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:

Shuhe Wang	Date

Development of Analytic Energy Gradients for the Driven Similarity Renormalization Group Second-Order Perturbation Theory

By

Shuhe Wang
Doctor of Philosophy
Chemistry

Dr. Francesco Evangelista, Ph.D.
Advisor

Dr. Joel Bowman, Ph.D.
Committee Member

Dr. Michael Heaven, Ph.D.
Committee Member

Accepted:

Kimberly Jacob Arriola, Ph.D, MPH
Dean of the James T. Laney School of Graduate Studies

Date

Development of Analytic Energy Gradients for the Driven Similarity Renormalization Group Second-Order Perturbation Theory

By

Shuhe Wang
B.S., University of Science and Technology of China, 2017

Advisor: Dr. Francesco Evangelista, Ph.D.

An abstract of
A dissertation 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
Doctor of Philosophy
in Chemistry
2022

Abstract

Development of Analytic Energy Gradients for the Driven Similarity Renormalization Group Second-Order Perturbation Theory

By Shuhe Wang

We develop the theory and efficient implementations of analytic energy gradients for the driven similarity renormalization group second-order perturbation theory (DSRG-PT2). Three DSRG formalisms are considered: single-reference (SR) DSRG-PT2, multireference (MR) DSRG-PT2, and density-fitted (DF) DSRG-MRPT2. The Z-vector equation approach and the method of Lagrange multipliers are utilized for the theory development. The analytic gradients possess high accuracy and immunity to the intruder state problem. Within DSRG-SRPT2, we demonstrate that the exponentially regularized DSRG parameters do not introduce intractable difficulties in the gradient derivation. The analytic gradients possess an asymptotic scaling same as that of the second-order Møller–Plesset perturbation theory (MP2), and are applied to geometry optimizations for small systems, such as water. For DSRG-MRPT2, we apply the gradients to investigate the singlet and triplet state *p*-benzyne. The resulting equilibrium structures are similar to those computed via other MRPT2s and Mukherjee's multireference coupled cluster theory. We also conduct pilot DSRG-MRPT2 benchmark studies, investigating the adiabatic singlet-triplet gap as a function of the DSRG flow parameter. An approximate DSRG-MRPT2 regime that neglects the three-body density cumulant contributions is further proposed and tested. The high accuracy achieved in associated results indicates that this pruned DSRG formalism could serve as a promising alternative for systems using large active spaces. Within DF-DSRG-MRPT2, we factorize the four-index integrals with density fitting, circumventing the expensive integral backtransformation process. The implemented DF gradients show decent accuracy and provide remarkable computational speedup. Overall, the DF gradient theory leverages the applicability of DSRG-MRPT2 for large strongly correlated systems.

Development of Analytic Energy Gradients for the Driven Similarity Renormalization Group Second-Order Perturbation Theory

By

Shuhe Wang
B.S., University of Science and Technology of China, 2017

Advisor: Dr. Francesco Evangelista, Ph.D.

A dissertation 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
Doctor of Philosophy
in Chemistry
2022

Acknowledgment

First, I would like to thank my undergraduate advisors, Dr. Xin Tao and Dr. Liang Han. I met Dr. Tao in freshman physics class, and I was touched by his humor and novel teaching style. While he intriguingly taught important physics concepts, he also helped me develop the skill of critical thinking. I could not complete my undergraduate thesis without his mentorship, and I missed those exciting days. Dr. Liang Han is an extraordinarily knowledgeable and humble professor. He offered precious advice to me when I was in deep frustration, and I was motivated by him in academic studies and research. They are both eminent scientists and mentors, and I may not start my Ph.D. career without heeding their words.

I would like to thank my advisor, Dr. Francesco Evangelista, for helping me overcome tough challenges and making me a professional researcher. He is a brilliant scientist, teacher, and instructor, with infinite interesting ideas and great enthusiasm. I still remember the day he chose me, a physics-major undergraduate, among other talented chemistry-major applicants. I was not familiar with the quantum chemistry field at the beginning, but he kept teaching me with patience. He introduced quantum chemistry basics, gave weekly tutorials on how to use group packages, and always answered my questions. Those professional research discussions along with the daily conversation from math olympiad problems to the latest news will be cherished. Again, I deeply appreciate the time he invested to change me into a more knowledgeable scientist. I also thank my committee members, Dr. Joel Bowman and Dr. Michael Heaven, for inspirational ideas and helpful discussion of curriculum topics to research proposals. I would like to thank Dr. Yuanzhe Xi from the math department, for being an epic master of numerical analysis, and for all the sincere career and life suggestions he gave me.

I want to show my sincere gratitude towards my colleagues in the Evangelista group. Dr. Chenyang Li taught me how to code in FORTE and provided indispensable help on my gradient project. Dr. Meng Huang, also an alumnus from USTC, offered essential guides to my research implementations and gave me priceless advice on future career choices. Dr. Jonathon Misiewicz, a talented programmer and an outstandingly knowledgeable researcher, always inspired me with elegant ideas. Nan He and Dr.Tianyuan Zhang, my dear colleague, thanks for all their academic help and daily life support with kindness. I am also thankful to all other group members, Dr. Nick Stair, Dr. Ilias Magoulas, Renke Huang, for creating beautiful memories of my Ph.D. studies. And for new students, Shuhang Li and Kevin Marin, thanks for the vitality they brought to the lab.

Finally, I would like to show my greatest gratitude to my parents, Xiaowei Wang and Fuzhen Xu. They always provided wholehearted support to any decisions I made, and they helped me grow as a real man with endless patience and love. I also thank my beloved soulmate, Lulu Mao, for her priceless care and company. I would not make it today without her tolerance and support.

We live in a twilight world, learning from the past and envisioning the future. May the force be with everyone.

- Shuhe Wang

Table of Contents

List of Figures

List of Tables

Chapter 1 Introduction

1.1 Electronic Structure Theory

In quantum chemistry, electronic structure theories are developed so that the electronic motion in atomic or molecular systems could be accurately described, using physics, chemistry, and mathematical techniques. To explain chemical properties and guide experiments, a major goal of electronic structure theories is to be highly predictive. Although various theories could fulfill the requirement of high accuracy by yielding exact solutions, such as the full configuration interaction method,[1] the exponentially growing cost with respect to the system size inherent to those approaches limits their applications to small systems. To address this computational bottleneck, a tradeoff solution is to incorporate a set of manageable approximations into exact or near-exact theories.

We first introduce the nonrelativistic time-independent Schrödinger equation.[2] This corresponds to the eigenvalue problem,

$$\hat{H} \left|\Psi_i\right\rangle = E_i \left|\Psi_i\right\rangle \tag{1.1}$$

where $\hat{H}$ is the Hamiltonian operator, Ψ_i is the electronic wave function and E_i is the energy of the state i. The many-body wave function Ψ must be antisymmetric under exchange of any two electrons, according to the Pauli exclusion principle,

$$\Psi(x_1, \cdots, x_i, \cdots, x_j, \cdots, x_n) = -\Psi(x_1, \cdots, x_j, \cdots, x_i, \cdots, x_n) \tag{1.2}$$

We further apply the well-known Born-Oppenheimer approximation, which assumes that the wave function of electrons and nuclei could be separated because the nuclei are much heavier than electrons.[3] Therefore, the electronic Hamiltonian operator

could be simplified to,

$$\hat{H} = -\frac{1}{2}\sum_{i}^{n}\nabla_i^2 - \sum_{i}^{n}\sum_{I}^{N}\frac{Z_I}{r_{I,i}} + \sum_{i<j}^{n}\frac{1}{r_{i,j}} \tag{1.3}$$

where indices i and I label electrons and nuclei, respectively, ∇ is the Laplace operator, Z is the nuclear charge and r is the distance between two particles. In brief, the first term represents the electronic kinetic energy, the second term describes the electron-nuclei attraction, and the last term is the electron-electron repulsion. The electron-electron repulsion term indicates that a wave function incorporating up to n-body correlations may be necessary for a system of n electrons. Therefore, obtaining the exact solution to the Schrödinger equation [Eq. (1.1)] for real molecular systems is intractable. It is worth mentioning that in second quantization, the Hamiltonian operator $\hat{H}$ is written as,

$$\hat{H} = \sum_{pq} h_p^q \hat{a}_p^\dagger \hat{a}_q + \frac{1}{4}\sum_{pqrs} v_{pq}^{rs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_s \hat{a}_r \tag{1.4}$$

where $h_p^q = \langle \phi_p | \hat{h} | \phi_q \rangle$ represents the one-electron integral, and $v_{pq}^{rs} = \langle \phi_p \phi_q | \phi_r \phi_s \rangle - \langle \phi_p \phi_q | \phi_s \phi_r \rangle$ is the antisymmetrized two-electron counterpart. The second quantization formalism will be discussed later.

1.1.1 One- and Many-Electron Orbital Basis

First, we briefly review the concept of molecular orbitals (MOs). A MO $\phi_p(r)$ is a linear combination of nonorthogonal atomic orbitals (AOs) $\phi_\mu(\mathbf{r})$,

$$\phi_p(\mathbf{r}) = \sum_{\mu}^{\mathrm{AO}} \phi_\mu(\mathbf{r})\, C_{\mu p} \tag{1.5}$$

where $\mathbf{r}$ are the spatial coordinates, and $C_{\mu p}$ are components of a MO coefficient matrix obtained by an orbital-optimization procedure. As a convention, we distinguish the general MO indices from the AO indices using individual labels as p, q, r, s, and Geek letters μ, ν, ρ, τ, respectively.

A molecular spin orbital (MSO) is the product of a MO $\phi(\mathbf{r})$ and a spin function $\sigma(\omega)$, with ω for spin coordinates.

$$\psi_p(\mathbf{r},\omega) = \phi_p(\mathbf{r})\sigma_p(\omega) \tag{1.6}$$

Unlike AOs, orthonormality is assumed for MSOs, yielding an identity MSO overlap matrix S,

$$S_p^q = \langle\psi_p|\psi_q\rangle = \delta_p^q \tag{1.7}$$

where δ is the Kronecker delta.

A many-body basis for an N-electron system is the set of all possible Slater determinants,

$$\Phi(x_1, x_2, \cdots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(x_1) & \psi_2(x_1) & \cdots & \psi_N(x_1) \\ \psi_1(x_2) & \psi_2(x_2) & \cdots & \psi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(x_N) & \psi_2(x_N) & \cdots & \psi_N(x_N) \end{vmatrix} \tag{1.8}$$

The coefficient $(N!)^{-1/2}$ is a normalization factor considering all possible particle permutations. The rows and columns of a Slater determinant are labeled by electrons and spin orbitals, respectively. Interchanging coordinates x of two electrons corresponds to swapping two rows of the determinant and results in a change of sign, which satisfies the antisymmetry principle. For convenience, a short-hand notation with only diagonal components is often used,

$$\Phi(x_1, x_2, \cdots, x_N) = |\psi_1, \psi_2, \cdots, \psi_N\rangle \tag{1.9}$$

In full configuration interaction (FCI), the wave function Ψ_{FCI} is expressed as a linear combination of multiple Slater determinants Φ_I with associated CI coefficients c_I,

$$\Psi_{\text{FCI}} = \sum_I c_I \Phi_I \tag{1.10}$$

This approach considers all possible configurations within the space spanned by a specific orbital basis, and provides exact solutions to the Schrödinger equation.[1,4,5] Unfortunately, FCI is only applicable to the smallest systems due to its high computational cost.[6]

1.1.2 Second Quantization

Second quantization is a formalism to represent the many-body configurations in a more succinct way.[2,7,8] Herein, we briefly review the essential terminology of second quantization for fermions.

The Fermi vacuum state, noted as $|-\rangle$, is introduced as a state containing no particles. We then define a creation operator $\hat{a}_\mu^\dagger$ and an annihilation operator $\hat{a}_\mu$, which respectively creates or annihilates a particle in the spin orbital $|\phi_\mu\rangle$. Applying these operators to a single particle state would modify the occupation status in such a way,

$$\hat{a}_\mu^\dagger |-\rangle = |\phi_\mu\rangle \tag{1.11}$$

$$\hat{a}_\mu |\phi_\mu\rangle = |-\rangle \tag{1.12}$$

$$\hat{a}_\mu^\dagger |\phi_\mu\rangle = 0 \tag{1.13}$$

$$\hat{a}_\mu |-\rangle = 0 \tag{1.14}$$

The last two equations indicate that creating fermions in an occupied state or annihilating a particle from a vacuum state is physically prohibited. Furthermore, we enforce that the pairwise permutations of operators introduce changes of sign of the resulting determinant,

$$\hat{a}_\mu^\dagger \hat{a}_\nu^\dagger |-\rangle = |\phi_\mu \phi_\nu\rangle = -|\phi_\nu \phi_\mu\rangle = -\hat{a}_\nu^\dagger \hat{a}_\mu^\dagger |-\rangle \tag{1.15}$$

Therefore, two creation operators or two annihilation operators satisfy the following *anticommutation* relation,

$$\{\hat{a}_\mu^\dagger, \hat{a}_\nu^\dagger\} = \hat{a}_\mu^\dagger \hat{a}_\nu^\dagger + \hat{a}_\nu^\dagger \hat{a}_\mu^\dagger = 0 \tag{1.16}$$

$$\{\hat{a}_\mu, \hat{a}_\nu\} = 0 \tag{1.17}$$

It could be verified that a creation operator and an annihilation operator obey the anticommutation relation, consistent with the antisymmetric character of the wave

function,

$$\{\hat{a}_\mu, \hat{a}_\nu^\dagger\} = \delta_{\mu\nu} \tag{1.18}$$

where δ is the Kronecker delta, that only equals 1 when μ equals ν and equals 0 for all other cases.

In terms of these second-quantized operators, a Slater determinant may be written as a set of creation operators applied to a vacuum state,

$$\hat{a}_\mu^\dagger \hat{a}_\nu^\dagger \cdots \hat{a}_\sigma^\dagger \, |-\rangle = |\phi_\mu \phi_\nu \cdots \phi_\sigma\rangle \tag{1.19}$$

1.1.3 Multireference Methods

Although the Hartree-Fock (HF) method is of low cost and easy to implement,[9–11] it might not be an appropriate choice if high accuracy is entailed. The HF theory intrinsically omits explicit electron-electron interactions, also named *electron correlation*, making the electronic description significantly deviate from the actual systems where the mean-field approximation breaks down.[12,13] A plethora of post-HF theories were proposed to improve the HF performance by recovering correlation energy contributions ΔE,

$$\Delta E = E^{\text{true}} - E^{\text{HF}} \tag{1.20}$$

Some prominent examples of such wave-function-based theories are Møller-Plesset perturbation theory (MP),[14–17] configuration interaction theory (CI),[4,18,19] and coupled-cluster theory (CC).[20–22] Typically, these post-HF approaches yield more accurate results than HF, though at the cost of much higher computational complexity.

However, these methods assume a single reference configuration, and provide a qualitatively wrong description for systems in which strong correlation exists.[23,24] Examples of strongly-correlated systems are bond dissociation, diradicals, transition metal complexes, and electronically excited states.[13,25–27] An example of such failure could be ascribed to the orbital degeneracies or quasi-degeneracies among the highest

occupied molecular orbitals (HOMO) and the lowest unoccupied molecular orbitals (LUMO).

A prominent multireference approach is the multi-configuration self-consistent field (MCSCF) method.[28–30] A MCSCF wave function Ψ_0 is a linear combination of Slater determinants (Φ_I),

$$|\Psi_0\rangle = \sum_I^{\mathcal{M}_0} c_I |\Phi_I\rangle \tag{1.21}$$

The CI coefficients c_I and the MO coefficients vary, and they are simultaneously optimized to approximate a wave function that satisfies the variational principle. Since incorporating all determinants is computationally intractable, restrictions on the orbital occupations are imposed. The orbitals with variable occupation are called *active*. The remaining MOs are classified as *core* or *virtual* if they are doubly occupied or completely unoccupied, respectively [Fig. 1.1]. Subsequently, only determinants that differ within the active space are used to formulate the reference wave function. Such approximate MCSCF method with active orbital partitioning is called complete active space self-consistent field (CASSCF).[31,32]

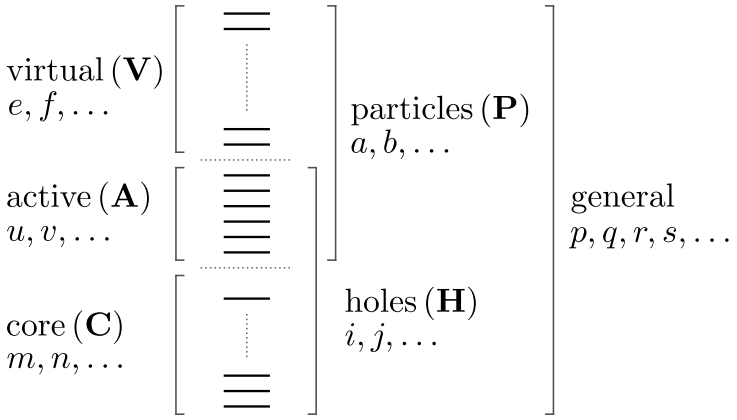

Figure 1.1: Orbital partitioning in CASSCF.

The CASSCF energy could be written in terms of the 1- and 2-particle reduced density matrices γ, the one-electron integral h, and the antisymmetrized ERIs v,

$$E = \sum_m^{\mathbb{C}} h_m^m + \frac{1}{2}\sum_{mn}^{\mathbb{C}} v_{mn}^{mn} + \sum_{uv}^{\mathbb{A}} \left(h_v^u + \sum_m^{\mathbb{C}} v_{vm}^{um} \right) \gamma_v^u + \frac{1}{4}\sum_{uvxy}^{\mathbb{A}} v_{uv}^{xy}\gamma_{xy}^{uv} \tag{1.22}$$

Note that the n-particle reduced density matrices (n-pRDMs) are defined as the expectation value of a product of second-quantized operators with respect to the CASSCF wave function,

$$\gamma_{ij\cdots}^{kl\cdots} = \langle \Psi_0 | \underbrace{\hat{a}_k^\dagger \hat{a}_l^\dagger \cdots}_{n \text{ operators}} \underbrace{\cdots \hat{a}_j \hat{a}_i}_{n \text{ operators}} | \Psi_0 \rangle \tag{1.23}$$

In CASSCF, dynamic correlations ascribed to determinants outside the active space are neglected. There has been various multireference approaches developed to incorporate such dynamic correlations, such as complete active space perturbation theory to the n-th order (CASPTn),[33–36] n-electron valence state perturbation theory,[37–39] multireference CI (MRCI),[40–42] and multireference coupled cluster (MRCC).[13,26] The details of these methods are beyond the scope of this thesis since they are overly comprehensive. Generally, these MR theories could yield reliable results for strongly correlated systems, though at the cost of high computational complexity.[42,43]

1.2 Driven Similarity Renormalization Group

1.2.1 Methodology

Multireference electronic structure theories have gained great success in accurately evaluating near-degenerate electronic states, by simultaneously capturing static and dynamic correlations with multideterminantal wave functions and associated perturbation theories. However, besides the extremely high cost, many multireference approaches suffer from the infamous intruder state problem. The intruder states may be introduced when near-degenerate determinants exist within the reference space and its orthogonal complements, causing small energy denominators.[44–46] This issue is often reflected in potential energy surfaces with unwanted discontinuities [see Fig. 1.2].[47]

One common treatment to address this issue is applying parametrized level shifts

to the diagonal components of the zeroth-order one-body Hamiltonian matrix.[48,49] Nevertheless, the empirical level shifts may affect spectroscopic constants and cause to predict the wrong ordering of states.[44,50] Another well-known solution assumes a Dyall's regularized zeroth-order Hamiltonian with bielectronic terms, as done in NEVPT2.[37,38] Unfortunately, in this case, intruders may be reintroduced if approximations are applied to three- and four-body density cumulants.[51]

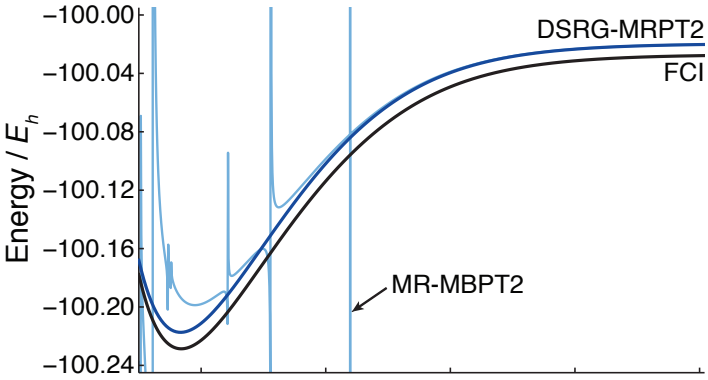

Figure 1.2: Potential energy curves for the $X\,^1\Sigma^+$ state of HF computed using various methods and the cc-pVDZ basis set. All multireference perturbation theories employed a CASSCF(2,2) reference. The fluorine $1s$ orbital was excluded from the correlation treatment. DSRG-MRPT2 employed a flow parameter of $s = 0.5\ E_\mathrm{h}^{-2}$, and the MR-MBPT2 curve is identical to the curve of DSRG-MRPT2($s \to \infty$).[47]

Our group recently developed the multireference driven similarity renormalization group (MR-DSRG), an alternative MR formalism that is immune to the intruder state problem.[47,52–54] The DSRG was originally motivated by the SRG formalism in quantum field theory and condensed matter physics, proposed by Głazek, Wilson, and Wegner.[55–57] Similar to SRG, within DSRG, the bare Hamiltonian ($\hat{H}$) is continuously transformed into an effective Hamiltonian [$\bar{H}(s)$] with a set of parametrized unitary operators [$\hat{U}(s)$],

$$\bar{H}(s) = \hat{U}^\dagger(s)\hat{H}\hat{U}(s) \tag{1.24}$$

where $s \in [0, +\infty)$ is defined as the DSRG flow parameter. The meaning of s will be discussed later when we review the multireference DSRG second-order perturbation theory (DSRG-MRPT2).[47] In general, such a flow parameter affects the Hamiltonian

in the following way (shown in Fig. 1.3): (1) when s equals 0, the unitary operator $[\hat{U}(0)]$ acts like an identity operator, thus the DSRG effective Hamiltonian is identical to the bare Hamiltonian; (2) as s increases, the transformation gradually decouples the reference state from its excited configurations; (3) in the limit that s reaches infinity, the unitary transformation is postulated to completely decouple the reference from excited determinants outside the active space, that is, the effective Hamiltonian is fully block-diagonalized. Conclusively, the transformed Hamiltonian incorporates the correlation contribution from those decoupled excited configurations and could be diagonalized to yield wave functions of both ground and excited states.

