.678	527.878
8	529.872	532.911	525.732	528.124
9	529.920	533.277	525.803	529.803
10	529.819	532.493	526.507	529.851
11	529.565	533.093	526.709	529.926
12	529.938	532.843	527.651	530.088
13	530.465	533.044	527.686	531.403
14	530.001	533.693	527.756	531.693
15	531.486	532.841	527.800	531.773
16	530.576	535.397	528.750	531.82
17	530.263	535.260	528.756	531.949
18	530.612	535.428	529.071	531.993
19	530.384	534.954	529.068	532.071
20	530.365	534.573	529.084	532.218
21	529.967	534.179	529.219	532.498
22	530.919	534.601	529.234	532.510
23	529.686	534.651	529.326	532.520
24	530.246	535.911	529.817	532.542
25	530.588	535.087	529.831	532.543

## Active Space Information

Table 3.2 and Table 3.3 list the active orbitals selection results of one  $a_1$  or  $b_2$  root through CIS-NO selection. The number of active virtual orbitals changes from 1 to 25. All computations us aug-cc-pCVTZ basis set.

Table 3.2:  $a_1$  root: active orbitals with different amount of virtual active orbitals selected by CIS-NO method.

Number of active virtual orbitals	$A_1$	$A_2$	$B_1$	$B_2$
1	2	0	0	0
2	3	0	0	0
3	4	0	0	0
4	5	0	0	0
5	6	0	0	0
6	7	0	0	0
7	8	0	0	0
8	9	0	0	0
9	10	0	0	0
10	11	0	0	0
11	12	0	0	0
12	13	0	0	0
13	14	0	0	0
14	15	0	0	0
15	16	0	0	0
16	17	0	0	0
17	18	0	0	0
18	19	0	0	0
19	20	0	0	0
20	20	0	0	1
21	20	0	0	2
22	21	0	0	2
23	21	0	0	3
24	20	0	0	5
25	21	0	0	5

Table 3.3:  $b_2$  root: active orbitals with different amount of virtual active orbitals selected by CIS-NO method.

Number of active virtual orbitals	$A_1$	$A_2$	$B_1$	$B_2$
1	1	0	0	1
2	1	0	0	2
3	1	0	0	3
4	1	0	0	4
5	1	0	0	5
6	1	0	0	6
7	1	0	0	7
8	1	0	0	8
9	1	0	0	9
10	1	0	0	10
11	1	0	0	11
12	1	0	0	12
13	1	0	0	13
14	1	0	0	14
15	1	0	0	15
16	2	0	0	15
17	3	0	0	15
18	5	0	0	14
19	5	0	0	15
20	6	0	0	15
21	7	0	0	15
22	8	0	0	15
23	9	0	0	15
24	10	0	0	15
25	10	0	0	16

## Chapter 4 Summary

In this thesis, I investigated an active space selection scheme, in valence and core excited states, based on configuration interaction with singles natural orbitals (CIS-NO).

In Chapter 2, CIS-NO is used to automatically select active spaces for VCIS-/VCISD-/SA-DSRG-PT2 computations. The active space changes with cumulative NO occupancies. When the threshold is 0.95, the corresponding active space becomes nearly half of that when the threshold is 0.99. Then active space changes little when the threshold changes to 0.85 or smaller values. The CIS-NO method is helpful to select active space automatically and smaller threshold helps to truncate active space without increasing errors. In addition, we test the Thiel's cases with three DSRG methods: VCIS-, VCISD- and SA-DSRG-PT2. All these three methods reproduce good results compared to CC3 values. They also yield smaller errors than some other methods, such as VCIS/VCISD-DSRG-PT2 with canonical orbitals, whose active space is chosen manually. The SA-DSRG-PT2 gives the best results among these three methods. It becomes a black-box method by using CIS-NO.

In Chapter 3, CIS-NO is used to select active space for H<sub>2</sub>O core excitations. By comparing energies of 4a<sub>1</sub>, 2b<sub>2</sub> and Rydberg states obtained with TDDFT, CASCI, and CAS-DSRG-PT2 with canonical orbitals, the more similar shapes and positions of peaks with two CAS methods show that multireference methods are necessary for core excitation energy calculations. Then we compared results between canonical and natural orbitals. The results of only one specific state, a<sub>1</sub> or b<sub>2</sub>, with increasing number of virtual orbitals computations indicate that CAS calculations with canonical orbitals need large active space to achieve accurate energies. On the contrary, CAS calculations with NOs only need small active space to achieve accurate energies. The

small active space can be automatically chosen by using CIS-NO method. Moreover, the shift value with CIS-NO is also smaller than that of TDDFT. However, Rydbergs generally requires a larger active space. In this case, I find that canonical orbitals are more suitable than NOs. Further investigations are required to assess the deficiency of CIS-NO for core excited states.