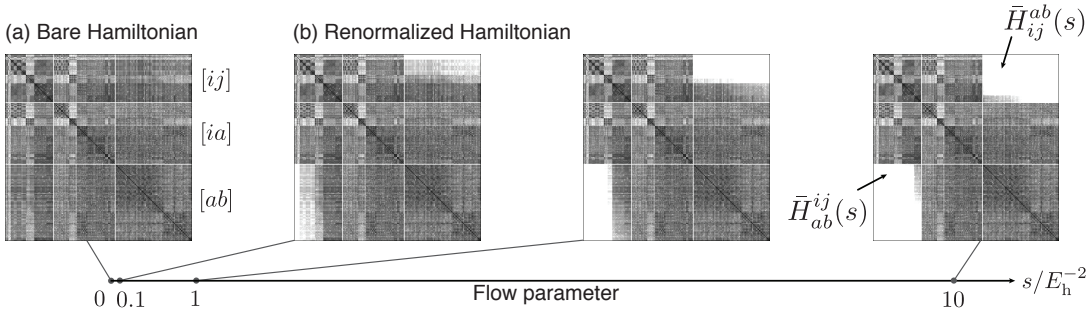

Figure 1.3: Example of the evolution of the two-body components of the transformed Hamiltonian $[\bar{H}^{rs}_{pq}(s)]$ as a function of the flow parameter s in the single-reference driven similarity renormalization group. $\bar{H}^{rs}_{pq}(s)$ is represented as a plot of the matrix $M_{[pq],[rs]}(s) = \bar{H}^{rs}_{pq}(s)$ where $[pq]$ and $[rs]$ are composite indices. The composite indices are divided into three sets: occupied-occupied ($[ij]$), occupied-virtual ($[ia]$), and virtual-virtual ($[ab]$), and the matrix plot shows nine distinct blocks that originate from various combinations of composite indices. For increasing values of s, the DSRG achieves an increasing decoupling of the block corresponding to $\bar{H}^{ij}_{ab}(s) = \langle \Phi^{ab}_{ij} | \bar{H}(s) | \Phi \rangle$, which is responsible for the coupling of the reference (Φ) to doubly excited determinants (Φ^{ab}_{ij}). Taken from Ref. 47.

Unlike the SRG routine in which the transformation is realized by solving a set of ordinary differential equations (ODEs), the DSRG formalism consists of a set of non-linear equations. Specifically, the DSRG unitary operator is expressed in exponential form, similar to the unitary coupled cluster (uCC) ansatz,[58]

$$\hat{U}(s) = e^{\hat{A}(s)} \tag{1.25}$$

where $\hat{A}(s)$ is an anti-Hermitian operator. The operator $\hat{A}(s)$ is a sum of k-body

operators truncated at rank n,

$$\hat{A}(s) = \sum_{k=1}^{n} \hat{A}_k(s) = \sum_{k=1}^{n} \left[\hat{T}_k(s) - \hat{T}_k^\dagger(s)\right] \tag{1.26}$$

where the operator $\hat{T}_k(s)$ and its Hermitian conjugate $\hat{T}_k^\dagger(s)$ correspond to excitation and de-excitation operators, respectively. The k-body DSRG cluster amplitudes are defined as,

$$\hat{T}_k(s) = \frac{1}{(k!)^2} \sum_{ij\cdots}^{\mathbf{H}} \sum_{ab\cdots}^{\mathbf{P}} t_{ab\cdots}^{ij\cdots}(s) \left\{\hat{a}_{ij\cdots}^{ab\cdots}\right\} \tag{1.27}$$

where $\{\cdots\}$ represents a series of normal-ordered creation and annihilation operators. It is worth mentioning that the cluster amplitudes ($t_{ab\cdots}^{ij\cdots}$) are antisymmetrized, in other words, any single permutation between either two upper or lower indices will introduce a minus sign. Nevertheless, the permutation between an upper and a lower index is generally prohibited. Another important point is that all internal excitations or de-excitations are assumed to be excluded, implying that if all indices of an amplitude belong to active orbitals, then the amplitude is zero. Specifically, $t_{uv\cdots}^{xy\cdots} = 0, \forall\, u, v, x, y, \cdots \in \mathbb{A}$. To fulfill the DSRG many-body conditions, we normally truncate $\hat{A}(s)$ rank to $n = 2$, considering only one- and two-body operators. Once these s-dependent operators are solved for, the electronic energy could be expressed as the expectation value of the effective DSRG Hamiltonian,

$$E(s) = \langle\Phi|\bar{H}(s)|\Phi\rangle = \langle\Phi|e^{-\hat{A}(s)}\hat{H}e^{\hat{A}(s)}|\Phi\rangle \tag{1.28}$$

We would like to clarify that this energy is only evaluated at a specific value of s without numerical integration as done in SRG.

1.2.2 DSRG-MRPT2

Among developed MR-DSRG approaches, DSRG-MRPT2 offers the best compromise between cost and accuracy. This method could yield PESs of comparable accuracy to other MRPT2 theories, and meanwhile circumvent the use of four-body reduced density matrices (4-RDMs) of the CASCI wave function, that are required in

CASPT2 and NEVPT2. Furthermore, DSRG-MRPT2 is applicable to large systems with more than two thousand basis functions and large active spaces, if appropriate integral-factorized techniques and CASCI approximations are applied.[59,60]

In DSRG-MRPT2, the zeroth-order Hamiltonian is chosen to be a diagonal normal-ordered Fock operator. The electronic energy can then be partitioned into a reference energy E_0 and a second-order correction term $E^{(2)}(s)$,

$$E(s) = E_0 + E^{(2)}(s) \tag{1.29}$$

In our study, a CASSCF reference wave function is used, thus E_0 is the CASSCF energy [Eq. (1.22)]. The second-order contribution is evaluated in terms of a full contraction with first-order quantities,

$$E^{(2)}(s) = \langle \Psi_0 | [\tilde{H}^{(1)}(s), \hat{T}^{(1)}(s)] | \Psi_0 \rangle \tag{1.30}$$

where $\tilde{H}^{(1)}(s)$ is a first-order effective Hamiltonian [Eq. (1.31)] and $\hat{T}^{(1)}(s)$ is a first-order cluster operator [Eq. (1.32)], both expressed in terms of first-order DSRG modified integrals $[\tilde{h}_i^{a,(1)}(s), \tilde{v}_{ij}^{ab,(1)}(s)]$ and first-order DSRG cluster amplitudes $[t_a^{i,(1)}(s), t_{ab}^{ij,(1)}(s)]$,

$$\tilde{H}^{(1)}(s) = \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \tilde{h}_i^{a,(1)}(s) \{\hat{a}_a^i\} + \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \tilde{v}_{ij}^{ab,(1)}(s) \{\hat{a}_{ab}^{ij}\} \tag{1.31}$$

$$\hat{T}^{(1)}(s) = \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} t_a^{i,(1)}(s) \{\hat{a}_i^a\} + \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} t_{ab}^{ij,(1)}(s) \{\hat{a}_{ij}^{ab}\} \tag{1.32}$$

The superscript "(1)" stands for "first-order" and the label "(s)" indicates that a parameter is s-dependent. In the following text these are omitted for simplicity. The quantities that enter into the DSRG-MRPT2 equations are then:

$$t_a^i = \left(f_a^i + \Delta_u^x t_{ax}^{iu} \gamma_u^x\right) \mathcal{R}_s(\Delta_a^i) \tag{1.33}$$

$$t_{ab}^{ij} = v_{ij}^{ab} \mathcal{R}_s(\Delta_{ab}^{ij}) \tag{1.34}$$

$$\tilde{h}_i^a = f_i^a \mathcal{P}_s(\Delta_a^i) + \Delta_u^x t_{ax}^{iu} \gamma_u^x \left[\mathcal{P}_s(\Delta_a^i) - 1\right] \tag{1.35}$$

$$\tilde{v}_{ij}^{ab} = v_{ij}^{ab} \mathcal{P}_s(\Delta_{ab}^{ij}) \tag{1.36}$$

where we define the auxiliary functions $\mathcal{R} = (1 - e^{-s\Delta^2})/\Delta$ and $\mathcal{P} = 1 + e^{-s\Delta^2}$. Recalling the discussions on the flow parameter s, we note that: (1) Both $\mathcal{R}$ and $\mathcal{P}$ are not numerically divergent given near-zero Møller–Plesset denominators, indicating that the DSRG energy is immune to intruder states regardless of the s value; (2) Although arbitrary values of s will not cause numerical issues, a large value of s must be avoided. It could be mathematically verified that $\mathcal{R}$ has a maximum value around $0.6382\sqrt{s}$, and intruder states might be reintroduced when using a large s value. Our previous study has shown that values of s around 1 E_h^{-2} recovers a sufficient amount of dynamic correlation without overestimation.[47]

In brief, we summarize expressions for the reference energy E_0 and the second-order correlation energy $E^{(2)}(s)$ in Table 1.1 , where λ denotes the $n-$body density

Table 1.1: DSRG-MRPT2 energy expressions.

Term	Energy Expression
E_0	$f^q_p\gamma^p_q - \frac{1}{2}v^{rs}_{pq}\gamma^p_r\gamma^q_s + \frac{1}{4}v^{rs}_{pq}\lambda^{pq}_{rs}$
$E^{(2)}$	$+\tilde{h}^b_j t^i_a \gamma^j_i \gamma^a_b + \frac{1}{4}\tilde{v}^{cd}_{kl} t^{ij}_{ab} \gamma^k_i \gamma^l_j \eta^a_c \eta^b_d + \frac{1}{4}\left(\tilde{v}^{uv}_{iz} t^{iw}_{xy} + \tilde{v}^{wa}_{xy} t^{uv}_{az}\right)\lambda^{xyz}_{uvw}$ $+\frac{1}{2}\left(\tilde{v}^{ev}_{xy} t^u_e - \tilde{v}^{uv}_{my} t^m_x + \tilde{h}^e_x t^{uv}_{ey} - \tilde{h}^v_m t^{um}_{xy}\right)\lambda^{xy}_{uv}$ $+\frac{1}{8}\left(\tilde{v}^{cd}_{xy} t^{uv}_{ab} \eta^a_c \eta^b_d + \tilde{v}^{uv}_{kl} t^{ij}_{xy} \gamma^k_i \gamma^l_j + 8\tilde{v}^{vb}_{jx} t^{iu}_{ay} \gamma^j_i \eta^a_b\right)\lambda^{xy}_{uv}$

cumulant written in terms of $m-$pRDMs [Eq. (1.23)] ($m \in \{1, 2, \cdots, n\}$),[61,62]

$$\lambda^{xy}_{uv} = \gamma^{xy}_{uv} - \gamma^x_u\gamma^y_v + \gamma^x_v\gamma^y_u \tag{1.37}$$

$$\lambda^{xyz}_{uvw} = \gamma^{xyz}_{uvw} - \sum_\pi (-1)^{\mathcal{N}(\pi)}\gamma^x_u\lambda^{yz}_{vw} - \det(\gamma^x_u\gamma^y_v\gamma^z_w) \tag{1.38}$$

In Eq. (1.38), $\det(\cdot)$ represents a sum of all upper or lower index permutations associated with individual sign factor (1 or -1) based on the parity of permutations, and $\sum_\pi(-1)^{\mathcal{N}(\pi)}$ represents a sum over all permutations of the upper and lower indices with a sign factor determined by the number of inversions in π [$\mathcal{N}(\pi)$].

1.3 Analytic Gradient Theory

In the previous sections, we reviewed basic concepts and prominent examples of electronic structure theories. Aside from electronic energy, there are other important physical quantities often substantially overlooked, which also help quantum chemists understand or predict chemical phenomena.

Herein, we discuss the gradient, defined as the vector of first-order energy derivatives with respect to nuclear displacements,

$$\overrightarrow{g} = \nabla E = \sum_{i=1}^{n} \frac{\partial E}{\partial x_i} \overrightarrow{e}_i \tag{1.39}$$

where the nabla symbol ∇ denotes the vector differential operator, and $\overrightarrow{e}_i$ is a normalized vector in a n dimensional space. Since $\overrightarrow{g}$ is the force acting on the nuclei, the gradient plays an important role in first-principle geometry optimizations and molecular dynamics simulations.[63–65] The gradient could further be used to evaluate the Hessian, the matrix of second-order energy derivatives, which allows one to evaluate spectroscopic properties, including vibrational frequencies and IR intensities.[66,67]

One straightforward way to obtain gradients is to use the so-called finite-difference approximation (FD). For example, a five-point stencil technique involves evaluating the following expression,

$$g_x \approx \frac{-E(x+2\Delta x) + 8E(x+\Delta x) - 8E(x-\Delta x) + E(x-2\Delta x)}{12\Delta x} \tag{1.40}$$

where E is the electronic energy computed at discrete grid points, and Δx denotes the displacement. The error is of order $(\Delta x)^4$, thus reliable gradient estimations could be obtained for small displacements (e.g. $\Delta x = 0.005$Å). Nevertheless, the required number of single-point energy computations increases linearly with system size, making FD methods impractical for large molecules. Another practical issue is making sure that FD computations all converge to the same electronic states, especially when geometric displacements lower the molecular symmetry.[68,69]

To effective address these bottlenecks, computing analytic gradients from reliable wave functions has been a preoccupation in quantum chemistry for decades.[70] In contrast to FD, the computational cost of the analytic routine scales with system size as an energy computation and circumvents the evaluation of energies at distorted structures. For example, for the DSRG-MRPT2 approach, an analytic gradient computation on the pentacene molecule is about 70 times faster than a FD one [see chapter 4]. The analytic derivative theory is based on the Hellmann-Feyman theorem,[71] which states that for an exact eigenstate or a variational wave function, the energy derivative with respect to a perturbative parameter λ is given by the derivatives of the Hamiltonian with respect to λ,

$$\frac{dE_\lambda}{d\lambda} = \langle \psi_\lambda | \frac{d\hat{H}_\lambda}{d\lambda} | \psi_\lambda \rangle \tag{1.41}$$

Due to the fact that for many electronic structure theories the energy implicitly depends on the MO coefficients [Eq. (1.5)] and CI coefficients [Eq. (1.10)], the analytic gradient can be solved from a series of reference-dependent coupled-perturbed (CP) equations, such as the CP-SCF and CP-CASSCF equations.[72,73] Specifically, the CP equations assume that the perturbed MOs are expanded in the unperturbed basis,

$$|i\rangle \rightarrow \sum_j |j\rangle \, U^r_{ji} \tag{1.42}$$

where U denotes coupled-perturbed coefficients. This theoretical approach is applicable to any methods whose orbitals are formulated from SCF or MCSCF reference computations. Despite its generality and effectiveness, the CP formalism is limited to small systems. For a molecule with N atoms, there are $3N$ CP equations, showing a computational complexity similar to FD gradients. An alternative, the Z-vector approach, proposed by Handy and Schaefer, reduces the $3N$ CP equations to a single response equation independent of nuclear displacements, whose solution is sufficient to compute analytic energy derivatives of all nuclei.[74,75] As discussed by Helgaker and Jørgensen, methods that incorporate redundant or non-variational parameters require

a more general treatment based on Lagrange multipliers.[70] For example, the MO and CI coefficients used in the unrelaxed DSRG-MRPT2 are determined by the CASSCF and CASCI stationary conditions, but they are not variationally optimized. There are also redundant parameters in DSRG-MRPT2, such as modified n-body integrals and cluster amplitudes, which could be expressed in terms of other quantities.[47] The general method of Lagrange multipliers forms the Lagrangian function $\mathcal{L}$,

$$\mathcal{L}(x; \lambda_1, \cdots, \lambda_n) = E(x) + \sum_{i=1}^{n} \lambda_i f_i(x) \tag{1.43}$$

where f_i denotes a Lagrangian constraint, and λ_i is its associated Lagrange multiplier.

Unfortunately, although these schemes and associated equations are well established, developing the analytic energy gradients for MR theories is still challenging. It requires extensive mathematical derivations and comprehensive software implementations. To the best of our knowledge, the first MRPT2 gradient theory was done by Nakano *et al.* on the analytic gradients of multi-configurational quasi-degenerate perturbation theory (MC-QDPT).[76,77] Recently, the gradient theory for extended MC-QDPT (XMC-QDPT) was implemented to optimize conical intersections of a retinal model chromophore.[78–80] Subsequently, analytic derivatives and its excited-state extensions were extended by Werner and coworkers to partially internally contracted CASPT2, and by Hoffmann's generalized Van Vleck perturbation theory.[81–83] The gradients for fully contracted CASPT2 were implemented two decades after the debut of the original theory, using automatic code generations by Shiozaki *et al.*, circumventing laborious manual derivations.[84] Shiozaki and co-workers further developed gradient theories for CASPT2 with density fitting and imaginary shift,[85,86] multi-state extensions of CASPT2,[87] extend multi-state (XMS) CASPT2,[34] and explicitly-correlated CASPT2 (CASPT2-F12).[88,89] Another important achievement is the implementation of analytic gradients for strongly-contracted, partially contracted, and quasidegenerate NEVPT2 by Park and Nishimoto.[90,90–92] In addition, Nishimoto formulated the derivative theory for the restricted active space second-order perturba-

tion theory (RASPT2).[93] Most recently, using automatic code generation, Abbott and coworkers presented a state-of-art automatic differentiation technique, capable of yielding arbitrary-order electronic energy derivatives with respect to nuclear coordinates for common methods like HF, MPn and CCSD(T).[94] Although such approach has limited practical applications, it could be useful to generate reference data for developing efficient implementations of analytic gradients.

1.4 Prospectus

In chapter 2, we introduce the analytic gradient theory for the single-reference DSRG second-order perturbation theory (DSRG-SRPT2) using the method of Lagrange multipliers. We reveal that the DSRG parameters do not introduce insurmountable difficulties in the derivation of gradients, and the resulting first-order derivatives possess the same asymptotic computational scaling of MP2. We conduct geometry optimizations for fifteen small molecules, and the equilibrium structures are compared to the ones obtained by other single-reference theories. In chapter 3, we subsequently develop the analytic gradient theory for the unrelaxed DSRG-MRPT2 method, considering additional Lagrangian constraints. The equilibrium bond lengths and angles of p-benzyne for both singlet and triplet states are computed and compared with the ones yielded via other MR approaches. We also evaluate the s-dependency of the DSRG-MRPT2 energy, by investigating the adiabatic singlet-triplet splitting as a function of the DSRG flow parameter s. Furthermore, we study an approximate variant of DSRG-MRPT2, which completely neglects the contribution of three-body density cumulants. In chapter 4, we apply the density fitting technique to extend the DSRG-MRPT2 analytic energy gradients to large molecular systems. The optimized geometries, vibrational frequencies, and IR intensities of p-benzyne with large orbital basis and active spaces are reported, and compared with experimental data and other electronic structure theories. Finally, in chapter 5, we summarize these works and

discuss prospects for future developments.

Bibliography

[1] Knowles, P. J.; Handy, N. C. A determinant based full configuration interaction program. *Comput. Phys. Commun.* **1989**, *54*, 75–83.

[2] Szabo, A.; Ostlund, N. S. *Modern quantum chemistry: introduction to advanced electronic structure theory*; Courier Corporation, 2012.

[3] Woolley, R.; Sutcliffe, B. Molecular structure and the born—Oppenheimer approximation. *Chem. Phys. Lett.* **1977**, *45*, 393–398.

[4] Knowles, P. J.; Handy, N. C. A new determinant-based full configuration interaction method. *Chem. Phys. Lett.* **1984**, *111*, 315–321.

[5] Olsen, J.; Jørgensen, P.; Simons, J. Passing the one-billion limit in full configuration-interaction (FCI) calculations. *Chem. Phys. Lett.* **1990**, *169*, 463–472.

[6] Dutta, A.; Sherrill, C. D. Full configuration interaction potential energy curves for breaking bonds to hydrogen: An assessment of single-reference correlation methods. *J. Chem. Phys.* **2003**, *118*, 1610–1619.

[7] Harris, F. E.; Monkhorst, H. J.; Freeman, D. L. *Algebraic and diagrammatic methods in many-fermion theory*; Courier Dover Publications, 2020.

[8] Jørgensen, P. *Second quantization-based methods in quantum chemistry*; Elsevier, 2012.

[9] Lykos, P.; Pratt, G. Discussion on the Hartree-Fock approximation. *Rev. Mod. Phys.* **1963**, *35*, 496.

[10] Fischer, C. F. Hartree–Fock method for atoms. A numerical approach. **1977**,

[11] Baerends, E.; Ellis, D.; Ros, P. Self-consistent molecular Hartree—Fock—Slater calculations I. The computational procedure. *Chem. Phys.* **1973**, *2*, 41–51.

[12] Mok, D. K.; Neumann, R.; Handy, N. C. Dynamical and nondynamical correlation. *J. Phys. Chem.* **1996**, *100*, 6225–6230.

[13] Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.

[14] Schlegel, H. B. Møller-Plesset perturbation theory with spin projection. *J. Phys. Chem.* **1988**, *92*, 3075–3078.

[15] Cremer, D. Møller–Plesset perturbation theory: from small molecule methods to methods for thousands of atoms. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 509–530.

[16] Pulay, P.; Saebø, S. Orbital-invariant formulation and second-order gradient evaluation in Møller-Plesset perturbation theory. *Theor. Chim. Acta* **1986**, *69*, 357–368.

[17] Grüneis, A.; Marsman, M.; Kresse, G. Second-order Møller–Plesset perturbation theory applied to extended systems. II. Structural and energetic properties. *J. Chem. Phys.* **2010**, *133*, 074107.

[18] Shavitt, I. *Methods of electronic structure theory*; Springer, 1977; pp 189–275.

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

[20] Bartlett, R. J.; Musiał, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

[21] Crawford, T. D.; Schaefer, H. F. An introduction to coupled cluster theory for computational chemists. *Rev. Comput. Chem.* **2000**, *14*, 33–136.

[22] Bartlett, R. J. Coupled-cluster approach to molecular structure and spectra: a step toward predictive quantum chemistry. *J. Phys. Chem.* **1989**, *93*, 1697–1708.

[23] Hollett, J. W.; Gill, P. M. The two faces of static correlation. *J. Chem. Phys.* **2011**, *134*, 114111.

[24] Bulik, I. W.; Henderson, T. M.; Scuseria, G. E. Can single-reference coupled cluster theory describe static correlation? *J. Chem. Theory Comput.* **2015**, *11*, 3171–3179.

[25] Szalay, P. G.; Muller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration self-consistent field and multireference configuration interaction methods and applications. *Chem. Rev.* **2012**, *112*, 108–181.

[26] Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference nature of chemistry: The coupled-cluster view. *Chem. Rev.* **2012**, *112*, 182–243.

[27] Köhn, A.; Hanauer, M.; Mueck, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 176–197.

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

[29] Hedegård, E. D.; List, N. H.; Jensen, H. J. A.; Kongsted, J. The multi-configuration self-consistent field method within a polarizable embedded framework. *J. Chem. Phys.* **2013**, *139*, 044101.

[30] Hinze, J.; Roothaan, C. C. Multi-configuration self-consistent-field theory. *Prog. Theor. Phys.* **1967**, *40*, 37–51.

[31] Roos, B. O. The complete active space self-consistent field method and its applications in electronic structure calculations. *Adv. Chem. Phys.* **1987**, *69*, 399–445.

[32] Olsen, J. The CASSCF method: A perspective and commentary. *Int. J. Quantum Chem* **2011**, *111*, 3267–3272.

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

[34] Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **2011**, *135*, 081106.

[35] Menezes, F.; Kats, D.; Werner, H.-J. Local complete active space second-order perturbation theory using pair natural orbitals (PNO-CASPT2). *J. Chem. Phys.* **2016**, *145*, 124115.

[36] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

[37] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

[38] Angeli, C.; Pastore, M.; Cimiraglia, R. New perspectives in multireference perturbation theory: the n-electron valence state approach. *Theor. Chem. Acc.* **2007**, *117*, 743–754.

[39] Angeli, C.; Bories, B.; Cavallini, A.; Cimiraglia, R. Third-order multireference

perturbation theory: The n-electron valence state perturbation-theory approach. *J. Chem. Phys.* **2006**, *124*, 054108.

[40] Szalay, P. G.; Bartlett, R. J. Multi-reference averaged quadratic coupled-cluster method: a size-extensive modification of multi-reference CI. *Chem. Phys. Lett.* **1993**, *214*, 481–488.

[41] Bruna, P. J.; Peyerimhoff, S. D.; Buenker, R. J. The ground state of the CN+ ion: a multi-reference CI study. *Chem. Phys. Lett.* **1980**, *72*, 278–284.

[42] Khedkar, A.; Roemelt, M. Modern multireference methods and their application in transition metal chemistry. *Phys. Chem. Chem. Phys.* **2021**,

[43] Park, J. W.; Al-Saadon, R.; MacLeod, M. K.; Shiozaki, T.; Vlaisavljevich, B. Multireference electron correlation methods: Journeys along potential energy surfaces. *Chem. Rev.* **2020**, *120*, 5878–5909.

[44] Camacho, C.; Witek, H. A.; Yamamoto, S. Intruder states in multireference perturbation theory: The ground state of manganese dimer. *J. Comput. Chem.* **2009**, *30*, 468–478.

[45] Evangelisti, S.; Daudey, J.; Malrieu, J. Qualitative intruder-state problems in effective Hamiltonian theory and their solution through intermediate Hamiltonians. *Phys. Rev. A* **1987**, *35*, 4930.

[46] Chattopadhyay, S.; Chaudhuri, R. K.; Mahapatra, U. S.; Ghosh, A.; Ray, S. S. State-specific multireference perturbation theory: development and present status. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 266–291.

[47] Li, C.; Evangelista, F. A. Multireference driven similarity renormalization group: A second-order perturbative analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.

[48] Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. Intruder state avoidance multireference Møller–Plesset perturbation theory. *J. Comput. Chem.* **2002**, *23*, 957–965.

[49] Choe, Y.-K.; Witek, H. A.; Finley, J. P.; Hirao, K. Identifying and removing intruder states in multireference Møller–Plesset perturbation theory. *J. Chem. Phys.* **2001**, *114*, 3913–3918.

[50] Camacho, C.; Cimiraglia, R.; Witek, H. A. Multireference perturbation theory can predict a false ground state. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5058–5060.

[51] Zgid, D.; Ghosh, D.; Neuscamman, E.; Chan, G. K.-L. A study of cumulant approximations to n-electron valence multireference perturbation theory. *J. Chem. Phys.* **2009**, *130*, 194107.

[52] Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.

[53] Li, C.; Evangelista, F. A. Multireference theories of electron correlation based on the driven similarity renormalization group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245–273.

[54] Li, C.; Evangelista, F. A. Driven similarity renormalization group: Third-order multireference perturbation theory. *J. Chem. Phys.* **2017**, *146*, 124132.

[55] Wegner, F. Flow-equations for Hamiltonians. *Ann. Phys.* **1994**, *506*, 77–91.

[56] Wegner, F. *Advances in Solid State Physics*; Springer, 2000; pp 133–142.

[57] Głazek, S. D.; Wilson, K. G. Renormalization of hamiltonians. *Phys. Rev. D* **1993**, *48*, 5863.

[58] Taube, A. G.; Bartlett, R. J. New perspectives on unitary coupled-cluster theory. *Int. J. Quantum Chem.* **2006**, *106*, 3393–3401.

[59] Schriber, J. B.; Hannon, K. P.; Li, C.; Evangelista, F. A. A combined selected configuration interaction and many-body treatment of static and dynamical correlation in oligoacenes. *J. Chem. Theory Comput.* **2018**, *14*, 6295–6305.

[60] Hannon, K. P.; Li, C.; Evangelista, F. A. An integral-factorized implementation of the driven similarity renormalization group second-order multireference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 204111.

[61] Kutzelnigg, W.; Mukherjee, D. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.* **1997**, *107*, 432–449.

[62] Kutzelnigg, W.; Mukherjee, D. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* **1999**, *110*, 2800–2809.

[63] Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci.* **2005**, *102*, 6654–6659.

[64] Park, J. W.; Shiozaki, T. On-the-fly CASPT2 surface-hopping dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 3676–3683.

[65] Curchod, B. F.; Martínez, T. J. Ab initio nonadiabatic quantum molecular dynamics. *Chem. Rev.* **2018**, *118*, 3305–3336.

[66] Lotrich, V.; Flocke, N.; Ponton, M.; Yau, A.; Perera, A.; Deumens, E.; Bartlett, R. Parallel implementation of electronic structure energy, gradient, and Hessian calculations. *J. Chem. Phys.* **2008**, *128*, 194104.

[67] Spicher, S.; Grimme, S. Single-point Hessian calculations for improved vibrational frequencies and rigid-rotor-harmonic-oscillator thermodynamics. *J. Chem. Theory Comput.* **2021**, *17*, 1701–1714.

[68] Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. Problematic p-benzyne: Orbital instabilities, biradical character, and broken symmetry. *J. Chem. Phys.* **2001**, *114*, 10638–10650.

[69] Li, X.; Paldus, J. Force field of para-and metabenzyne diradicals: A multireference coupled-cluster study. *J. Chem. Phys.* **2010**, *132*, 114103.

[70] Helgaker, T.; Jørgensen, P. Analytical calculation of geometrical derivatives in molecular electronic structure theory. *Adv. Quantum Chem.* **1988**, *19*, 183–245.

[71] Politzer, P.; Murray, J. S. The Hellmann-Feynman theorem: a perspective. *J. Mol. Model.* **2018**, *24*, 1–7.

[72] Yamaguchi, Y.; Osamura, Y.; Goddard, J.; Schaefer, H. Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory, A New Dimension to Quantum Chemistry. **1994**,

[73] Osamura, Y.; Yamaguchi, Y.; Schaefer III, H. F. Generalization of analytic configuration interaction (CI) gradient techniques for potential energy hypersurfaces, including a solution to the coupled perturbed Hartree–Fock equations for multiconfiguration SCF molecular wave functions. *J. Chem. Phys.* **1982**, *77*, 383–390.

[74] Handy, N. C.; Schaefer III, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[75] Amos, R. D.; Rice, J. E. Implementation of analytic derivative methods in quantum chemistry. *Comput. Phys. Rep.* **1989**, *10*, 147–187.

[76] Nakano, H.; Hirao, K.; Gordon, M. S. Analytic energy gradients for multiconfigurational self-consistent field second-order quasidegenerate perturbation theory (MC-QDPT). *J. Chem. Phys.* **1998**, *108*, 5660–5669.

[77] Nakano, H. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

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

[79] Park, J. W. Analytical first-order derivatives of second-order extended multi-configuration quasi-degenerate perturbation theory (XMCQDPT2): Implementation and application. *J. Chem. Theory Comput.* **2020**, *16*, 5562–5571.

[80] Park, J. W. Analytical Gradient Theory for Resolvent-Fitted Second-Order Extended Multiconfiguration Perturbation Theory (XMCQDPT2). *J. Chem. Theory Comput.* **2021**, *17*, 6122–6133.

[81] Celani, P.; Werner, H.-J. Analytical energy gradients for internally contracted second-order multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 5044–5057.

[82] Dudley, T. J.; Khait, Y. G.; Hoffmann, M. R. Molecular gradients for the second-order generalized Van Vleck variant of multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 651–660.

[83] Theis, D.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 energy gradient using a Lagrangian formulation. *J. Chem. Phys.* **2011**, *135*, 044117.

[84] MacLeod, M. K.; Shiozaki, T. Communication: Automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory. *J. Chem. Phys.* **2015**, *142*, 051103.

[85] Park, J. W.; Al-Saadon, R.; Strand, N. E.; Shiozaki, T. Imaginary shift in

CASPT2 nuclear gradient and derivative coupling theory. *J. Chem. Theory Comput.* **2019**, *15*, 4088–4098.

[86] Győrffy, W.; Shiozaki, T.; Knizia, G.; Werner, H.-J. Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* **2013**, *138*, 104104.

[87] Vlaisavljevich, B.; Shiozaki, T. Nuclear energy gradients for internally contracted complete active space second-order perturbation theory: Multistate extensions. *J. Chem. Theory Comput.* **2016**, *12*, 3781–3787.

[88] Shiozaki, T.; Werner, H.-J. Communication: Second-order multireference perturbation theory with explicit correlation: CASPT2-F12. *J. Chem. Phys.* **2010**, *133*, 141103.

[89] Shiozaki, T. BAGEL: brilliantly advanced general electronic-structure library. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1331.

[90] Park, J. W. Analytical gradient theory for strongly contracted (SC) and partially contracted (PC) N-electron valence state perturbation theory (NEVPT2). *J. Chem. Theory Comput.* **2019**, *15*, 5417–5425.

[91] Nishimoto, Y. Analytic first-order derivatives of partially contracted n-electron valence state second-order perturbation theory (PC-NEVPT2). *J. Chem. Phys.* **2019**, *151*, 114103.

[92] Nishimoto, Y. Locating conical intersections using the quasidegenerate partially and strongly contracted NEVPT2 methods. *Chem. Phys. Lett.* **2020**, *744*, 137219.

[93] Nishimoto, Y. Analytic gradients for restricted active space second-order perturbation theory (RASPT2). *J. Chem. Phys.* **2021**, *154*, 194103.

[94] Abbott, A. S.; Abbott, B. Z.; Turney, J. M.; Schaefer III, H. F. Arbitrary-order derivatives of quantum chemical methods via automatic differentiation. *J. Phys. Chem. Lett.* **2021**, *12*, 3232–3239.

Chapter 2 Analytic Gradients for DSRG-SRPT2

Chapter Abstract

We derive and implement analytic energy gradients for the single-reference driven similarity renormalization group second-order perturbation theory (DSRG-PT2). The resulting equations possess an asymptotic scaling that is identical to that of the second-order Møller–Plesset perturbation theory (MP2), indicating that the exponential regularizer in the DSRG equations does not introduce formal difficulties in the gradient theory. We apply the DSRG-PT2 method to optimize the geometries of fifteen small molecules. The equilibrium bond lengths computed with DSRG-PT2 are found similar to those of MP2, yielding a mean absolute error of 0.0033 Å and a standard deviation of 0.0045 Å when compared with coupled cluster with singles, doubles and perturbative triples.

2.1 introduction

Second-order Møller–Plesset perturbation theory (MP2) features prominently in quantum chemistry—it is the simplest wave-function method that includes electron cor relation effects. Highly efficient algorithms have been developed that enable MP2 computations on molecular systems with up to thousands of atoms.[1,2] Meanwhile, numerous variants have been proposed to improve the accuracy or applicability of MP2. For example, spin-component scaled MP2 improves the energetics for non-covalent interactions,[3–5] while orbital-optimized (OO) MP2 is applicable to symmetry-breaking systems.[6–9] More recently, it has been shown that spin-projected regularized OO-MP2

Reproduced from Wang, S., Li, C., & Evangelista, F. A. *J. Chem. Phys.* **2019**, *151*, 044118, with the permission of AIP Publishing

can provide reliable results even for diradical systems, which is intractable with MP2 due to static correlation effects.[10]

The developments of analytic energy derivatives for MP2 and its variants have provided an efficient way to explore ground-state potential energy surfaces (PESs) of closed-shell or high-spin open-shell systems.[11–17] However, as other single-reference methods,[18–35] reliable predictions are typically obtained only near the equilibrium geometry. A multireference (MR) generalization of MP2 is usually required at stretched geometries or when two or more electronic states become near degenerate.[36–44] Unlike the case of single-reference approaches, analytic derivatives have been formulated only for a few MR methods. Some of the more recent contributions in the context of nuclear energy gradients include MR configuration interaction,[45,46] several second-order MR perturbation theories (MRPTs),[47–52] and the state-specific MR coupled cluster theory.[53,54]

A potential problem in the development of analytic energy gradients for MR formalisms is the presence of discontinuities in the PESs. For instance, the removal of linearly dependent excitations in internally contracted MR methods may lead to discontinuities on PESs.[55,56] Intruder states can also break the continuity of PESs and add extra complexity in the derivation of analytic gradients[48] due to the use of regularization techniques.[57–60] To address these problems, we have recently introduced the MR driven similarity renormalization group (DSRG) approach,[61,62] a many-body formalism that is robust to intruder states and immune to the discontinuity problem of internally-contracted MR theories. Numerical results show that the DSRG-MRPTs yield smooth PESs and results that are similar in accuracy to those obtained by other MRPTs of equal order.[63–66] Given the desirable properties of the MR-DSRG methods, it is then useful to develop the corresponding nuclear energy gradients, especially for the lost-cost second-order scheme (DSRG-MRPT2).[63]

Deriving energy derivatives under the MR framework, in itself, can be an in-

tractable task. For example, analytic gradients for the fully internally contracted second-order complete-active-space perturbation theory (CASPT2) were only very recently attained via automatic code generations,[52] twenty-three years after its debut.[37] Therefore, as an initial step towards the realization of DSRG-MRPT2 gradients, we have decided to address the question: How does the presence of a regularization function in the DSRG equations affect the formulation of analytic gradients? In this work, we will focus on the single-reference DSRG second-order perturbation theory (DSRG-PT2),[67] which is equivalent to DSRG-MRPT2 in the limit of an active space of dimension zero.[63] The DSRG-PT2 gradient theory is considerably simpler than the multireference version, but it should reveal any difficulty that may arise from the presence of a regularizer in the DSRG-MRPT2 equations and enable us to investigate the numerical robustness of the resulting linear-response equations. We also note that the DSRG-PT2 energy expression is equivalent to the second-order perturbative energy (assuming Møller–Plesset partitioning) of the in-medium similarity renormalization group (IM-SRG) approach.[67,68] Hence, this work also provides some insights into the IM-SRG linear response theory.

We derive the DSRG-PT2 gradients using the method of Lagrangian multipliers.[29,48,69–71] Since the DSRG-PT2 energy expressions must be formulated in the canonical basis, we explicitly impose constraints in the Lagrangian to guarantee that the orbitals remain canonical in response to nuclear displacements. This approach has recently been applied to develop analytic gradients for imaginary-shift CASPT2,[72] where a similar problem arises. Alternatively, an orbital-invariant formalism must be adopted, an option that has been explored by Lee and Head-Gordon in their regularized OO-MP2 methods.[9]

We start our discussion by deriving analytic expressions for the DSRG-PT2 gradients and comparing them to the MP2 gradient theory. In Sec. 2.3, we benchmark our implementation by computing the equilibrium molecular geometries of fifteen small

molecules using DSRG-PT2. These results are compared against the geometries obtained at the MP2, coupled cluster with singles and doubles (CCSD),[73] and CCSD with perturbative triples [CCSD(T)][74] levels of theories. We then discuss our findings and future work in Sec. 2.4.

2.2 Theory

2.2.1 Energy and amplitude expressions

We first introduce a set of orthonormal molecular spin orbitals (MSOs) $\mathbf{G} \equiv \{\psi_p(\mathbf{r}, \sigma), p = 1, 2, \ldots, N_{\mathbf{G}}\}$, where $\mathbf{r}$ and σ indicate spatial and spin coordinates, respectively. Each MSO is a product of a spin function (σ_p) and a molecular orbital (MO) $\phi_p(\mathbf{r})$, i.e., $\psi_p(\mathbf{r}, \sigma) = \phi_p(\mathbf{r})\sigma_p(\sigma)$. The MOs are expanded using a set of atomic orbitals (AOs) $\{\chi_\mu(\mathbf{r}), \mu = 1, 2, \ldots, N\}$ as

$$\phi_p(\mathbf{r}) = \sum_{\mu}^{\mathrm{AO}} \chi_\mu(\mathbf{r})\, C_{\mu p}. \tag{2.1}$$

The orbital coefficients ($C_{\mu p}$) are obtained from a self-consistent-field (SCF) computation. The MSO set is separated into occupied ($\mathbf{O}$) and virtual ($\mathbf{V}$) orbitals according to their occupation in the reference determinant. For post-Hartree–Fock theories, each of the two orbital sets is further partitioned into frozen and active parts, where excitations are only allowed among active orbitals. Accordingly, we partition the occupied orbitals into frozen core ($\mathbf{FC}$) and hole ($\mathbf{H}$) orbitals (with $\mathbf{O} = \mathbf{FC} \cup \mathbf{H}$) and the virtual orbitals into particle ($\mathbf{P}$) and frozen virtual ($\mathbf{FV}$) orbitals (with $\mathbf{V} = \mathbf{P} \cup \mathbf{FV}$). The MSO indices used to label orbitals belonging to these sets are summarized in Table 2.1, while AOs are labeled by the indices μ, ν, ρ, τ.

The DSRG-PT2 method can be considered as a variant of MP2 theory, modified using a regularizer that depends on the flow parameter $s \in [0, \infty)$.[67] The DSRG-PT2

Table 2.1: Definition of the molecular spin orbital spaces used in this work.

Space	Symbol	Size	Indices	Correlated[a]	Occupied[b]
Frozen core	**FC**	$N_{\mathbf{FC}}$	I, J	F	T
Hole	**H**	$N_{\mathbf{H}}$	i, j, k, l	T	T
Particle	**P**	$N_{\mathbf{P}}$	a, b, c, d	T	F
Frozen virtual	**FV**	$N_{\mathbf{FV}}$	A, B	F	F
Occupied	**O**	$N_{\mathbf{O}}$	m, n	$\mathbf{H} \cup \mathbf{FC}$	
Virtual	**V**	$N_{\mathbf{V}}$	e, f	$\mathbf{P} \cup \mathbf{FV}$	
General	**G**	$N_{\mathbf{G}}$	p, q, r, s	$\mathbf{O} \cup \mathbf{V}$	

[a] True (T) if electron correlation is considered in the orbital space, otherwise false (F).