CIS-NO is a black-box procedure for active space selection. The occupation number of NOs helps to choose more important orbitals into active space. We can also truncate active space with the smaller CIS-NO threshold. This method helps DSRG calculations become more convenient and inexpensive.

## Bibliography

- [1] Szalay, P. G.; Mller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. *Chem. Rev.* **2011**, *112*, 108–181.
- [2] K nn r, D. n.; Szalay, P. t. G. *J. Chem. Theory Comput.* **2014**, *10*, 3757–3765.
- [3] Nilsson, A.; Nordlund, D.; Waluyo, I.; Huang, N.; Ogasawara, H.; Kaya, S.; Bergmann, U.; slund, L. Å. N.; m, H. Å. s.; Wernet, P.; Andersson, K. J.; Schiros, T.; Pettersson, L. G. M. *J. Electron. Spectrosc. Relat. Phenom.* **2010**, *177*, 99–129.
- [4] Schreiber, M.; Silva-Junior, M. R.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *128*, 134110–26.
- [5] Lu, Z.; Matsika, S. *J. Phys. Chem. A* **2013**, *117*, 7421–7430.
- [6] Schlegel, H. *Adv. Chem. Phys.* **1987**, *67*, 249.
- [7] Ruedenberg, K.; Schmidt, M. W.; Gilbert, M. M.; Elbert, S. T. *Chem. Phys.* **1982**, *71*, 41–49.
- [8] Werner, H.-J. *Adv. Chem. Phys.* *69*.
- [9] Roos, B. O. *Adv. Chem. Phys.* *69*.
- [10] Hirao, K. *Chem. Phys.Lett.* **1992**, *190*, 374 – 380.
- [11] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. *J. Chem. Phys.* **1992**, *96*, 1218–1226.
- [12] Nakano, H. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

- [13] Finley, J.; Malmqvist, P.; Roos, B. O.; Serrano-Andrés, L. *Chem. Phys. Lett.* **1998**, *288*, 299 – 306.
- [14] Celani, P.; Werner, H.-J. *J. Chem. Phys.* **2000**, *112*, 5546–5557.
- [15] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. *J. Chem. Phys.* **2001**, *114*, 10252–10264.
- [16] Khait, Y. G.; Song, J.; Hoffmann, M. R. *J. Chem. Phys.* **2002**, *117*, 4133–4145.
- [17] Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. 2011.
- [18] Roos, B. O. *Acc. Chem. Res.* **1999**, *32*, 137–144.
- [19] Löwdin, P.-O. *Phys. Rev* **1955**, *97*, 1474.
- [20] Davidson, E. R. *Rev. Mod. Phys.* **1972**, *44*, 451.
- [21] Löwdin, P.-O.; Shull, H. *Phys. Rev* **1956**, *101*, 1730.
- [22] Bender, C. F.; Davidson, E. R. *J. Chem. Phys.* **1967**, *47*, 4972–4978.
- [23] Desjardins, S.; Bawagan, A.; Liu, Z.; Tan, K.; Wang, Y.; Davidson, E. *J. Chem. Phys.* **1995**, *102*, 6385–6399.
- [24] Buenker, R. J.; Peyerimhoff, S. D. *Theor. Chem. Acc.* **1974**, *35*, 33–58.
- [25] Thunemann, K.; Römelt, J.; Peyerimhoff, S.; Buenker, R. *Int. J. Quantum Chem.* **1977**, *11*, 743–752.
- [26] Sherrill, C. D.; Schaefer III, H. F. *Adv. Quantum Chem.*; Elsevier, 1999; Vol. 34; pp 143–269.
- [27] Ivanic, J.; Ruedenberg, K. *Theor. Chem. Acc.* **2002**, *107*, 220–228.

- [28] Jensen, H. J. A.; Jo/rgensen, P.; Ågren, H.; Olsen, J. *J. Chem. Phys.* **1988**, *88*, 3834–3839.
- [29] Pulay, P.; Hamilton, T. P. *J. Chem. Phys.* **1988**, *88*, 4926–4933.
- [30] Bofill, J. M.; Pulay, P. *J. Chem. Phys.* **1989**, *90*, 3637–3646.
- [31] Grev, R. S.; Schaefer III, H. F. *J. Chem. Phys.* **1992**, *96*, 6850–6856.
- [32] Gordon, M. S.; Schmidt, M. W.; Chaban, G. M.; Glaesemann, K. R.; Stevens, W. J.; Gonzalez, C. *J. Chem. Phys.* **1999**, *110*, 4199–4207.
- [33] Abrams, M. L.; Sherrill, C. D. *J. Chem. Phys.* **2003**, *118*, 1604–1609.
- [34] Abrams, M. L.; Sherrill, C. D. *Chem. Phys. Lett.* **2004**, *395*, 227–232.
- [35] Shavitt, I.; Rosenberg, B. J.; Palalikit, S. *Int. J. Quantum Chem.* **1976**, *10*, 33–46.
- [36] Barr, T. L.; Davidson, E. R. *Phys. Rev. A* **1970**, *1*, 644.
- [37] Taube, A. G.; Bartlett, R. J. *Collect. Czech. Chem. Commun.* **2005**, *70*, 837–850.
- [38] Aquilante, F.; Todorova, T. K.; Gagliardi, L.; Pedersen, T. B.; Roos, B. O. *J. Chem. Phys.* **2009**, *131*, 034113.
- [39] DePrince III, A. E.; Sherrill, C. D. *J. Chem. Theory Comput.* **2013**, *9*, 2687–2696.
- [40] Shu, Y.; Hohenstein, E. G.; Levine, B. G. *J. Chem. Phys.* **2015**, *142*, 024102–15.
- [41] Jensen, H. J. r. A.; Jo rgensen, P.; Ågren, H.; Olsen, J. *J. Chem. Phys.* **1988**, *88*, 3834–3839.

- [42] Keller, S.; Boguslawski, K.; Janowski, T.; Reiher, M.; Pulay, P. *J. Chem. Phys.* **2015**, *142*, 244104–12.
- [43] Dutta, A. K.; Nooijen, M.; Neese, F.; Izsk, R. b. *J. Chem. Phys.* **2017**, *146*, 074103–12.
- [44] Li, C.; Evangelista, F. A. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.
- [45] Li, C.; Verma, P.; Hannon, K. P.; Evangelista, F. A. *J. Chem. Phys.* **2017**, *147*, 074107.
- [46] Li, C.; Evangelista, F. A. *J. Chem. Phys.* **2018**, *148*, 124106–16.
- [47] Jeziorski, B.; Monkhorst, H. J. *Phys. Rev. A* **1981**, *24*, 1668.
- [48] Davidson, E. R. *J. Comput. Phys.* **1975**, *17*, 87 – 94.
- [49] Liu, B. **1978**,
- [50] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. *J. Phys. Chem.* **1992**, *96*, 135–149.
- [51] Weigend, F.; Ahlrichs, R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 3297–3305.
- [52] Andersson, K.; Malmqvist, P. A.; Roos, B. O.; Sadlej, A. J.; Wolinski, K. *J. Phys. Chem.* **1990**, *94*, 5483–5488.
- [53] Hirao, K. *Chem. Phys. Lett.* **1992**, *190*, 374–380.
- [54] Kozłowski, P. M.; Davidson, E. R. *J. Chem. Phys.* **1994**, *100*, 3672–3682.
- [55] Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *Chem. Phys. Lett.* **2001**, *350*, 297–305.
- [56] Angeli, C.; Cimiraglia, R.; Malrieu, J.-P. *J. Chem. Phys.* **2002**, *117*, 9138–9153.
- [57] Angeli, C.; Pastore, M.; Cimiraglia, R. *Theor. Chem. Acc.* **2007**, *117*, 743–754.