[b] True (T) if orbitals are occupied in the reference determinant, otherwise false (F).

first-order amplitude equation reads as

$$t_{ab}^{ij,(1)}(s) = \frac{v_{ab}^{ij}}{\Delta_{ab}^{ij}} \left[1 - e^{-s(\Delta_{ab}^{ij})^2}\right], \quad \forall ij \in \mathbf{H}, \forall ab \in \mathbf{P}, \tag{2.2}$$

where $v_{ab}^{ij} = \langle \psi_a \psi_b \| \psi_i \psi_j \rangle$ are antisymmetrized two-electron integrals and $\Delta_{ab}^{ij} = \epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b$ indicate Møller–Plesset denominators defined by the canonical orbital energies ϵ_p. These orbital energies are given by the diagonal elements of the Fock matrix ($\epsilon_p = f_p^p$). The elements of the Fock matrix are defined in terms of one-electron integrals (h_q^p) and v_{pq}^{rs},

$$f_q^p = h_q^p + \sum_m^{\mathbf{O}} v_{qm}^{pm}. \tag{2.3}$$

Note that the Fock matrix is diagonal in the canonical orbital basis, that is, $f_q^p = 0$ if $p \neq q$.

The DSRG-PT2 correlation energy may be expressed as

$$E^{(2)}(s) = \frac{1}{4} \sum_{ij}^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ij}^{ab} t_{ab}^{ij,(1)}(s) \left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right]. \tag{2.4}$$

Equations (2.2) and (2.4) imply that for any finite value of s, the DSRG-PT2 energy does not diverge even when the denominators Δ_{ab}^{ij} go to zero. If no denominator is null (i.e., $\Delta_{ab}^{ij} \neq 0$), in the limit of s that goes to infinity Eqs. (2.2) and (2.4) approach the MP2 amplitude and energy equations, respectively. Thus, the DSRG-PT2 energy gradient becomes identical to the MP2 gradient in the limit of $s \to \infty$.

Note that the DSRG-PT2 amplitude and energy equations are not invariant under separate rotations of hole and particle orbitals because the exponential regularizer is defined using canonical orbital energies. Consequently, two options are available for formulating DSRG-PT2 gradients: i) generalizing the DSRG-PT2 equations to an orbital-invariant form, or ii) enforcing diagonality of the Fock matrix in the linear response equations. The former solution is problematic, because it increases the computational cost of the DSRG-PT2 due to the need to transform and back-transform the amplitudes to the canonical basis.[62] We instead follow the latter approach which, as we shall demonstrate, leads to a numerically robust computational procedure that has the same asymptotic cost of the MP2 gradient theory.

2.2.2 Energy gradient expression

Since the DSRG-PT2 scheme assumes canonical SCF orbitals, it is convenient to derive equations for analytic gradients using the Lagrangian approach.[29,48,69–71] To this end, we write the DSRG-PT2 Lagrangian as

$$\begin{aligned}\mathcal{L}(s) = & E^{(2)}(s) \\ & + \frac{1}{4}\sum_{ij}^{\mathbf{H}}\sum_{ab}^{\mathbf{P}} \bar{t}_{ij}^{ab,(1)}(s)\left[\tilde{v}_{ab}^{ij}(s) - \Delta_{ab}^{ij} t_{ab}^{ij,(1)}(s)\right] \\ & + \frac{1}{2}\sum_{pq}^{\mathbf{G}} z_q^p (f_p^q - \delta_p^q \epsilon_p) - \frac{1}{2}\sum_{pq}^{\mathbf{G}} \omega_q^p (S_p^q - \delta_p^q),\end{aligned} \tag{2.5}$$

where $S_p^q = \langle \psi_p | \psi_q \rangle$ are overlap integrals and δ_p^q denotes the Kronecker delta. For brevity, we also introduce the shorthand notation $\tilde{v}_{ab}^{ij}(s) = v_{ab}^{ij}[1 - e^{-s(\Delta_{ab}^{ij})^2}]$. In Eq. (2.5), we have enforced three constraints: i) the DSRG-PT2 amplitude conditions (second line), ii) the diagonality of the Fock matrix (first term in the third line), and iii) the orthonormality of the MSOs (last term). To these three constraints we associate the corresponding Lagrange multipliers, $\bar{t}_{ij}^{ab,(1)}(s)$, z_q^p, and ω_q^p, respectively. Following convention, we refer to z_q^p as the DSRG-PT2 one-particle density matrix (OPDM) and to ω_q^p as the energy-weighted density matrix (EWDM), noticing,

however, that other derivations may absorb a factor 1/2 into these expressions.

The Lagrangian [Eq. (2.5)] is equal to the energy expression [Eq. (2.4)] when $\mathcal{L}(s)$ is stationary with respect to all its parameters [$\bar{t}_{ij}^{ab,(1)}(s)$, z_q^p, ω_q^p, $t_{ab}^{ij,(1)}(s)$, ϵ_p, and $C_{\mu p}$], that is, when the following conditions are satisfied

$$\begin{cases} \partial\mathcal{L}(s)/\partial[\bar{t}_{ij}^{ab,(1)}(s)] = 0, & (2.6) \\ \partial\mathcal{L}(s)/\partial z_q^p = 0, & (2.7) \\ \partial\mathcal{L}(s)/\partial\omega_q^p = 0, & (2.8) \\ \partial\mathcal{L}(s)/\partial[t_{ab}^{ij,(1)}(s)] = 0, & (2.9) \\ \partial\mathcal{L}(s)/\partial\epsilon_p = 0, & (2.10) \\ \partial\mathcal{L}(s)/\partial C_{\mu p} = 0. & (2.11) \end{cases}$$

Consequently, the DSRG-PT2 energy gradient with respect to a perturbation x can be written as

$$\left.\frac{\mathrm{d}E^{(2)}(s)}{\mathrm{d}x}\right|_{x=0} = \left.\frac{\mathrm{d}\mathcal{L}(s)}{\mathrm{d}x}\right|_{x=0} = \left.\frac{\partial\mathcal{L}(s)}{\partial x}\right|_{x=0}. \tag{2.12}$$

Note that the stationarity conditions [Eqs. (2.6)–(2.11)] apply in the unperturbed situation ($x = 0$).

Equations (2.6), (2.7), and (2.8) correspond to the three constraints imposed on $\mathcal{L}(s)$ and they are trivial to derive. The amplitude response equation [Eq. (2.9)] may be expressed as:

$$\frac{\partial\mathcal{L}}{\partial\, t_{ab}^{ij,(1)}} = \frac{1}{4}v_{ij}^{ab}\left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right] - \frac{1}{4}\bar{t}_{ij}^{ab,(1)}\Delta_{ab}^{ij} = 0, \tag{2.13}$$

where here and in the following text we drop the label "(s)" for all s-dependent quantities. The amplitude Lagrange multipliers [$\bar{t}_{ij}^{ab,(1)}$] are thus given by

$$\bar{t}_{ij}^{ab,(1)} = \frac{v_{ij}^{ab}}{\Delta_{ab}^{ij}}\left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right] = t_{ab}^{ij,(1)}\frac{1 + e^{-s(\Delta_{ab}^{ij})^2}}{1 - e^{-s(\Delta_{ab}^{ij})^2}}. \tag{2.14}$$

Note that according to Eq. (2.14) the value of $\bar{t}_{ij}^{ab,(1)}$ diverges when $\Delta_{ab}^{ij} \to 0$; however, the Lagrangian terms that depend on $\bar{t}_{ij}^{ab,(1)}$ will *not* diverge. It is easy to show that

terms that contribute to the second line of Eq. (2.5) individually go to zero:

$$\lim_{\Delta_{ab}^{ij}\to 0} \bar{t}_{ij}^{ab,(1)} \tilde{v}_{ab}^{ij} = \lim_{\Delta_{ab}^{ij}\to 0} v_{ij}^{ab} v_{ab}^{ij} \frac{1 - e^{-2s(\Delta_{ab}^{ij})^2}}{\Delta_{ab}^{ij}} = 0, \tag{2.15}$$

$$\lim_{\Delta_{ab}^{ij}\to 0} \bar{t}_{ij}^{ab,(1)} \Delta_{ab}^{ij} t_{ab}^{ij,(1)} = \lim_{\Delta_{ab}^{ij}\to 0} v_{ij}^{ab} t_{ab}^{ij,(1)} \left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right] = 0. \tag{2.16}$$

We can now use Eq. (2.14) to remove $\bar{t}_{ij}^{ab,(1)}$ from the Lagrangian [Eq. (2.5)]:

$$\begin{aligned}
\mathcal{L} = & \frac{1}{2} \sum_{ij}^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ij}^{ab} t_{ab}^{ij,(1)} \left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right] \\
& - \frac{1}{4} \sum_{ij}^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} \left(t_{ab}^{ij,(1)}\right)^2 \Delta_{ab}^{ij} \frac{1 + e^{-s(\Delta_{ab}^{ij})^2}}{1 - e^{-s(\Delta_{ab}^{ij})^2}} \\
& + \frac{1}{2} \sum_{pq}^{\mathbf{G}} z_q^p (f_p^q - \delta_p^q \epsilon_p) - \frac{1}{2} \sum_{pq}^{\mathbf{G}} \omega_q^p (S_p^q - \delta_p^q).
\end{aligned} \tag{2.17}$$

Next, we consider the canonical orbitals constraint. Adding orbital energies to the optimization procedure is convenient because once Eq. (2.10) is satisfied, terms of the form $\partial \epsilon_p / \partial C_{\mu p}$ do not contribute to the gradient. Imposing this constraint leads to the following expressions for the diagonal elements of the OPDM:

$$z_I^I = z_A^A = 0, \tag{2.18}$$

$$z_i^i = \sum_{j}^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} (v_{ij}^{ab})^2 \left[4se^{-2s(\Delta_{ab}^{ij})^2} - \frac{1 - e^{-2s(\Delta_{ab}^{ij})^2}}{(\Delta_{ab}^{ij})^2}\right], \tag{2.19}$$

$$z_a^a = -\sum_{ij}^{\mathbf{H}} \sum_{b}^{\mathbf{P}} (v_{ij}^{ab})^2 \left[4se^{-2s(\Delta_{ab}^{ij})^2} - \frac{1 - e^{-2s(\Delta_{ab}^{ij})^2}}{(\Delta_{ab}^{ij})^2}\right]. \tag{2.20}$$

We are now ready to derive the orbital response terms by requiring stationarity with respect to orbital coefficients [Eq. (2.11)]. Here, we follow the approach of Levchenko *et al.*[29] where derivatives $\partial \mathcal{L} / \partial C_{\mu q}$ are contracted with the MO coefficient matrix, i.e., by imposing

$$X_p^q = \sum_{\mu}^{\mathrm{AO}} C_{\mu p}^* \frac{\partial \mathcal{L}}{\partial C_{\mu q}} = 0, \quad \forall pq \in \mathbf{G}. \tag{2.21}$$

Expressions for all the DSRG-PT2 OPDM blocks are reported in Table 2.2. Here we illustrate the approach used to derive these expressions by considering the elements

Table 2.2: The DSRG-PT2 one-particle density matrix elements. Einstein convention of summation over repeated indices is adopted. The big $\mathcal{O}$ notation is assumed for computational cost analysis.

Term	Expression[a]	Cost
z_B^A	0	$N_{\mathbf{FV}}^2$
z_J^I	0	$N_{\mathbf{FC}}^2$
z_A^I	$v_{pA}^{qI} z_q^p / \Delta_A^I$	$N_{\mathbf{G}}^2 N_{\mathbf{FC}} N_{\mathbf{FV}}$
z_I^i	$v_{Ij}^{ab} w_{ab}^{ij} / \Delta_I^i$	$N_{\mathbf{P}}^2 N_{\mathbf{H}}^2 N_{\mathbf{FC}}$
z_a^A	$v_{ij}^{Ab} w_{ab}^{ij} / \Delta_A^a$	$N_{\mathbf{P}}^2 N_{\mathbf{H}}^2 N_{\mathbf{FV}}$
z_a^I	$(v_{pa}^{qI} z_q^p - v_{ij}^{Ib} w_{ab}^{ij}) / \Delta_a^I$	$N_{\mathbf{P}}^2 N_{\mathbf{H}}^2 N_{\mathbf{FC}}$
z_A^i	$(v_{pA}^{qi} z_q^p + v_{Aj}^{ab} w_{ab}^{ij}) / \Delta_A^i$	$N_{\mathbf{P}}^2 N_{\mathbf{H}}^2 N_{\mathbf{FV}}$
z_j^i	$(v_{ab}^{ik} w_{jk}^{ab} + w_{ab}^{ik} v_{jk}^{ab}) / \Delta_j^i$	$N_{\mathbf{H}}^3 N_{\mathbf{P}}^2$
z_b^a	$(v_{ij}^{ac} w_{bc}^{ij} + w_{ij}^{ac} v_{bc}^{ij}) / \Delta_a^b$	$N_{\mathbf{P}}^3 N_{\mathbf{H}}^2$
z_i^a	$(v_{pi}^{qa} z_q^p - v_{cd}^{aj} w_{ij}^{cd} + v_{ib}^{kl} w_{kl}^{ab}) / \Delta_a^i$	$N_{\mathbf{P}}^3 N_{\mathbf{H}}^2$

[a] $w_{pq}^{rs} = v_{pq}^{rs}[1 - e^{-2s(\Delta_{pq}^{rs})^2}] / \Delta_{pq}^{rs}$ and thus $w_{rs}^{pq} = -w_{pq}^{rs}$.

of the **H**-**H** block (z_j^i). Imposing the condition $X_i^j = 0$ leads to the following system of equations

$$\sum_k^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ik}^{ab} t_{ab}^{jk,(1)} \left[1 + e^{-s(\Delta_{ab}^{jk})^2}\right] + z_i^j \epsilon_i + \sum_{pq}^{\mathbf{G}} z_q^p v_{pi}^{qj} - \omega_i^j = 0, \tag{2.22}$$

for $\forall i, j \in \mathbf{H}$. Since both the OPDM and EWDM are symmetric (i.e., $z_p^q = z_q^p$ and $\omega_p^q = \omega_q^p$), we can eliminate the ω_i^j terms by evaluating the difference $X_i^j - X_j^i = 0$ and express z_i^j in closed form as

$$z_i^j = \frac{1}{\Delta_i^j} \sum_k^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ik}^{ab} v_{ab}^{jk} \left[\frac{1 - e^{-2s(\Delta_{ab}^{jk})^2}}{\Delta_{ab}^{jk}} - \frac{1 - e^{-2s(\Delta_{ab}^{ik})^2}}{\Delta_{ab}^{ik}}\right]. \tag{2.23}$$

At first glance, Eq. (2.23) may appear to diverge when orbitals ψ_i and ψ_j have degenerate orbital energies, yet its value is bounded since

$$\lim_{\Delta_i^j \to 0} z_i^j = \sum_k^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ik}^{ab} v_{ab}^{jk} \left[4s e^{-2s(\Delta_{ab}^{ik})^2} - \frac{1 - e^{-2s(\Delta_{ab}^{ik})^2}}{(\Delta_{ab}^{ik})^2}\right]. \tag{2.24}$$

Equation (2.24) is bounded even in the limit of $\Delta_{ab}^{ik} \to 0$, in which case it is equal to $z_i^j = 2s \sum_k^{\mathbf{H}} \sum_{ab}^{\mathbf{P}} v_{ik}^{ab} v_{ab}^{jk}$. Interestingly, the diagonal elements of Eq. (2.24) have expression identical to those of z_i^i as defined in Eq. (2.19). It can be shown that terms in Table 2.2 are universal for both diagonal and off-diagonal elements of the

OPDM. We note that the MP2 density can also be retrieved using equations in Table 2.2 by setting $s \to \infty$. In this limit, the denominators associated with z_j^i and z_b^a can be easily removed. For example, in the case of z_j^i the term in square brackets appearing in Eq. (2.23) may be written as $1/\Delta_{ab}^{jk} - 1/\Delta_{ab}^{ik} = \Delta_j^i/(\Delta_{ab}^{jk}\Delta_{ab}^{ik})$.[75]

Table 2.3: The DSRG-PT2 energy-weighted density matrix elements. Einstein convention of summation over repeated indices is adopted. The big $\mathcal{O}$ notation is assumed for computational cost analysis.

Term	Expression[a]	Cost
ω_B^A	0	$N_{\mathbf{FV}}^2$
ω_I^J	$z_q^p v_{pI}^{qJ}$	$N_{\mathbf{G}}^2 N_{\mathbf{FC}}^2$
ω_I^A	$\epsilon_I z_I^A$	$N_{\mathbf{FC}} N_{\mathbf{FV}}$
ω_i^I	$\epsilon_i z_i^I + z_q^p v_{pi}^{qI}$	$N_{\mathbf{G}}^2 N_{\mathbf{H}} N_{\mathbf{FC}}$
ω_A^a	$\epsilon_a z_A^a$	$N_{\mathbf{P}} N_{\mathbf{FV}}$
ω_a^I	$\epsilon_a z_a^I + z_p^q v_{qa}^{pI}$	$N_{\mathbf{G}}^2 N_{\mathbf{P}} N_{\mathbf{FC}}$
ω_i^A	$\epsilon_i z_i^A$	$N_{\mathbf{H}} N_{\mathbf{FV}}$
ω_j^i	$\epsilon_j z_j^i + z_p^q v_{qj}^{pi} + v_{jk}^{ab} w_{ab}^{ik}$	$N_{\mathbf{H}}^3 N_{\mathbf{P}}^2$
ω_b^a	$\epsilon_a z_b^a + v_{ij}^{ac} w_{bc}^{ij}$	$N_{\mathbf{P}}^3 N_{\mathbf{H}}^2$
ω_i^a	$\epsilon_i z_i^a - v_{ib}^{kl} w_{kl}^{ab}$	$N_{\mathbf{P}}^2 N_{\mathbf{H}}^3$

[a] $w_{pq}^{rs} = v_{pq}^{rs}[1 - e^{-2s(\Delta_{pq}^{rs})^2}]/\Delta_{pq}^{rs}$.

Once the DSRG-PT2 OPDM is determined, we can solve for ω_q^p by imposing $X_q^p = 0$ or $X_p^q = 0$. In Table 2.3, we report the most efficient way to compute the EWDM, implicitly assuming its symmetry ($\omega_p^q = \omega_q^p$). At this point, the DSRG-PT2 Lagrangian is fully specified and the energy gradient [Eq. (2.12)] can be evaluated as:

$$\frac{\mathrm{d}E^{(2)}}{\mathrm{d}x} = \frac{1}{2}\sum_{pq}^{\mathbf{G}} z_q^p \left[(h_p^q)^x + \sum_m^{\mathbf{O}} (v_{pm}^{qm})^x\right] + \frac{1}{2}\sum_{ij}^{\mathbf{H}}\sum_{ab}^{\mathbf{P}} \Gamma_{ab}^{ij}(v_{ij}^{ab})^x - \frac{1}{2}\sum_{pq}^{\mathbf{G}} \omega_q^p (S_p^q)^x, \tag{2.25}$$

where $(h_p^q)^x$, $(S_p^q)^x$, and $(v_{pq}^{rs})^x$ are the skeleton one-electron, overlap, and antisymmetrized two-electron derivative MSO integrals,[29] respectively. In Eq. (2.25), Γ_{ab}^{ij} is the DSRG-PT2 two-particle density matrix given by:

$$\Gamma_{ab}^{ij} = t_{ab}^{ij,(1)}\left[1 + e^{-s(\Delta_{ab}^{ij})^2}\right]. \tag{2.26}$$

Finally, we briefly discuss the computational complexity of the DSRG-PT2 analytic gradients and compare our derivation to MP2 gradient theory. The cost of computing z_q^p and ω_q^p is given in Tables 2.2 and 2.3, respectively. The most expensive step in the computation of the DSRG-PT2 gradient is evaluating the z_i^a and ω_b^a terms, which scales as $\mathcal{O}(N_{\mathbf{P}}^3 N_{\mathbf{H}}^2)$. This cost is found to be *identical* to that of MP2 gradient theory. In comparison to the MP2 densities,[75] the DSRG-PT2 counterparts [Tables 2.2 and 2.3 and Eq. (2.26)] all possess additional exponential factors that arise from the regularizer. In the limit of $s \to \infty$, the DSRG-PT2 densities become identical to the MP2 ones, assuming all denominators differ from zero.

The derivation presented in this section *explicitly* imposes the canonical-orbital constraints in the Lagrangian. This approach can also be used to derive MP2 gradients. A more convenient method to obtain MP2 gradients uses equations that are invariant with respect to separate unitary rotations of occupied and virtual orbitals. In this case, there is no need to impose the canonical-orbital conditions in the Lagrangian and no singularities need to be removed in the unrelaxed OPDM (z_j^i and z_b^a). For DSRG-PT2, the presence of an exponential function $e^{-s(\Delta_{ab}^{ij})^2}$ in the energy expression [Eq. (2.4)] complicates the formulation of orbital-invariant energy and amplitude equations, and as such, enforcing the canonical-orbital constraints in the Lagrangian is preferable. Note that the DSRG-PT2 derivatives can also be obtained without using the method of Lagrange multiplier following the approach used for MP2 gradients.[75] Nevertheless, this route appears to be more complex than the Lagrangian approach.

2.3 Results

We have developed a proof-of-principle implementation of DSRG-PT2 gradient theory as a plugin of Psi4.[76] The correctness of this code was verified by two internal tests using H_2O and NH_3. First, we verified the equivalence of the electronic dipole

moments and equilibrium geometries between MP2 and DSRG-PT2 when $s \geq 10^7$ E_h^{-2}. Second, the analytic DSRG-PT2 energy gradients were tested against those from five-point finite-difference calculations using a 0.005 a.u. step size. All components of the analytic and finite-difference gradients differed by less than 10^{-10} a.u.

To illustrate the performance of the DSRG-PT2 gradients, we computed the equilibrium structures of fifteen small molecules, taken from Ref.[31] These molecules include H_2, N_2, HF, CO, CO_2, H_2O, HOF, H_2O_2, HNC, HCN, NH_3, CH_4, C_2H_2, C_2H_4 and CH_2O. The DSRG-PT2 optimized geometries were compared against those obtained from MP2, CCSD,[73] and CCSD(T).[74] Analytic gradients were employed for MP2 and CCSD, while the CCSD(T) gradients were obtained using a five-point finite-difference formula. For geometry optimizations, we used a set of tight convergence criteria where the maximum force and displacement between two consecutive steps differ by no more than 2×10^{-6} and 6×10^{-6} a.u., respectively. In all computations, we adopted the cc-pVQZ basis set[77] and the MOs mainly built from 1s orbitals of the second-row atoms were excluded for post-Hartree–Fock treatment. The flow parameter of DSRG-PT2 was set to 1.0 E_h^{-2}. The MP2, CCSD and CCSD(T) geometries were optimized using Molpro 2015.1.[78,79]

Tables 2.4 and 2.5 report bond length and bond angle deviations for DSRG-PT2, MP2, and CCSD from the CCSD(T) values. The geometries optimized using DSRG-PT2 and MP2 yield almost identical mean absolute deviations, 0.0033 and 0.0035 Å, respectively. This result is expected for the following reasons: i) both methods are based on the second-order perturbation theory, and ii) the molecules considered here are closed shells and the magnitude of the Møller–Plesset denominators is sufficiently large that regularization has a minor effect on the energy. Interestingly, the CCSD results show slightly worse statistics than those of DSRG-PT2, as previously also observed for MP2.[80]

Table 2.4: Equilibrium bond lengths (in Ångström) of DSRG-PT2, MP2, and CCSD relative to those of CCSD(T). The absolute values of CCSD(T) are shown in the rightmost column. The flow parameter of DSRG-PT2 is set to 1.0 E_h^{-2}. All computations employ the cc-pVQZ basis set and the frozen-core approximation.

Molecule	Bond	DSRG-PT2	MP2	CCSD	CCSD(T)
H_2	H−H	-0.0059	-0.0058	-0.0000	0.7419
HF	F−H	0.0009	0.0009	-0.0025	0.9162
H_2O	O−H	-0.0002	-0.0002	-0.0026	0.9579
HOF	O−H	-0.0002	-0.0002	-0.0035	0.9665
H_2O_2	O−H	0.0004	0.0004	-0.0033	0.9627
HNC	N−H	-0.0004	-0.0003	-0.0026	0.9961
NH_3	N−H	-0.0027	-0.0026	-0.0024	1.0124
C_2H_2	C−H	-0.0023	-0.0023	-0.0018	1.0634
HCN	C−H	-0.0027	-0.0026	-0.0020	1.0668
C_2H_4	C−H	-0.0029	-0.0028	-0.0017	1.0823
CH_4	C−H	-0.0039	-0.0038	-0.0014	1.0879
CH_2O	C−H	-0.0029	-0.0030	-0.0022	1.1022
N_2	N−N	0.0087	0.0101	-0.0072	1.1003
CO	C−O	0.0025	0.0032	-0.0071	1.1314
CO_2	C−O	0.0031	0.0036	-0.0072	1.1626
CH_2O	C−O	0.0003	0.0016	-0.0068	1.2066
HCN	C−N	0.0042	0.0070	-0.0074	1.1564
HNC	C−N	-0.0004	0.0011	-0.0072	1.1720
C_2H_2	C−C	-0.0016	0.0021	-0.0065	1.2065
C_2H_4	C−C	-0.0073	-0.0047	-0.0060	1.3343
HOF	O−F	-0.0129	-0.0121	-0.0198	1.4347
H_2O_2	O−O	-0.0069	-0.0061	-0.0182	1.4525
Mean absolute deviation		0.0033	0.0035	0.0054	
Standard deviation		0.0045	0.0047	0.0050	

Table 2.5: Equilibrium bond angles (in degrees) of DSRG-PT2, MP2, and CCSD relative to those of CCSD(T). The absolute values of CCSD(T) are shown in the rightmost column. The flow parameter of DSRG-PT2 is set to 1.0 E_h^{-2}. All computations employ the cc-pVQZ basis set and the frozen-core approximation.

Molecule	Bond angle	DSRG-PT2	MP2	CCSD	CCSD(T)
H_2O	H−O−H	-0.09	-0.10	0.28	104.12
HOF	H−O−F	0.16	0.14	0.76	97.78
H_2O_2	H−O−O	-0.22	-0.25	0.73	99.91
NH_3	H−N−H	0.29	0.29	0.28	106.18
C_2H_4	H−C−H	0.20	0.23	-0.12	117.12
CH_2O	H−C−H	-0.04	-0.00	-0.06	116.44

2.4 Conclusions

We have presented a derivation of the DSRG-PT2 analytic energy gradients using the method of Lagrange multipliers that includes canonical orbital constraints, as required by the formulation of the DSRG-PT2 method. Importantly, this additional constraint does not introduce any difficulty in the derivation of the DSRG-PT2 one-particle density matrix and the energy-weighted density matrix. The DSRG-PT2 Lagrangian is also found to be bounded for small energy denominators. The evaluation of DSRG-PT2 gradients has essentially the same computational cost of MP2 gradients, since additional terms due to the regularizer have a negligible cost.

We have employed the DSRG-PT2 analytic gradients to compute the equilibrium structures of fifteen small molecules. The DSRG-PT2 optimized geometries are in excellent agreement with those obtained by MP2. This observation indicates that $s = 1.0\ E_\mathrm{h}^{-2}$ is sufficiently large for DSRG-PT2 to recover the most of the electron correlation captured by MP2. In general, both DSRG-PT2 and MP2 theories yield a good balance between accuracy and cost for optimizing molecular structures, and they are in fortuitously good agreement with CCSD(T) for the molecules considered in this work.

Based on the current work, we envision the following challenges for developing gradient theory for the multireference DSRG-PT2 (DSRG-MRPT2). Obviously, the multireference partitioning of non-frozen orbitals into core, active, and virtual sets increases the number of orbital response equations. Perhaps, the major difficulty is to implement an efficient algorithm for optimizing the DSRG-MRPT2 Lagrangian with respect to the reference wave function. This challenge results from the fact that the DSRG-MRPT2 energy is written in terms of the one-, two-, and three-particle density matrices of the reference. The derivatives of the higher-order reduced density matrices introduce a memory bottleneck that may require the use of a batching algorithm, as is done in the context of CASPT2.[81] Additional complications may also come from the

DSRG-MRPT2 amplitude equations where singles depend on semi-internal doubles.

There are several other directions worth explorations. The simple structure of the DSRG-PT2 gradient equations suggests that energy derivatives of the nonperturbative version of single-reference DSRG theory may be easily attained.[67] Note that, as any other regularized MP2 approach, the DSRG-PT2 energy lies in between the Hartree–Fock and MP2 energies. In particular, for systems with one or more small denominators, the DSRG-PT2 method will ignore large unphysical contributions to the correlation energy due to the presence of a regularizing function. Thus, the DSRG-PT2 in itself is not applicable to systems that require a multireference treatment. However, the DSRG-PT2 may be useful in other contexts where the performance of MP2 can be improved by regularizations, like in the case of dispersion interactions.[60,82] It would also be interesting to see how an orbital-optimized version of DSRG-PT2 may perform against other regularized OO-MP2 methods recently proposed by Lee and Head-Gordon.[9] In the near future, we plan to optimize the value of the flow parameter s by developing test sets in which static linear response properties are considered along with energies.

2.5 Acknowledgments

This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004 and a Camille Dreyfus Teacher-Scholar Award. C.L. would like to acknowledge Dr. Justin M. Turney and Dr. Alexander Yu. Sokolov for helpful discussions on implementing gradient theory using Psi4.

Bibliography

[1] Doser, B.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Linear-scaling atomic orbital-based second-order Møller–Plesset perturbation theory by rigorous integral screening criteria. *J. Chem. Phys.* **2009**, *130*, 064107.

[2] Maurer, S. A.; Lambrecht, D. S.; Kussmann, J.; Ochsenfeld, C. Efficient distance-including integral screening in linear-scaling Møller-Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 014101.

[3] Grimme, S. Improved second-order Møller–Plesset perturbation theory by separate scaling of parallel- and antiparallel-spin pair correlation energies. *J. Chem. Phys.* **2003**, *118*, 9095.

[4] Jung, Y.; Lochan, R. C.; Dutoi, A. D.; Head-Gordon, M. Scaled opposite-spin second order Møller–Plesset correlation energy: An economical electronic structure method. *J. Chem. Phys.* **2004**, *121*, 9793–9802.

[5] Distasio JR, R. A.; Head-Gordon, M. Optimized spin-component scaled second-order Møller-Plesset perturbation theory for intermolecular interaction energies. *Mol. Phys.* **2007**, *105*, 1073–1083.

[6] Lochan, R. C.; Head-Gordon, M. Orbital-optimized opposite-spin scaled second-order correlation: An economical method to improve the description of open-shell molecules. *J. Chem. Phys.* **2007**, *126*, 164101.

[7] Neese, F.; Schwabe, T.; Kossmann, S.; Schirmer, B.; Grimme, S. Assessment of Orbital-Optimized, Spin-Component Scaled Second-Order Many-Body Perturbation Theory for Thermochemistry and Kinetics. *J. Chem. Theory Comput.* **2009**, *5*, 3060–3073.

[8] Bozkaya, U. Orbital-optimized third-order Møller-Plesset perturbation theory and its spin-component and spin-opposite scaled variants: Application to symmetry breaking problems. *J. Chem. Phys.* **2011**, *135*, 224103.

[9] Lee, J.; Head-Gordon, M. Regularized Orbital-Optimized Second-Order Møller-Plesset Perturbation Theory: A Reliable Fifth-Order-Scaling Electron Corre-

lation Model with Orbital Energy Dependent Regularizers. *J. Chem. Theory Comput.* **2018**, *14*, 5203–5219.

[10] Lee, J.; Head-Gordon, M. Two Single-Reference Approaches to Singlet Biradicaloid Problems: Complex, Restricted Orbitals and Approximate Spin-Projection Combined With Regularized Orbital-Optimized Møller-Plesset Perturbation Theory. *arXiv* **2019**,

[11] Pople, J. A.; Krishnan, R.; Schlegel, H. B.; Binkley, J. S. Derivative Studies in Hartree–Fock and Møller–Plesset Theories. *Int. J. Quantum Chem.* **1979**, *16*, 225–241.

[12] Pulay, P.; Saeb, S. Orbital-invariant formulation and second-order gradient evaluation in Møller–Plesset perturbation theory. *Theor. Chim. Acta* **1986**, *69*, 357–368.

[13] Jørgensen, P.; Helgaker, T. Møller–Plesset energy derivatives. *J. Chem. Phys.* **1988**, *89*, 1560.

[14] Frisch, M. J.; Head-Gordon, M.; Pople, J. A. A direct MP2 gradient method. *Chem. Phys. Lett.* **1990**, *166*, 275–280.

[15] DiStasio, R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An improved algorithm for analytical gradient evaluation in resolution-of-the-identity second-order Møller-Plesset perturbation theory: application to alanine tetrapeptide conformational analysis. *J. Comput. Chem.* **2007**, *28*, 839–856.

[16] Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the orbital-optimized second-order Møller–Plesset perturbation theory. *J. Chem. Phys.* **2013**, *138*, 184103.

[17] Song, C.; Martínez, T. J. Analytical gradients for tensor hyper-contracted MP2 and SOS-MP2 on graphical processing units. *J. Chem. Phys.* **2017**, *147*, 161723.

[18] Gerratt, J.; Mills, I. M. Force Constants and Dipole-Moment Derivatives of Molecules from Perturbed Hartree–Fock Calculations. I. *J. Chem. Phys.* **1968**, *49*, 1719–1729.

[19] Pulay, P. In *Applications of Electronic Structure Theory*; Schaefer, H. F., Ed.; Springer US: Boston, MA, 1977; pp 153–185.

[20] Krishnan, R.; Schlegel, H. B.; Pople, J. A. Derivative studies in configuration–interaction theory. *J. Chem. Phys.* **1980**, *72*, 4654–4655.

[21] Handy, N. C.; Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[22] Salter, E. A.; Trucks, G. W.; Bartlett, R. J. Analytic energy derivatives in many-body methods. I. First derivatives. *J. Chem. Phys.* **1989**, *90*, 1752–1766.

[23] Foresman, J. B.; Head-Gordon, M.; Pople, J. A.; Frisch, M. J. Toward a systematic molecular orbital theory for excited states. *J. Phys. Chem.* **1992**, *96*, 135–149.

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

[25] (a) Van Caillie, C.; Amos, R. D. Geometric derivatives of excitation energies using SCF and DFT. *Chem. Phys. Lett.* **1999**, *308*, 249–255; (b) Van Caillie, C.; Amos, R. D. Geometric derivatives of density functional theory excitation energies using gradient-corrected functionals. *Chem. Phys. Lett.* **2000**, *317*, 159–164.

[26] Gauss, J.; Stanton, J. F. Analytic gradients for the coupled-cluster singles, doubles, and triples (CCSDT) model. *J. Chem. Phys.* **2002**, *116*, 1773–1782.

[27] Köhn, A.; Hättig, C. Analytic gradients for excited states in the coupled-cluster model CC2 employing the resolution-of-the-identity approximation. *J. Chem. Phys.* **2003**, *119*, 5021–5036.

[28] Kállay, M.; Gauss, J.; Szalay, P. G. Analytic first derivatives for general coupled-cluster and configuration interaction models. *J. Chem. Phys.* **2003**, *119*, 2991–3004.

[29] Levchenko, S. V.; Wang, T.; Krylov, A. I. Analytic gradients for the spin-conserving and spin-flipping equation-of-motion coupled-cluster models with single and double substitutions. *J. Chem. Phys.* **2005**, *122*, 224106.

[30] Rappoport, D.; Furche, F. Lagrangian approach to molecular vibrational Raman intensities using time-dependent hybrid density functional theory. *J. Chem. Phys.* **2007**, *126*, 201104.

[31] Sokolov, A. Y.; Wilke, J. J.; Simmonett, A. C.; Schaefer, H. F. Analytic gradients for density cumulant functional theory: the DCFT-06 model. *J. Chem. Phys.* **2012**, *137*, 054105.

[32] Zhang, D.; Peng, D.; Zhang, P.; Yang, W. Analytic gradients, geometry optimization and excited state potential energy surfaces from the particle-particle random phase approximation. *Phys. Chem. Chem. Phys.* **2015**, *17*, 1025–1038.

[33] Schutz, M. Oscillator strengths, first-order properties, and nuclear gradients for local ADC(2). *J. Chem. Phys.* **2015**, *142*, 214103.

[34] Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the coupled-cluster

singles and doubles with perturbative triples method with the density-fitting approximation. *J. Chem. Phys.* **2017**, *147*, 044104.

[35] Győrffy, W.; Werner, H.-J. Analytical energy gradients for explicitly correlated wave functions. II. Explicitly correlated coupled cluster singles and doubles with perturbative triples corrections: CCSD(T)-F12. *J. Chem. Phys.* **2018**, *148*, 114104.

[36] Hirao, K. Multireference Møller—Plesset method. *Chem. Phys. Lett.* **1992**, *190*, 374–380.

[37] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

[38] Hoffmann, M. R. Canonical Van Vleck Quasidegenerate Perturbation Theory with Trigonometric Variables. *J. Phys. Chem.* **1996**, *100*, 6125–6130.

[39] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J. P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

[40] Chaudhuri, R. K.; Freed, K. F.; Hose, G.; Piecuch, P.; Kowalski, K.; Włoch, M.; Chattopadhyay, S.; Mukherjee, D.; Rolik, Z.; Szabados, Á.; Tóth, G.; Surján, P. R. Comparison of low-order multireference many-body perturbation theories. *J. Chem. Phys.* **2005**, *122*, 134105.

[41] Hoffmann, M. R.; Datta, D.; Das, S.; Mukherjee, D.; Szabados, Á.; Rolik, Z.; Surján, P. R. Comparative study of multireference perturbative theories for ground and excited states. *J. Chem. Phys.* **2009**, *131*, 204104.

[42] Evangelista, F. A.; Simmonett, A. C.; Schaefer, H. F.; Mukherjee, D.; Allen, W. D. A companion perturbation theory for state-specific multireference coupled cluster methods. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4728–4741.

[43] Sinha Ray, S.; Ghosh, P.; Chaudhuri, R. K.; Chattopadhyay, S. Improved virtual orbitals in state specific multireference perturbation theory for prototypes of quasidegenerate electronic structure. *J. Chem. Phys.* **2017**, *146*, 064111.

[44] Giner, E.; Angeli, C.; Garniron, Y.; Scemama, A.; Malrieu, J.-P. A Jeziorski-Monkhorst fully uncontracted multi-reference perturbative treatment. I. Principles, second-order versions, and tests on ground state potential energy curves. *J. Chem. Phys.* **2017**, *146*, 224108.

[45] Shepard, R.; Lischka, H.; Szalay, P. G.; Kovar, T.; Ernzerhof, M. A general multireference configuration interaction gradient program. *J. Chem. Phys.* **1992**, *96*, 2085–2098.

[46] Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications. *Chem. Rev.* **2012**, *112*, 108–181.

[47] Nakano, H.; Hirao, K.; Gordon, M. S. Analytic energy gradients for multiconfigurational self-consistent field second-order quasidegenerate perturbation theory (MC-QDPT). *J. Chem. Phys.* **1998**, *108*, 5660–5669.

[48] Celani, P.; Werner, H.-J. Analytical energy gradients for internally contracted second-order multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 5044–5057.

[49] Theis, D.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 energy gradient using a Lagrangian formulation. *J. Chem. Phys.* **2011**, *135*, 044117.

[50] Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **2011**, *135*, 081106.

[51] Győrffy, W.; Shiozaki, T.; Knizia, G.; Werner, H.-J. Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* **2013**, *138*, 104104.

[52] MacLeod, M. K.; Shiozaki, T. Communication: automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory. *J. Chem. Phys.* **2015**, *142*, 051103.

[53] Prochnow, E.; Evangelista, F. A.; Schaefer, H. F.; Allen, W. D.; Gauss, J. Analytic gradients for the state-specific multireference coupled cluster singles and doubles model. *J. Chem. Phys.* **2009**, *131*, 064109.

[54] Jagau, T.-C.; Prochnow, E.; Evangelista, F. A.; Gauss, J. Analytic gradients for Mukherjee's multireference coupled-cluster method using two-configurational self-consistent-field orbitals. *J. Chem. Phys.* **2010**, *132*, 144110.

[55] Evangelista, F. A.; Gauss, J. An orbital-invariant internally contracted multireference coupled cluster approach. *J. Chem. Phys.* **2011**, *134*, 114102.

[56] Hanauer, M.; Köhn, A. Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly. *J. Chem. Phys.* **2011**, *134*, 204111.

[57] Roos, B. O.; Andersson, K. Multiconfigurational perturbation theory with level shift — the Cr_2 potential revisited. *Chem. Phys. Lett.* **1995**, *245*, 215–223.

[58] Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.* **1997**, *274*, 196–204.

[59] Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. Intruder state avoidance multireference Møller-Plesset perturbation theory. *J. Comput. Chem.* **2002**, *23*, 957–965.

[60] Ohnishi, Y.-y.; Ishimura, K.; Ten-no, S. Interaction Energy of Large Molecules from Restrained Denominator MP2-F12. *J. Chem. Theory Comput.* **2014**, *10*, 4857–4861.

[61] (a) Li, C.; Evangelista, F. A. Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators. *J. Chem. Phys.* **2016**, *144*, 164114; (b) Li, C.; Evangelista, F. A. Erratum: "Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators" [J. Chem. Phys. 144, 164114 (2016)]. *J. Chem. Phys.* **2018**, *148*, 079903.

[62] Li, C.; Evangelista, F. A. Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245–273.

[63] Li, C.; Evangelista, F. A. Multireference Driven Similarity Renormalization Group: A Second-Order Perturbative Analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.

[64] Li, C.; Verma, P.; Hannon, K. P.; Evangelista, F. A. A low-cost approach to electronic excitation energies based on the driven similarity renormalization group. *J. Chem. Phys.* **2017**, *147*, 074107.

[65] (a) Li, C.; Evangelista, F. A. Driven similarity renormalization group: Third-order multireference perturbation theory. *J. Chem. Phys.* **2017**, *146*, 124132; (b) Li, C.; Evangelista, F. A. Erratum: "Driven similarity renormalization group:

Third-order multireference perturbation theory" [J. Chem. Phys. 146, 124132 (2017)]. *J. Chem. Phys.* **2018**, *148*, 079902.