- [58] others., et al. *J. Chem. Phys.* **2005**, *122*, 134105.
- [59] Hoffmann, M. R.; Datta, D.; Das, S.; Mukherjee, D.; Szabados, Á.; Rolik, Z.; Surján, P. R. *J. Chem. Phys.* **2009**, *131*, 204104.
- [60] Evangelista, F. A.; Simmonett, A. C.; Schaefer III, H. F.; Mukherjee, D.; Allen, W. D. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4728–4741.
- [61] Knowles, P. J.; Werner, H.-J. *Chem. Phys. Lett.* **1988**, *145*, 514–522.
- [62] Werner, H.-J.; Knowles, P. J. *J. Chem. Phys.* **1988**, *89*, 5803–5814.
- [63] Shiozaki, T.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2011**, *134*, 034113.
- [64] Mahapatra, U. S.; Datta, B.; Mukherjee, D. *J. Chem. Phys.* **1999**, *110*, 6171–6188.
- [65] Hanrath, M. *J. Chem. Phys.* **2005**, *123*, 084102.
- [66] Datta, D.; Kong, L.; Nooijen, M. *J. Chem. Phys.* **2011**, *134*, 214116.
- [67] Nooijen, M.; Demel, O.; Datta, D.; Kong, L.; Shamasundar, K.; Lotrich, V.; Huntington, L. M.; Neese, F. Communication: Multireference equation of motion coupled cluster: A transform and diagonalize approach to electronic structure. 2014.
- [68] Yanai, T.; Chan, G. K.-L. *J. Chem. Phys.* **2006**, *124*, 194106.
- [69] Neuscammann, E.; Yanai, T.; Chan, G. K.-L. *Int. Rev. Phys. Chem.* **2010**, *29*, 231–271.
- [70] Chen, Z.; Hoffmann, M. R. *J. Chem. Phys.* **2012**, *137*, 014108.
- [71] Evangelisti, S.; Daudey, J.; Malrieu, J. *Phys. Rev. A* **1987**, *35*, 4930.
- [72] Kowalski, K.; Piecuch, P. *Phys. Rev. A* **2000**, *61*, 052506.

- [73] Roos, B. O.; Andersson, K. *Chem. Phys. Lett.* **1995**, *245*, 215–223.
- [74] Evangelista, F. A. *J. Chem. Phys.* **2014**, *141*, 054109.
- [75] Kehrein, S. *The Flow Equation Approach to Many-Particle Systems*; Springer, 2006; pp 137–168.
- [76] Wegner, F. *Ann. Phys.* **1994**, *506*, 77–91.
- [77] Tsukiyama, K.; Bogner, S.; Schwenk, A. *Phys. Rev. Lett.* **2011**, *106*, 222502.
- [78] Tsukiyama, K.; Bogner, S.; Schwenk, A. *Phys. Rev. C* **2012**, *85*, 061304.
- [79] Hergert, H.; Bogner, S.; Morris, T.; Binder, S.; Calci, A.; Langhammer, J.; Roth, R. *Phys. Rev. C* **2014**, *90*, 041302.
- [80] Jurgenson, E.; Navratil, P.; Furnstahl, R. *Phys. Rev. Lett.* **2009**, *103*, 082501.
- [81] Mukherjee, D. *Chem. Phys. Lett.* **1997**, *274*, 561–566.
- [82] Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **1997**, *107*, 432–449.
- [83] Kong, L.; Nooijen, M.; Mukherjee, D. *J. Chem. Phys.* **2010**, *132*, 234107.
- [84] Hannon, K. P.; Li, C.; Evangelista, F. A. *J. Chem. Phys.* **2016**, *144*, 204111.
- [85] others., et al. *Wiley Interdiscip Rev Comput Mol Sci.* **2012**, *2*, 556–565.
- [86] Weigend, F. *J. Comput. Chem.* **2008**, *29*, 167–175.
- [87] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2008**, *129*, 104103–15.
- [88] Silva-Junior, M. R.; Schreiber, M.; Sauer, S. P. A.; Thiel, W. *J. Chem. Phys.* **2010**, *133*, 174318–14.

- [89] Schapiro, I.; Sivalingam, K.; Neese, F. *J. Chem. Theory Comput.* **2013**, *9*, 3567–3580.
- [90] Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098.
- [91] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [92] Xu, X.; Goddard, W. A. *PNAS* **2004**, *101*, 2673–2677.
- [93] Koningsberger, D.; Prins, R. **1988**,
- [94] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.
- [95] Roemelt, M.; Maganas, D.; DeBeer, S.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 204101.
- [96] Derricotte, W. D.; Evangelista, F. A. *Phys. Chem. Chem. Phys.* **2015**, *17*, 14360–14374.
- [97] Woon, D. E.; Dunning Jr, T. H. *J. Chem. Phys.* **1995**, *103*, 4572–4585.
- [98] Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648–5652.
- [99] Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- [100] Vosko, S. H.; Wilk, L.; Nusair, M. *Can. J. Phys.* **1980**, *58*, 1200–1211.
- [101] Stephens, P.; Devlin, F.; Chabalowski, C.; Frisch, M. J. *J. Phys. Chem.* **1994**, *98*, 11623–11627.