[66] Li, C.; Evangelista, F. A. Driven similarity renormalization group for excited states: A state-averaged perturbation theory. *J. Chem. Phys.* **2018**, *148*, 124106.

[67] Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.

[68] Hergert, H.; Bogner, S. K.; Morris, T. D.; Schwenk, A.; Tsukiyama, K. The In-Medium Similarity Renormalization Group: A novel ab initio method for nuclei. *Phys. Rep.* **2016**, *621*, 165–222.

[69] Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; John Wiley & Sons, 2000.

[70] Helgaker, T. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 1998; pp 1157–1169.

[71] Hohenstein, E. G. Analytic formulation of derivative coupling vectors for complete active space configuration interaction wavefunctions with floating occupation molecular orbitals. *J. Chem. Phys.* **2016**, *145*, 174110.

[72] Park, J. W.; Al-Saadon, R.; Strand, N. E.; Shiozaki, T. Imaginary Shift in CASPT2 Nuclear Gradient and Derivative Coupling Theory. *J. Chem. Theory Comput.* **2019**, *15*, 4088–4098.

[73] Purvis, G. D.; Bartlett, R. J. A full coupled-cluster singles and doubles model: The inclusion of disconnected triples. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[74] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. A fifth-order perturbation comparison of electron correlation theories. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

[75] Aikens, C. M.; Webb, S. P.; Bell, R. L.; Fletcher, G. D.; Schmidt, M. W.; Gordon, M. S. A derivation of the frozen-orbital unrestricted open-shell and restricted closed-shell second-order perturbation theory analytic gradient expressions. *Theor. Chem. Acc.* **2003**, *110*, 233–253.

[76] Parrish, R. M. et al. Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability. *J. Chem. Theory Comput.* **2017**, *13*, 3185–3197.

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

[78] Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schuetz, M. Molpro: a general-purpose quantum chemistry program package. *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2012**, *2*, 242–253.

[79] Werner, H.-J. et al. MOLPRO, version 2015.1, a package of *ab initio* programs. 2015; see http://www.molpro.net.

[80] Bak, K. L.; Gauss, J.; Jørgensen, P.; Olsen, J.; Helgaker, T.; Stanton, J. F. The accurate determination of molecular equilibrium structures. *J. Chem. Phys.* **2001**, *114*, 6548–6556.

[81] Vlaisavljevich, B.; Shiozaki, T. Nuclear Energy Gradients for Internally Contracted Complete Active Space Second-Order Perturbation Theory: Multistate Extensions. *J. Chem. Theory Comput.* **2016**, *12*, 3781–3787.

[82] Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D. Estimates of the Ab Initio Limit for π-π Interactions: The Benzene Dimer. *J. Am. Chem. Soc.* **2002**, *124*, 10887–10893.

Chapter 3 Analytic Gradients for DSRG-MRPT2

Chapter Abstract

We derive analytic energy gradients of the driven similarity renormalization group (DSRG) multireference second-order perturbation theory (MRPT2) using the method of Lagrange multipliers. In the Lagrangian, we impose constraints for a complete-active-space self-consistent-field reference wave function and the semicanonical orthonormal molecular orbitals. Solving the associated Lagrange multipliers is found to share the same asymptotic scaling of a single DSRG-MRPT2 energy computation. A pilot implementation of the DSRG-MRPT2 analytic gradients is used to optimize the geometry of the singlet and triplet states of *p*-benzyne. The equilibrium bond lengths and angles are similar to those computed via other MRPT2s and Mukherjee's multireference coupled cluster theory. An approximate DSRG-MRPT2 method that neglects the contributions of three-body density cumulant is found to introduce negligible errors in the geometry of *p*-benzyne, lending itself to a promising low-cost approach for molecular geometry optimizations using large active spaces.

3.1 introduction

Analytic energy derivatives play a central role in modern quantum chemistry.[1] They enable efficient geometry optimizations and *ab initio* (including non-adiabatic) molecular dynamics simulations,[2–4] two tasks that require rapid evaluation of energy gradients with respect to nuclear coordinates. Recent developments of analytic gradients for local correlation methods have extended first principles geometry optimiza-

tions to weakly correlated molecules with hundreds of nuclear degrees of freedom.[5–7] For strongly correlated systems (e.g., diradicals and transition-metal complexes), multireference (MR) methods[8–11] are generally necessary to obtain accurate global potential energy surfaces (PESs). Unfortunately, the development of analytic energy gradients for multireference theories has trailed that of single-reference methods both in terms of the underlying mathematical formalism and the broad availability of software implementations, limiting studies of strongly correlated systems.

Within the domain of MR methods, the sweet spot between accuracy and computational cost is found in second-order perturbation theory (MRPT2). Various MRPT2 methods have been proposed over the years,[12–25] among which the most widely applied are the complete-active-space (CAS) second-order perturbation theory (CASPT2)[14] and n-electron valence second-order perturbation theory (NEVPT2).[18] The CASPT2 scheme based on a single CAS configuration interaction (CI) state is known to suffering from the intruder-state problem. This issue is commonly addressed by applying level shifts to the diagonal elements of the one-body zeroth-order Hamiltonian.[26,27] A different, parameter-free approach is used in NEVPT2 to deal with intruder states, whereby the zeroth-order Hamiltonian is augmented with bi-electronic terms, as proposed by Dyall.[28] Nonetheless, both CASPT2 and NEVPT2 in principle require the four-body reduced density matrix (4-RDM) of the CASCI wave function that are both costly to compute and store in memory. Numerous efforts have been made to reduce the cost of high-order density matrices. For example, building and storing the 4-RDM can be avoided using a cumulant decomposition[29–35] and subsequently neglecting contributions from the 4-body density cumulant. This approach has lent itself to efficient and robust implementations of CASPT2 that can handle up to thirty active orbitals.[36,37] Such approximations are less successful in NEVPT2 and "false intruders" may appear due to the density dependencies of the Koopman's matrices in the energy denominators.[38,39] For a similar reason, the use of Cholesky decomposed

integrals may also destabilize the numerical robustness of the NEVPT2 method.[40] For certain formulations of MRPT2, it is possible to avoid computing the 4-RDM by introducing appropriate intermediates, employing an uncontracted formalism, or a matrix product state reference.[24,41–45]

The developments of analytic energy gradients for MRPT2s were largely overlooked for a long time. The very first derivation were reported by Nakano and coworkers in 1998,[46] on the multi-configurational quasi-degenerate perturbation theory (MC-QDPT).[47] However, applications of MC-QDPT gradient theory were restricted to small systems[48] until the recent work of Park that employs the analytic gradient theory of the extended MC-QDPT[49] to optimize the conical intersections of a retinal model chromophore.[50,51] The analytic first derivatives have also been developed for Werner's partially contracted CASPT2[52] and Hoffmann's generalized Van Vleck perturbation theory,[53,54] along with their extensions for excited states.[55–57]

More recently, significant advances have been made in developing the analytic gradients for CASPT2 and NEVPT2. The analytic gradients for the fully internally contracted CASPT2 were first achieved by MacLeod and Shiozaki via automatic code generation.[58] Multi-state generalizations of CASPT2 have also been derived by Shiozaki and co-workers,[59–61] and made publicly available through the BAGEL package.[62] Song, Martínez, and Neaton developed the analytic gradients for the reduced scaling CASPT2 based on supporting subspace method.[63,64] Gradient theory for NEVPT2 was introduced independently by Park[65,66] and Nishimoto.[67,68] Analytic gradients for the restricted active space second-order perturbation theory have also been developed by Nishimoto.[69]

The driven similarity renormalization group (DSRG) provides an alternative framework to formulate MR theories that avoid the intruder-state problem and yield smooth PESs.[70,71] In the DSRG, the many-body Hamiltonian is unitarily transformed in such a way that interactions that couple the reference state and the excited configurations

are zeroed (this is equivalent to a unitary internally contracted theory). Importantly, this decoupling depends on the magnitude of the energy denominator of each interaction removed, and it is gradually suppressed when a denominator approaches zero. This feature of the DSRG introduces a separation of energy scales, the extent of which is controlled via the so-called flow parameter s. For finite values of s, the MR-DSRG methods yield continuous potential energy surfaces that are free from the characteristic "spikes" caused by intruder states.

Over the past few years, we have proposed and implemented several practical MR-DSRG ansätze.[72–75] The least computational demanding member of this family is the DSRG-MRPT2 method.[74] This approach uses a diagonal normal-ordered Fock operator as the zeroth-order Hamiltonian. As a result, the DSRG-MRPT2 energy only depends on the reference 1-, 2-, and 3-RDMs. Previous benchmarks on small molecules show that the DSRG-MRPT2 approach yields PESs of similar accuracy to other MRPT2s.[72,76] When combined with factorization of the two-electron integrals, DSRG-MRPT2 be routinely applied to systems with more than two thousand basis functions.[77] The DSRG-MRPT2 approach has also been combined with approximate CASCI methods to target large active spaces.[78,79] These encouraging results motivate us to further extend its applicability.

Herein, we report a pilot implementation of the analytic energy gradients for the state-specific unrelaxed DSRG-MRPT2 method.[74] The DSRG-MRPT2 energy is not variationally optimized with respect to the orbital and CI coefficients, nor the cluster amplitudes. Following a standard approach,[80,81] we construct a Lagrangian function ($\mathcal{L}$) and incorporate constraints for non-variational quantities. As anticipated from previous experiences,[52,58,82] the computational bottleneck of the gradient procedure is solving the coupled Z-vector equations[83,84] for the orbital and CI coefficients. Due to the complexity of the DSRG-MRPT2 analytic gradients, in this work we restrict our derivation to the original unrelaxed approach.[74] Variants of the DSRG-MRPT2

that include reference relaxation[72] will be considered in future works.

This paper is organized as follows. We start by introducing the DSRG-MRPT2 energy expressions and the amplitude equations in Sec. 3.2.1, followed by a general discussion of gradient theory using the method of Lagrange multipliers in Sec. 3.2.2. We report expressions for all the constraints and the corresponding Lagrange multipliers in Sec. 3.2.3 and 3.2.4, respectively. The theory section is concluded with a brief discussion on the computational cost and limitations of the current implementation (see Sec. 3.2.5). In Sec. 3.3, we report the adiabatic singlet–triplet splittings of *p*-benzyne computed from the DSRG-MRPT2 optimized geometries using analytic gradients. Finally, we conclude this work in Sec. 3.4 by pointing out several future directions and applications of DSRG-MRPT2 gradient theory.

3.2 Theory

We first introduce the orbital notation adopted in this work. Consider a set of CASSCF orthonormal molecular spin orbitals (MSOs) $\mathbb{G} \equiv \{\psi_p(\mathbf{r},\omega) = \phi_p(\mathbf{r})\sigma_p(\omega), p = 1,2,\ldots,N_\mathrm{G}\}$. Each MSO is a product of a molecular orbital (MO) $\phi_p(\mathbf{r})$ and a spin function $\sigma_p(\omega)$, and the spatial and spin coordinates are indicated with $\mathbf{r}$ and ω, respectively. An MO is a linear combination of nonorthogonal atomic orbitals (AOs) $\chi_\mu(\mathbf{r})$:

$$\phi_p(\mathbf{r}) = \sum_\mu^{\mathrm{AO}} \chi_\mu(\mathbf{r})\, C_{\mu p}, \tag{3.1}$$

where $C_{\mu p}$ is the orbital coefficient matrix. The MSOs are assumed to be orthonormal, in which case the MSO overlap integral (S_p^q) is an identity matrix:

$$S_p^q = \langle \psi_p | \psi_q \rangle = \delta_p^q, \tag{3.2}$$

where δ_p^q is the Kronecker delta. We partition the MSOs into three subsets: core ($\mathbb{C}$, doubly occupied), active ($\mathbb{A}$, partially occupied), and virtual ($\mathbb{V}$, unoccupied). For convenience, we also introduce composite orbital spaces, namely, hole ($\mathbb{H} = \mathbb{C} \cup \mathbb{A}$)

and particle ($\mathbb{P} = \mathbb{A} \cup \mathbb{V}$). The indices labeling MSOs are summarized in Table 3.1, and Greek letters μ, ν, ρ, τ are utilized to index AOs.

Table 3.1: Partition of the spin orbital spaces.

Space	Symbol	Size	Indices	Description
Core	$\mathbb{C}$	N_C	m, n, o	Occupied
Active	$\mathbb{A}$	N_A	u, v, w, x, y, z	Partially occupied
Virtual	$\mathbb{V}$	N_V	e, f	Unoccupied
Hole	$\mathbb{H}$	N_H	i, j, k, l	$\mathbb{C} \cup \mathbb{A}$
Particle	$\mathbb{P}$	N_P	a, b, c, d	$\mathbb{A} \cup \mathbb{V}$
General	$\mathbb{G}$	N_G	p, q, r, s	$\mathbb{C} \cup \mathbb{A} \cup \mathbb{V}$

The CASSCF reference wave function Ψ_0 (often referred to as the "reference" in the following) is a linear combination of Slater determinants Φ_I:

$$|\Psi_0\rangle = \sum_I^{\mathcal{M}_0} c_I \, |\Phi_I\rangle \,, \tag{3.3}$$

with c_I being the vector of CI coefficients. These determinants form a complete active space (CAS) denoted by $\mathcal{M}_0$. Any $\Phi_I \in \mathcal{M}_0$ can be expressed as

$$|\Phi_I\rangle = \hat{\mathcal{I}}^\dagger \prod_m^{\mathbb{C}} \hat{a}_m^\dagger \, |-\rangle \,, \tag{3.4}$$

where $|-\rangle$ is the true vacuum and $\hat{a}_p^\dagger$ ($\hat{a}_p$) is a fermionic creation (annihilation) operator. In Eq. (3.4), the operator $\hat{\mathcal{I}}^\dagger = \hat{a}_u^\dagger \hat{a}_v^\dagger \cdots$ is one of the $|\mathcal{M}_0|$ choices of creating n_a electrons in the N_A active orbitals in such a way that Φ_I has the desired spin and spatial symmetry. In the following, we use capital letters I and J to label the index of determinants in $\mathcal{M}_0$.

It is convenient to express the properties of the CASSCF reference in terms of general n-particle reduced density matrices (n-pRDMs), with nonzero elements defined as

$$\gamma_{ij\cdots}^{kl\cdots} = \langle \Psi_0 | \underbrace{\hat{a}_k^\dagger \hat{a}_l^\dagger \cdots}_{n \text{ operators}} \underbrace{\cdots \hat{a}_j \hat{a}_i}_{n \text{ operators}} | \Psi_0 \rangle \,. \tag{3.5}$$

For example, the reference energy $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ may be expressed in terms of the 1- and 2-pRDMs (γ_v^u and γ_{xy}^{uv}), and the one-electron (h_p^q) and antisymmetrized

two-electron (v^{rs}_{pq}) integrals. In particular, we have

$$E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle = E_0^{\mathrm{c}} + E_0^{\mathrm{a}}, \tag{3.6}$$

where the core (E_0^{c}) and active (E_0^{a}) parts of the energy are defined as:

$$E_0^{\mathrm{c}} = \sum_{m}^{\mathbb{C}} h^m_m + \frac{1}{2} \sum_{mn}^{\mathbb{C}} v^{mn}_{mn}, \tag{3.7}$$

$$E_0^{\mathrm{a}} = \sum_{uv}^{\mathbb{A}} \bar{f}^v_u \gamma^u_v + \frac{1}{4} \sum_{uvxy}^{\mathbb{A}} v^{xy}_{uv} \gamma^{uv}_{xy}. \tag{3.8}$$

In Eq. (3.8), we have introduced the core Fock matrix ($\bar{f}^q_p$):

$$\bar{f}^q_p = h^q_p + \sum_{m}^{\mathbb{C}} v^{qm}_{pm}. \tag{3.9}$$

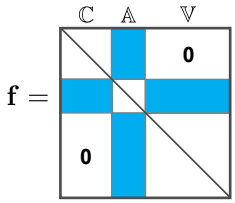

Figure 3.1: Generalized Fock matrix in the semicanonical CASSCF basis. The blocks colored in blue are dense while the three diagonal blocks contain only diagonal elements.

We further define a generalized Fock matrix of the reference with components f^p_q given by:

$$f^p_q = \bar{f}^p_q + \sum_{uv}^{\mathbb{A}} v^{pu}_{qv} \gamma^v_u. \tag{3.10}$$

The DSRG-MRPT2 method is formulated in the semicanonical orbital basis, such that the core, active, and virtual blocks of the generalized Fock matrix (see Fig. 3.1) are diagonal:

$$f^q_p = f^p_p \delta^q_p, \quad \forall\, p, q \in \mathbb{O},\ \forall\, \mathbb{O} \in \{\mathbb{C}, \mathbb{A}, \mathbb{V}\}. \tag{3.11}$$

The diagonal entries f^p_p can thus be viewed as orbital energies and they are denoted as ϵ_p. From here, the Møller–Plesset denominators are defined as:

$$\Delta^{ij\cdots}_{ab\cdots} = \epsilon_i + \epsilon_j + \cdots - \epsilon_a - \epsilon_b - \cdots. \tag{3.12}$$

In this semicanonical basis, the zeroth-order Hamiltonian $[\hat{H}^{(0)}]$ of DSRG-MRPT2 has a simple form:

$$\hat{H}^{(0)} = E_0 + \sum_p^{\mathbb{G}} \epsilon_p \{\hat{a}_p^\dagger \hat{a}_p\}, \tag{3.13}$$

where the curly braces "$\{\cdot\}$" indicate operator normal ordering with respect to the state Ψ_0, following the approach of Mukherjee and Kutzelnigg.[31]

3.2.1 DSRG-MRPT2 Energy

In this section, we summarize the DSRG-MRPT2 energy expression within the unrelaxed formalism. The reader is encouraged to consult Ref.[74] for a detailed derivation. In Table 3.2, we summarize the scalar and tensorial quantities that enter in the DSRG-MRPT2 energy expression.

Table 3.2: Summary of notations used in the DSRG-MRPT2 energy.

Name	Expression	Description
E_0	Eq. (3.6)	CASCI reference energy
$E^{(2)}$	Eq. (3.15)	DSRG second-order energy correction
h_p^q	$\langle\psi_p\|\hat{h}\|\psi_q\rangle$	1-electron integrals
v_{pq}^{rs}	$\langle\psi_p\psi_q\|\psi_r\psi_s\rangle$	antisymmetrized 2-electron integrals
$\bar{f}_p^q$	Eq. (3.9)	core Fock matrix
f_p^q	Eq. (3.10)	generalized Fock matrix
$\check{f}_i^a$	Eq. (3.19)	modified first-order Fock matrix
$\gamma_{ij\cdots}^{kl\cdots}$	Eq. (3.5)	n-particle reduced density matrices
$\Delta_{ab\cdots}^{ij\cdots}$	Eq. (3.12)	Møller–Plesset denominators
$\tilde{h}_{ij\cdots}^{ab\cdots}$	Eqs. (3.17) & (3.18)	modified first-order integrals
$t_{ab\cdots}^{ij\cdots}$	Eqs. (3.21) & (3.22)	first-order cluster amplitudes

The unrelaxed DSRG-MRPT2 energy $\mathcal{E}(s)$ is the sum of the the reference energy E_0 [Eq. (3.6)] and a second-order correction $E^{(2)}(s)$:

$$\mathcal{E}(s) = E_0 + E^{(2)}(s), \tag{3.14}$$

where $s \in [0, +\infty)$ is the DSRG flow parameter, whose significance will be clarified later. The second-order energy correction in Eq. (3.14) is given by the fully

contracted terms from an effective first-order Hamiltonian $\widetilde{H}^{(1)}(s)$ and a first-order cluster operator $\hat{T}^{(1)}(s)$:

$$E^{(2)}(s) = \langle \Psi_0 | [\widetilde{H}^{(1)}(s), \hat{T}^{(1)}(s)] | \Psi_0 \rangle \,. \tag{3.15}$$

Detailed expressions for equation (3.15) are presented in Appendix 3.A. In general, $E^{(2)}(s)$ is a sum of tensor contractions of the first-order cluster amplitudes $[t_a^{i,(1)}(s), t_{ab}^{ij,(1)}(s)]$, the modified first-order integrals $[\tilde{h}_i^{a,(1)}(s), \tilde{h}_{ij}^{ab,(1)}(s)]$, and 1-, 2- and 3-pRDMs. For brevity, in the following we drop the superscript "(1)" for the first-order quantities and the label "(s)" for s-dependent amplitudes or integrals.

The effective first-order Hamiltonian possesses the form

$$\widetilde{H} = \sum_{i}^{\mathbb{H}} \sum_{a}^{\mathbb{P}} \tilde{h}_i^a \{\hat{a}_a^i\} + \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \tilde{h}_{ij}^{ab} \{\hat{a}_{ab}^{ij}\}, \tag{3.16}$$

where the modified first-order integrals are given by:[74]

$$\tilde{h}_i^a = f_i^a + \check{f}_i^a - \Delta_a^i t_a^i, \quad \neg(\forall\, i, a \in \mathbb{A}), \tag{3.17}$$

$$\tilde{h}_{ij}^{ab} = 2v_{ij}^{ab} - \Delta_{ab}^{ij} t_{ab}^{ij}. \tag{3.18}$$

In Eq. (3.16), we have introduced a compact notation for a string of creation and annihilation operators: $\hat{a}_{rs\cdots}^{pq\cdots} = \hat{a}_p^\dagger \hat{a}_q^\dagger \cdots \hat{a}_s \hat{a}_r$. In Eq. (3.17), we have also defined an auxiliary one-body intermediate $\check{f}_i^a$

$$\check{f}_i^a = f_i^a + \sum_{ux}^{\mathbb{A}} \Delta_u^x \gamma_u^x t_{ax}^{iu}. \tag{3.19}$$

The DSRG-MRPT2 cluster operator $\hat{T}$ is written as

$$\hat{T} = \sum_{i}^{\mathbb{H}} \sum_{a}^{\mathbb{P}} t_a^i \{\hat{a}_i^a\} + \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} t_{ab}^{ij} \{\hat{a}_{ij}^{ab}\} \tag{3.20}$$

where the cluster amplitudes are determined via:

$$t_a^i = \check{f}_i^a \, \mathcal{R}_s(\Delta_a^i), \qquad \neg(\forall\, i, a \in \mathbb{A}), \tag{3.21}$$

$$t_{ab}^{ij} = v_{ij}^{ab} \, \mathcal{R}_s(\Delta_{ab}^{ij}), \qquad \neg(\forall\, i, j, a, b \in \mathbb{A}). \tag{3.22}$$

Here, $\mathcal{R}_s(\Delta)$ is a function that regularizes the inverse of Δ:

$$\mathcal{R}_s(\Delta) = \frac{1 - e^{-s\Delta^2}}{\Delta}. \tag{3.23}$$

The internal amplitudes labeled solely by active indices do not follow Eqs. (3.21) and (3.22) as they are assumed to be zero,[74] that is, $t^{xy\cdots}_{uv\cdots} = 0, \forall\, u, v, x, y, \cdots \in \mathbb{A}$. For convenience of deriving the gradient theory, these internal excitations are excluded from the cluster operator without loss of generality. For this reason, along with the fact that $f^v_u (\forall\, u, v \in \mathbb{A})$ contribute only to $\hat{H}^{(0)}$, the one-body components of $\widetilde{H}$ should not include elements labeled by only active indices either [see Eq. (3.17)]. Contrarily, no index restrictions apply to $\tilde{h}^{ab}_{ij}$ [Eq. (3.18)] because $v^{xy}_{uv}, \forall\, u, v, x, y \in \mathbb{A}$ are proper contributions to the first-order Hamiltonian such that $\tilde{h}^{xy}_{uv} = 2v^{xy}_{uv}, \forall\, u, v, x, y \in \mathbb{A}$. We also point out that both $\tilde{h}^{ab}_{ij}$ and t^{ij}_{ab} are antisymmetric with respect to individual permutations of upper or lower indices, e.g., $t^{ij}_{ab} = -t^{ij}_{ba} = -t^{ji}_{ab} = t^{ji}_{ba}$.

The denominators $\Delta^{ij\cdots}_{ab\cdots}$ that enter into the DSRG-MRPT2 amplitudes [Eqs. (3.21) and (3.22)] may be positive or close to zero. When the latter occurs, these amplitudes remain bounded for finite values of s because the divergence of the denominator is suppressed by the regularizer $\mathcal{R}_s$ [Eq. (3.23)]. However, even if a denominator is not zero, the magnitude of $\mathcal{R}_s$ can be as large as $\approx 0.6382\sqrt{s}$.[74] Therefore, it is necessary to use a value of s that balances the amount of correlation captured by the DSRG-MRPT2 with the risk of reintroducing intruders. Previous work[74] has found a "Goldilocks zone" for s around $\sim 1\ E_\mathrm{h}^{-2}$ that yields accurate results and avoids intruders.

3.2.2 DSRG-MRPT2 Gradients

Analytic expressions for the DSRG-MRPT2 gradients are obtained by taking the total derivatives of the energy [Eq. (3.14)] with respect to external perturbations. Without loss of generality, in this work we take the external perturbations to be nuclear displacements. The difficulty of deriving the DSRG-MRPT2 gradient theory

can be easily appreciated. Consider the derivative of an amplitude t_i^a with respect to an atomic coordinate R in the gradient contribution $\frac{\partial \mathcal{E}}{\partial t_i^a} \frac{\partial t_i^a}{\partial R}$. As shown in Eq. (3.21), all quantities that enter in the equation for t_i^a (f_a^i, Δ_u^v, and t_{au}^{iv}) depend on R, leading to numerous contributions to the derivative equations.

More importantly, the DSRG-MRPT2 energy implicitly depends on both the orbital coefficients $C_{\mu p}$ [Eq. (3.1)] and the CI coefficients c_I [Eq. (3.3)]. These quantities are determined by the CASSCF stationary conditions but are not variationally optimized in DSRG-MRPT2. Thus, computing the DSRG-MRPT2 gradients requires the evaluation of $\partial C_{\mu p}/\partial R$ and $\partial c_I/\partial R$, which can be solved via the coupled-perturbed (CP) CASSCF equation.[84,85] For a molecule with M atoms, there are $3M$ CP-CASSCF equations, the solution of which becomes computationally impractical for large systems. As realized by Handy and Schaefer,[83] the $3M$ CP-CASSCF equations may be replaced with a single perturbation-independent response equation (Z-vector approach), whose solution suffices to compute the energy derivatives for all nuclei. The Lagrangian formulation of the gradient theory of Helgaker and Jørgensen[80] directly leads to a set of response equations equivalent to the Z-vector approach.

Herein, we follow the standard approach for deriving analytic energy gradients based on the method of Lagrange multipliers.[80,81] The DSRG-MRPT2 Lagrangian ($\mathcal{L}$) reads as

$$\mathcal{L} = \mathcal{E} + \sum_{n=1}^{2} (\mathcal{T}_n + \tilde{\mathcal{H}}_n) + \mathcal{F} + \mathcal{W} + \mathcal{X} + \mathcal{Y}, \tag{3.24}$$

with scalar terms reflecting the constraints on the n-body cluster amplitudes ($\mathcal{T}_n$), the n-body modified integrals ($\tilde{\mathcal{H}}_n$), the use of semicanonical CASSCF orbitals ($\mathcal{F}$), the orthonormality of the MSOs ($\mathcal{W}$), and the use of a CASCI reference ($\mathcal{X}$) subject to normalization ($\mathcal{Y}$). In general, each of these terms is written as a dot product between a vector (or tensor) of equality constraints and the associated Lagrange *multipliers*, where every constraint is a zero-valued function of some *parameters*. All terms of

Eq. (3.24) are summarized in Table 3.3 and explicit definitions are discussed in detail in section 3.2.3.

Table 3.3: Summary for the DSRG-MRPT2 Lagrangian constraints.

Term	Constraints		Multipliers	Description
$\mathcal{T}_n$	$T^{ij\cdots}_{ab\cdots}$	Eqs. (3.35) & (3.36)	$\tau^{ij\cdots}_{ab\cdots}$	n-body cluster amplitudes
$\tilde{\mathcal{H}}_n$	$\tilde{H}^{ab\cdots}_{ij\cdots}$	Eqs. (3.37) & (3.38)	$\kappa^{ab\cdots}_{ij\cdots}$	n-body modified integrals
$\mathcal{F}$	F^q_p	Eqs. (3.29)–(3.32)	ζ^q_p	CASSCF semicanonical orbitals
$\mathcal{W}$	W^q_p	Eq. (3.47)	ω^q_p	orthonormal orbitals
$\mathcal{X}$	X_I	Eq. (3.40)	ξ_I	CI coefficients from CASCI
$\mathcal{Y}$	Y	Eq. (3.41)	ι	normalized CI coefficients

When the Lagrangian is stationary with respect to variations of *all* the parameters and multipliers, the DSRG-MRPT2 analytic energy gradients (evaluated at the reference geometry R_0) can be computed as:

$$\begin{aligned}\left.\frac{\mathrm{d}\mathcal{E}}{\mathrm{d}R}\right|_{R=R_0} &= \left.\frac{\partial\mathcal{L}}{\partial R}\right|_{R=R_0} \\ &= \sum_{pq}^{\mathbb{G}} \Gamma^p_q (h^p_q)^x + \sum_{pqrs}^{\mathbb{G}} \Gamma^{pq}_{rs} (v^{rs}_{pq})^x + \sum_{pq}^{\mathbb{G}} \omega^p_q (S^q_p)^x. \end{aligned} \tag{3.25}$$

Here, $(h^p_q)^x$, $(v^{rs}_{pq})^x$, $(S^q_p)^x$ are skeleton one-electron, antisymmetrized two-electron, and overlap derivative MSO integrals, respectively.[84,86,87] These quantities are multiplied by the corresponding relaxed one-body density (Γ^p_q), relaxed two-body density (Γ^{pq}_{rs}), and the energy-weighted density ω^p_q, which can be obtained by collecting the respective terms in front of h^q_p, v^{rs}_{pq}, and S^q_p in $\mathcal{L}$. Contributions to the relaxed densities are given in Appendix 3.B.

3.2.3 DSRG-MRPT2 Lagrangian Constraints

CASSCF Semicanonical Orbitals

To impose that the orbitals are variationally optimized using CASSCF and satisfy the semicanonical condition [Eq. (3.11)], we include the term $\mathcal{F}$ in the Lagrangian

function [Eq. (3.24)]. This term is defined as:

$$\mathcal{F} = \sum_{pq}^{\mathbb{G}} \zeta_p^q F_p^q, \tag{3.26}$$

where ζ_p^q are the Lagrange multipliers associated with the constraints $F_p^q = 0$. For converged CASSCF orbitals, the following conditions are satisfied:[88,89]

$$f_m^e = 0, \quad \tilde{f}_e^u = 0, \quad f_m^u - \tilde{f}_m^u = 0, \tag{3.27}$$

with the intermediate $\tilde{f}_p^u$ defined by:

$$\tilde{f}_p^u = \sum_{v}^{\mathbb{A}} \bar{f}_p^v \gamma_v^u + \frac{1}{2} \sum_{vxy}^{\mathbb{A}} v_{pv}^{xy} \gamma_{xy}^{uv}. \tag{3.28}$$

Equation (3.27) is easily translated to the following constraints:

$$F_m^e = F_e^m = f_m^e, \qquad m \in \mathbb{C}, e \in \mathbb{V}, \tag{3.29}$$

$$F_u^e = F_e^u = -\tilde{f}_e^u, \qquad u \in \mathbb{A}, e \in \mathbb{V}, \tag{3.30}$$

$$F_u^m = F_m^u = f_m^u - \tilde{f}_m^u, \qquad u \in \mathbb{A}, m \in \mathbb{C}, \tag{3.31}$$

where the symmetry of F_p^q reflects the Hermiticity of f_p^q and $\tilde{f}_p^q$. To formulate the CASSCF constraints in a consistent manner, in Eq. (3.30) we define F_u^e to be the negative of the corresponding term in Eq. (3.27).

We impose the semicanonical orbital basis condition [Eq. (3.11)] by defining the diagonal blocks of F_p^q as:

$$F_p^q = f_p^q - \epsilon_p \delta_p^q, \quad \forall\, p, q \in \mathbb{O},\ \forall\, \mathbb{O} \in \{\mathbb{C}, \mathbb{A}, \mathbb{V}\}. \tag{3.32}$$

We point out that in our formulation the orbital energies (ϵ_p) in Eq. (3.32) are treated as *parameters* constrained to take the value of diagonal elements of the generalized Fock operator, as done in MC-QDPT2 and CASPT2 gradient theories.[46,61] It can be easily checked that the quantities F_p^q implicitly depend on three sets of parameters: 1) the MSO coefficients $\mathbf{C}$ (via the one- and two-electron integrals), 2) the reference CI coefficients $\mathbf{c}$ (via the 1- and 2-pRDMs), and 3) the MSO orbital energies $\boldsymbol{\epsilon}$.

Cluster Amplitudes and Modified Integrals

The DSRG-MRPT2 correlation energy $E^{(2)}$ [see Eq. (3.15)] is a function of cluster amplitudes $(\mathbf{t}_1, \mathbf{t}_2)$, modified integrals $(\tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2)$, and the reference n-pRDMs, as shown in Appendix 3.A. To shift the dependence of $\mathbf{C}$ away from $E^{(2)}$, we consider both cluster amplitudes and modified integrals as parameters in the DSRG-MRPT2 Lagrangian. This aspect is embodied in the $\mathcal{T}_n$ and $\tilde{\mathcal{H}}_n$ constraints in Eq. (3.24), which are given by

$$\mathcal{T}_n = \frac{1}{(n!)^2} \sum_{ij\cdots}^{\mathbb{H}} \sum_{ab\cdots}^{\mathbb{P}} \tau_{ab\cdots}^{ij\cdots} T_{ab\cdots}^{ij\cdots}, \tag{3.33}$$

$$\tilde{\mathcal{H}}_n = \frac{1}{(n!)^2} \sum_{ij\cdots}^{\mathbb{H}} \sum_{ab\cdots}^{\mathbb{P}} \kappa_{ij\cdots}^{ab\cdots} \tilde{H}_{ij\cdots}^{ab\cdots}. \tag{3.34}$$

The constraints for the one- and two-body cluster amplitudes are obtained by rearranging Eqs. (3.21) and (3.22):

$$T_a^i = \check{f}_i^a \, \mathcal{R}_s(\Delta_a^i) - t_a^i, \qquad \neg(\forall\, i, a \in \mathbb{A}), \tag{3.35}$$

$$T_{ab}^{ij} = v_{ij}^{ab} \, \mathcal{R}_s(\Delta_{ab}^{ij}) - t_{ab}^{ij}, \qquad \neg(\forall\, i, j, a, b \in \mathbb{A}), \tag{3.36}$$

with the associated Lagrange multipliers τ_a^i and τ_{ab}^{ij}, respectively. Similarly, Eqs. (3.17) and (3.18) result in constraints for the modified integrals:

$$\tilde{H}_i^a = f_i^a + \check{f}_i^a - \Delta_a^i t_a^i - \tilde{h}_i^a, \quad \neg(\forall\, i, a \in \mathbb{A}), \tag{3.37}$$

$$\tilde{H}_{ij}^{ab} = 2v_{ij}^{ab} - \Delta_{ab}^{ij} t_{ab}^{ij} - \tilde{h}_{ij}^{ab}, \tag{3.38}$$

with the corresponding multipliers denoted as κ_i^a and κ_{ij}^{ab}, respectively. Notice again that Eqs. (3.35)–(3.37) inherit the restrictions of indices from Eqs. (3.21), (3.22), and (3.17). As far as the implicit dependence on parameters concerned in these constraints, T_{ab}^{ij} depends on $\mathbf{t}_2$, $\mathbf{C}$, and $\boldsymbol{\epsilon}$, while T_a^i depends on $\mathbf{t}_1$, $\mathbf{C}$, $\boldsymbol{\epsilon}$, $\mathbf{c}$ and $\mathbf{t}_2$. Compared to the same-rank amplitude constraints, the one- and two-body constraints for modified integrals simply add additional dependences on $\tilde{\mathbf{h}}_1$ and $\tilde{\mathbf{h}}_2$, respectively.

Reference CI Coefficients

Next, we discuss constraints that arise from enforcing the variational condition on the reference and its normalization. The reference wave function Ψ_0 [Eq. (3.3)] satisfies the eigenvalue problem:

$$\sum_J^{\mathcal{M}_0} \langle \Phi_I | \hat{H} | \Phi_J \rangle \, c_J = E_0 c_I, \quad \forall\, I \in \mathcal{M}_0, \tag{3.39}$$

subject to the normalization condition $\|\mathbf{c}\|_2^2 = \sum_I^{\mathcal{M}_0} c_I^2 = 1$. As such, we may write out the CI constraints as a vector (X_I) and a scalar (Y) defined as

$$X_I = \langle \Phi_I | \hat{H} | \Psi_0 \rangle - E_0 c_I, \tag{3.40}$$

$$Y = 1 - \sum_I^{\mathcal{M}_0} c_I^2, \tag{3.41}$$

and associate each constraint of Eq. (3.40) with a multiplier ξ_I and Eq. (3.41) with the multiplier ι. In Eq. (3.40), we have used the fact that the c_I coefficients are real to symmetrize the expression for X_I. It is easily verified that the X_I and Y constraints only depend on the parameters $\mathbf{C}$ and $\mathbf{c}$.

In the DSRG-MRPT2 Lagrangian [Eq. (3.24)], the CI constraints are imposed via both $\mathcal{X}$ and $\mathcal{Y}$:

$$\mathcal{X} = \sum_I^{\mathcal{M}_0} \xi_I X_I = (E_0^{\mathrm{c}} - E_0) \sum_I^{\mathcal{M}_0} \xi_I c_I + \tilde{E}_0^{\mathrm{a}}, \tag{3.42}$$

$$\mathcal{Y} = \iota \Big(1 - \sum_I^{\mathcal{M}_0} c_I^2 \Big), \tag{3.43}$$

where E_0^{c} has been defined in Eq. (3.7). The term $\tilde{E}_0^{\mathrm{a}}$ in Eq. (3.42) is similar to Eq. (3.8) except that the 1- and 2-pRDMs in Eq. (3.8) should be replaced to the corresponding modified RDMs (mRDMs) given by:

$$\tilde{\gamma}^{uv\cdots}_{xy\cdots} = \sum_{IJ}^{\mathcal{M}_0} \xi_I c_J \, \langle \Phi_I | \hat{a}^{uv\cdots}_{xy\cdots} | \Phi_J \rangle \, . \tag{3.44}$$

As noted by Celani and Werner,[52] any multiple of c_I can be added to ξ_I without altering the Lagrangian contribution $\mathcal{X}$ (since $\sum_I^{\mathcal{M}_0} c_I X_I = 0$). It is thus convenient

to use this degree of freedom to make $\boldsymbol{\xi}$ and $\mathbf{c}$ orthogonal:

$$\sum_{I}^{\mathcal{M}_0} \xi_I c_I = 0. \tag{3.45}$$

Given such orthogonality condition, the CI constraint $\mathcal{X}$ [Eq. (3.42)] can be further simplified to only one term $\tilde{E}_0^{\mathrm{a}}$.

Orbital Orthonormality

Lastly, the orthonormality of MSOs is imposed via the Lagrangian term $\mathcal{W}$:

$$\mathcal{W} = \sum_{pq}^{\mathbb{G}} \omega_p^q W_p^q, \tag{3.46}$$

where the constraints are defined by

$$W_p^q = \delta_p^q - S_p^q. \tag{3.47}$$

The multipliers ω_p^q are identified as elements of the energy-weighted density matrix. The MSO orthonormality constraint [Eq. (3.47)] depends parametrically only on the orbital coefficients $\mathbf{C}$.

3.2.4 DSRG-MRPT2 Lagrange Multipliers

After defining each term in the DSRG-MRPT2 Lagrangian, we solve for the Lagrange multipliers by imposing stationarity with respect to all the parameters ($\mathbf{C}, \mathbf{c}, \tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2, \mathbf{t}_1, \mathbf{t}_2$, and $\boldsymbol{\epsilon}$). These parameters can be separated into two categories. The orbital and CI coefficients stationary conditions resemble the coupled perturbed CASSCF equations, which require an iterative procedure for the solution of the corresponding multipliers ($\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$). Instead, the multipliers associated with the remaining parameters can be obtained in a direct way.

Modified Integrals

We first solve the Lagrange multipliers $\boldsymbol{\kappa}_1$ and $\boldsymbol{\kappa}_2$ corresponding to the modified integrals constraints. Taking the derivative of $\mathcal{L}$ with respect to the modified integrals

and setting them to zero leads to:

$$\frac{\partial \mathcal{L}}{\partial \tilde{h}_i^a} = 0 \quad \Rightarrow \quad \kappa_i^a = \frac{\partial E^{(2)}}{\partial \tilde{h}_i^a} = \langle \Psi_0 | [\{\hat{a}_a^i\}, \hat{T}] | \Psi_0 \rangle \,, \tag{3.48}$$

$$\frac{\partial \mathcal{L}}{\partial \tilde{h}_{ij}^{ab}} = 0 \quad \Rightarrow \quad \kappa_{ij}^{ab} = 4\frac{\partial E^{(2)}}{\partial \tilde{h}_{ij}^{ab}} = \langle \Psi_0 | [\{\hat{a}_{ab}^{ij}\}, \hat{T}] | \Psi_0 \rangle \,. \tag{3.49}$$

We point out that 1) κ_{ij}^{ab} is antisymmetric with respect to individual permutations of upper or lower indices and 2) those elements labeled by active indices are zero ($\kappa_u^v = \kappa_{uv}^{xy} = 0, \forall\, u, v, x, y \in \mathbb{A}$). Explicit expressions of κ_i^a and κ_{ij}^{ab} are reported in Appendix 3.C.1, where we evaluate the fully connected terms of the commutators in Eqs. (3.48) and (3.49). Identical expressions can be alternatively obtained by directly taking the partial derivatives of $E^{(2)}$ with respect to the modified integrals ($\tilde{h}_i^a$ and $\tilde{h}_{ij}^{ab}$) and antisymmetrizing the resulting contributions to κ_{ij}^{ab} with respect to index permutations.

Cluster Amplitudes

The Lagrange multipliers for the one-body cluster amplitudes can be easily solved:

$$\frac{\partial \mathcal{L}}{\partial t_a^i} = 0 \quad \Rightarrow \quad \tau_a^i = \frac{\partial E^{(2)}}{\partial t_a^i} - \kappa_i^a \Delta_a^i. \tag{3.50}$$

For the two-body multipliers, we have

$$\frac{\partial \mathcal{L}}{\partial t_{ab}^{ij}} = 0 \;\Rightarrow\; \tau_{ab}^{ij} = 4\frac{\partial}{\partial t_{ab}^{ij}}\left(E^{(2)} + \mathcal{T}_1 + \tilde{\mathcal{H}}_1\right) - \kappa_{ij}^{ab}\Delta_{ab}^{ij}. \tag{3.51}$$

The derivatives of the second-order energy correction with respect to cluster amplitudes can be written as:

$$\frac{\partial E^{(2)}}{\partial t_a^i} = \langle \Psi_0 | [\widetilde{H}, \{\hat{a}_i^a\}] | \Psi_0 \rangle \,, \tag{3.52}$$

$$\frac{\partial E^{(2)}}{\partial t_{ab}^{ij}} = \frac{1}{4} \langle \Psi_0 | [\widetilde{H}, \{\hat{a}_{ij}^{ab}\}] | \Psi_0 \rangle \,, \tag{3.53}$$

and their explicit expressions are presented in Appendix 3.C.1. To continue, we evaluate the partial derivatives of $\check{f}_k^c$ [Eq. (3.19)] with respect to t_{ab}^{ij}:

$$\frac{\partial \check{f}_k^c}{\partial t_{ab}^{ij}} = \frac{1}{4}\mathcal{P}(ab)\mathcal{P}(ij)(\Delta_i^a \gamma_i^a \delta_b^c \delta_k^j), \tag{3.54}$$

where $\mathcal{P}(pq)$ is an antisymmetrizer with respect to indices p and q: $\mathcal{P}(pq)f(p,q,r,\dots) = f(p,q,r,\dots) - f(q,p,r,\dots)$. We may then calculate the $\mathcal{T}_1$ and $\tilde{\mathcal{H}}_1$ terms in Eq. (3.51) as

$$\frac{\partial \mathcal{T}_1}{\partial t^{ij}_{ab}} = \frac{1}{4}\mathcal{P}(ab)\mathcal{P}(ij)\big[\Delta^a_i \gamma^a_i \tau^j_b \, \mathcal{R}_s(\Delta^j_b)\big], \tag{3.55}$$

$$\frac{\partial \tilde{\mathcal{H}}_1}{\partial t^{ij}_{ab}} = \frac{1}{4}\mathcal{P}(ab)\mathcal{P}(ij)\big(\Delta^a_i \gamma^a_i \kappa^b_j\big). \tag{3.56}$$

Two aspects are worth mentioning. First, only the active-active block of the 1-pRDM contributes to Eqs. (3.54)–(3.56). Hence, γ^a_i may be replaced with γ^u_v ($\forall\, u, v \in \mathbb{A}$) after appropriate reindexing. Second, multipliers labeled only by active indices (τ^v_u and τ^{xy}_{uv}, $\forall\, u, v, x, y \in \mathbb{A}$) are not defined because internal excitations are forbidden and they are conveniently set to zero in our implementation.

Orbital Energies

The diagonal elements of $\boldsymbol{\zeta}$ can be obtained by making the Lagrangian stationary with respect to the semicanonical orbital energies:

$$\frac{\partial \mathcal{L}}{\partial \epsilon_p} = 0 \quad \Rightarrow \quad \zeta^p_p = \frac{\partial}{\partial \epsilon_p}\Big[\sum_{n=1}^{2}(\mathcal{T}_n + \tilde{\mathcal{H}}_n)\Big]. \tag{3.57}$$

Evaluating the derivatives that enter into Eq. (3.57) is straightforward and the resulting expressions are provided in Appendix 3.C.2.

Energy-Weighted Density, Orbital Rotations, and CI Coefficients

The remaining unknowns are the energy-weighted density ($\boldsymbol{\omega}$) and the orbital ($\boldsymbol{\zeta}$) and CI ($\boldsymbol{\xi}$ and ι) multipliers. In principle, these quantities are all coupled together, but as shown by Celani and Werner,[52] it is possible to write separate equations for $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ from those for $\boldsymbol{\omega}$. The equations for $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ form a coupled linear systems, whose solution may be then used to evaluate $\boldsymbol{\omega}$.

When differentiating the Lagrangian with respect to the orbital coefficients $\mathbf{C}$, it is convenient to express this quantity as a unitary transformation of the unperturbed orbitals ($\mathbf{C}_0$):

$$\mathbf{C} = \mathbf{C}_0 \exp(\boldsymbol{\vartheta}). \tag{3.58}$$

Here, $\boldsymbol{\vartheta}$ is an anti-Hermitian matrix whose elements become the actual variational parameters. This parameterization ensures that the perturbed orbitals remain orthonormal.

Imposing the stationarity of the Lagrangian with respect to orbital rotations

$$\left(\frac{\partial \mathcal{L}}{\partial \boldsymbol{\vartheta}}\right)_{\boldsymbol{\vartheta}=0} = \left(\mathbf{C}^\dagger \frac{\partial \mathcal{L}}{\partial \mathbf{C}}\right)_{\boldsymbol{\vartheta}=0} = 0, \tag{3.59}$$

yields a set of equations that depend on $\boldsymbol{\omega}$, $\boldsymbol{\zeta}$, and $\boldsymbol{\xi}$ (via the mRDMs $\tilde{\gamma}^{uv\cdots}_{xy\cdots}$), as reported in Appendix 3.C.3. A way to decouple $\boldsymbol{\omega}$ from the other variables is suggested by the structure of the MSO overlap contribution to Eq. (3.59)

$$\frac{\partial \mathcal{W}}{\partial \vartheta^q_p} = \frac{\partial \mathcal{W}}{\partial \vartheta^p_q} = -\sum_r (\omega^r_p S^r_q + \omega^p_r S^q_r) = -(\omega^q_p + \omega^p_q). \tag{3.60}$$

To remove the dependence on $\boldsymbol{\omega}$, it is sufficient to consider the antisymmetric part of $\partial \mathcal{L}/\partial \vartheta^q_p$,

$$\frac{\partial \mathcal{L}}{\partial \vartheta^q_p} - \frac{\partial \mathcal{L}}{\partial \vartheta^p_q} = 0, \tag{3.61}$$

which only depends on the unsolved orbital ($\boldsymbol{\zeta}$) and CI ($\boldsymbol{\xi}$) multipliers. Equation (3.61) forms a set of linear equations of the form

$$\mathbf{A}^{\mathrm{oo}}\boldsymbol{\zeta} + \mathbf{A}^{\mathrm{oc}}\boldsymbol{\xi} = \mathbf{b}^{\mathrm{o}}, \tag{3.62}$$

where $\mathbf{A}^{\mathrm{oo}}$ and $\mathbf{A}^{\mathrm{oc}}$ are matrices of dimension N^2_{indep} and $N_{\mathrm{indep}}N_{\mathrm{det}}$, where N_{indep} is the number of independent orbital rotation parameters and N_{det} the number of CI determinants. The vector $\mathbf{b}^{\mathrm{o}}$ collects all constant terms and is of dimension N_{indep}. Equation (3.62) alone is insufficient to determine $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$, and must be augmented with additional conditions obtained from imposing stationarity with respect to the CI coefficients.

The derivative of $\mathcal{L}$ with respect to the CI coefficients takes the form

$$\frac{\partial \mathcal{L}}{\partial c_I} = \frac{\partial}{\partial c_I}(\mathcal{E} + \mathcal{T}_1 + \tilde{\mathcal{H}}_1 + \mathcal{F} + \mathcal{X} + \mathcal{Y}) = 0, \quad \forall I \in \mathcal{M}_0. \tag{3.63}$$

Equation (3.63) consists of a large set of linear equations for the orbital multipliers $\boldsymbol{\zeta}$ and the CI multipliers $\boldsymbol{\xi}$, that is,

$$\mathbf{A}^{\mathrm{co}}\boldsymbol{\zeta} + \mathbf{A}^{\mathrm{cc}}\boldsymbol{\xi} = \mathbf{b}^{\mathrm{c}}, \tag{3.64}$$

where the matrices $\mathbf{A}^{\mathrm{co}}$ and $\mathbf{A}^{\mathrm{cc}}$ are of size $N_{\mathrm{indep}}N_{\mathrm{det}}$ and N_{det}^2, respectively, while the vector $\mathbf{b}^{\mathrm{c}}$ contains N_{det} entries. The Lagrange multiplier connected to the CI normalization condition [ι, see Eq. (3.43)] can be computed as:

$$\iota = \frac{1}{2}\sum_{I}^{\mathcal{M}_0} c_I \frac{\partial}{\partial c_I}(\mathcal{E} + \mathcal{T}_1 + \tilde{\mathcal{H}}_1 + \mathcal{F}), \tag{3.65}$$

which depends on the orbital multipliers $\boldsymbol{\zeta}$ (see Appendix 3.C.4). Equation (3.65) is obtained from Eq. (3.63) ($\sum_I^{\mathcal{M}_0} c_I \frac{\partial \mathcal{L}}{\partial c_I} = 0$) using the fact that $\|\mathbf{c}\|_2^2 = 1$ and $\hat{H}\,|\Psi_0\rangle = E_0\,|\Psi_0\rangle$.

The linear equations for the orbital and CI multipliers [Eqs. (3.62) and (3.64)] may be combined into a single linear system of the form $\mathbf{A}\mathbf{x} = \mathbf{b}$ with entries defined as follows

$$\mathbf{A} \equiv \begin{pmatrix} \mathbf{A}^{\mathrm{oo}} & \mathbf{A}^{\mathrm{oc}} \\ \mathbf{A}^{\mathrm{co}} & \mathbf{A}^{\mathrm{cc}} \end{pmatrix}, \quad \mathbf{x} \equiv \begin{pmatrix} \boldsymbol{\zeta} \\ \boldsymbol{\xi} \end{pmatrix}, \quad \mathbf{b} \equiv \begin{pmatrix} \mathbf{b}^{\mathrm{o}} \\ \mathbf{b}^{\mathrm{c}} \end{pmatrix}. \tag{3.66}$$

When written in this form, $\mathbf{A}$ may be identified as a Jacobian matrix. Expressions for all the blocks of $\mathbf{A}$ and $\mathbf{b}$ are reported in Appendix 3.D. We postpone the discussion of how this linear system is solved to Sec. 3.2.4.

Once $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$ are determined, the symmetric counterpart of Eq. (3.61) (i.e., $\partial\mathcal{L}/\partial\vartheta_p^q + \partial\mathcal{L}/\partial\vartheta_q^p = 0$) can be used to recover the energy-weighted density:

$$\omega_p^q = \frac{1}{4}\Big(\frac{\partial}{\partial\vartheta_p^q} + \frac{\partial}{\partial\vartheta_q^p}\Big)\Big[E_0 + \sum_{n=1}^{2}(\mathcal{T}_n + \tilde{\mathcal{H}}_n) + \mathcal{F} + \mathcal{X}\Big]. \tag{3.67}$$

Here, we have ignored the contribution from $E^{(2)}$ because it is independent on orbital rotations, as stated in Sec. 3.2.3.

Iterative Solution of the Orbital and CI Multipliers

The system of linear equations for the orbital and CI multipliers involve $N_{\text{indep}} + N_{\text{det}}$ variables. When either the number of independent pairs or the size of the CI space becomes too large, it is unfeasible to store $\mathbf{A}$ explicitly and solve the linear system by direct inversion. The standard solution to this problem is employ a direct iterative linear solver that directly builds a vector $\boldsymbol{\sigma} \equiv \mathbf{A}\mathbf{x}$, thus avoiding the storage problem.

As noted in Sec. 3.2.3, to find a unique solution to the CI multiplier equations, we impose the constraint $\boldsymbol{\xi} \cdot \mathbf{c} = 0$ [see Eq. (3.45)]. To enforce this constraint in the solution of the linear system, we define a projection matrix $\mathbf{P}$

$$\mathbf{P} = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} - \mathbf{c}\mathbf{c}^T \end{pmatrix}. \tag{3.68}$$

Then the constraint $\boldsymbol{\xi} \cdot \mathbf{c} = 0$ is equivalent to the condition $\mathbf{P}\mathbf{x} = \mathbf{x}$, and we may use this result to write the linear system in the form

$$(\mathbf{PAP})\mathbf{x} = \mathbf{P}\mathbf{b}. \tag{3.69}$$

The matrix $\mathbf{PAP}$ is rank deficient since the vector $\mathbf{x}^{\parallel} = (\mathbf{0}, \mathbf{c})^T$ is such that $\mathbf{P}\mathbf{x}^{\parallel} = \mathbf{0}$. This linear system can be solved using the generalized minimal residual method without explicitly storing the matrix $\mathbf{PAP}$.

3.2.5 Computational Cost

We end this section by briefly discussing the computational cost of the DSRG-MRPT2 analytic gradients. In general, the cost of solving the Lagrange multipliers has the same scaling of a DSRG-MRPT2 single-point computation with a slightly larger prefactor. A vanilla DSRG-MRPT2 energy computation based on the CASSCF orbitals can be largely separated into four steps:

1. Solve the CASSCF problem for the orbital and CI coefficients.

2. Compute the 1-, 2-, and 3-pRDMs using the CASSCF wave function.

3. Transform the one- and two-electron integrals to the MO basis.

4. Build the modified integrals and cluster amplitudes, and use these quantities to evaluate the DSRG-MRPT2 correlation energy via tensor contractions.

In comparison, the gradient computations take the following additional steps:

5. Compute the multipliers $\boldsymbol{\kappa}_1, \boldsymbol{\kappa}_2, \boldsymbol{\tau}_1, \boldsymbol{\tau}_2$, and ζ_p^p.

6. Setup the coupled linear system [Eq. (3.66)] and solve for the multipliers $\boldsymbol{\zeta}$ and $\boldsymbol{\xi}$.

7. Form the relaxed density matrices and energy-weighted density.

8. Transform the MO densities in step 7 to the AO basis and contract it with skeleton derivative integrals.

We can see a rough correspondence between these two procedures. For example, step 5 in the gradients computation corresponds to step 4 of the energy computation. In these two steps, with the assumption of using a small active space ($N_\mathrm{A} \ll N_\mathrm{C} < N_\mathrm{V}$), the computational cost is dominated by the tensor contraction of an MP2-like term with a scaling of $\mathcal{O}(N_\mathrm{C}^2 N_\mathrm{V}^2)$. The cost of solving the linear system (step 6) is slightly higher than the cost of second-order CASSCF optimization (step 1). In fact, the linear system [Eq. (3.66)] is analogous to the Newton optimization step in CASSCF, where $\mathbf{A}$ and $\mathbf{b}$ correspond to the Hessian and the gradient vector, respectively. However, the $\mathbf{b}^\mathrm{c}$ vector contains terms involving $\partial \gamma_{xyz}^{uvw} / \partial c_I$, which share the same $\mathcal{O}(N_\mathrm{A}^6 N_\mathrm{det})$ scaling of computing the 3-pRDM. This steep cost may be avoided by considering the additional tensor contractions and introducing clever intermediates, as suggested in Ref.[45] The computational scaling for the MO to AO transformation of the relaxed densities (step 8) is identical to the integral transformation step for the DSRG-MRPT2 energy (step 3). Overall, we see that the

computational cost to obtain the analytical gradients is similar to that of an energy computation.

3.3 Results

We implemented the DSRG-MRPT2 analytic energy gradients in the open-source program Forte.[90] The one- and two-electron integrals along with the corresponding derivative integrals were obtained from Psi4 1.4.[91] The correctness of the implementation was validated by comparing the analytic gradients against five-point finite-difference numerical gradients using a 0.005 a.u. step size. In particular, we tested the gradient and the optimized bond lengths of HF and N_2 using CASSCF(2,2) and CASSCF(6,6) reference wave functions, respectively. The cc-pCVDZ basis set[92,93] was used for all computations in this work.

As a pilot application of the DSRG-MRPT2 gradient theory, we optimized the geometry of *p*-benzyne for both singlet and triplet states. The resulting geometries were used to compute the adiabatic singlet–triplet gap ($\Delta E_{\mathrm{ST}} = E_{\mathrm{T}} - E_{\mathrm{S}}$). We compared the DSRG-MRPT2 results to those of CASPT2,[16] the partially contracted NEVPT2 (pc-NEVPT2),[18] and Mukherjee's state-specific multireference coupled cluster theory with singles, doubles, and perturbative triples [Mk-MRCCSD(T)].[94,95] We employed the minimal CAS(2,2) active space that consists of two electrons in the two σ orbitals located on the dehydrogenated carbon atoms. For geometry optimizations, the maximum component of the gradient was converged to less than 2×10^{-6} a.u. Both CASPT2 and pc-NEVPT2 results were obtained using Molpro 2015.1[96] while those from Mk-MRCCSD(T) were computed using Psi4 1.4.[91]

We notice that the minimal active space and the cc-pCVDZ basis set are both too small to generate convincing results for *p*-benzyne to compare with experiments. However, there are several limitations in the current implementation that prevent us from large-scale applications of the DSRG-MRPT2 gradient theory, including 1) the

storage of all four-index two-electron integrals in memory, 2) the explicit storage for the derivatives of 1-, 2-, and 3-pRDMs with respect to CI coefficients, and 3) the lack of functionality of freezing core orbitals. We plan to address all these issues in a forthcoming paper.

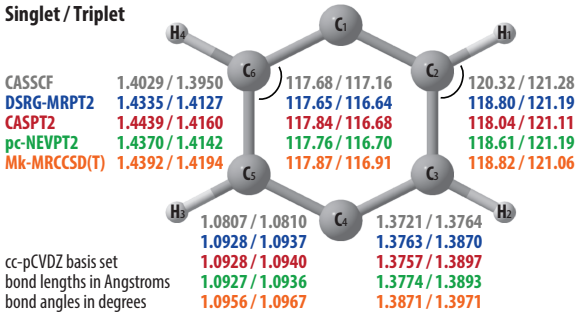

Figure 3.2: Equilibrium geometries of singlet and triplet *p*-benzyne optimized using various multireference methods using the cc-pCVDZ basis set. The DSRG flow parameter was set to 1 E_h^{-2}. The CASSCF(2,2) reference was used for all computations.

Figure 3.2 presents the DSRG-MRPT2 optimized geometries of singlet and triplet *p*-benzyne. Here, we set the flow parameter to $s = 1\ E_\mathrm{h}^{-2}$, a value that previously shown to yield reliable singlet–triplet gap of *p*-benzyne.[74] The DSRG-MRPT2 optimized geometries are in excellent agreements to those of CASPT2 and pc-NEVPT2. For example, the DSRG-MRPT2 bond lengths and angles deviate from those of CASPT2 by at most 1.0 pm ($C_5{-}C_6$ of the singlet) and 0.8° ($\angle H_1C_2C_3$ of the singlet), respectively. Compared to Mk-MRCCSD(T), DSRG-MRPT2 underestimates all C−C bonds by roughly 1 pm for both singlet and triplet states. Interestingly, all bonds are elongated when dynamical correlation effects are added on top of CASSCF, most prominently displayed by the 3.0–4.0 pm difference on the $C_5{-}C_6$ bond of the singlet.

In Fig. 3.3, we show the sensitivity of the DSRG-MRPT2 optimized bond distances and angles [computed as deviations from CASPT2 values] with respect to the flow parameter s. As s increases from 0, all geometric parameters vary quickly and converge roughly around $s = 1\ E_\mathrm{h}^{-2}$. Interestingly, the bond lengths first decrease

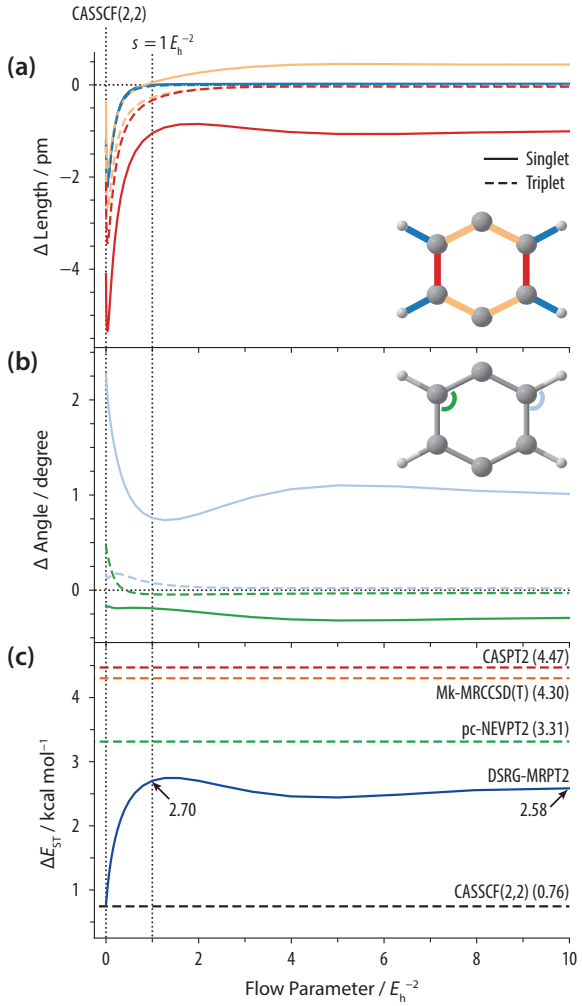

Figure 3.3: (a) Bond lengths and (b) bond angles from DSRG-MRPT2 optimized geometries relative to the CASPT2 optimized geometries as a function of the flow parameter. (c) DSRG-MRPT2 adiabatic singlet–triplet gap as a function of the flow parameter. Two flow parameters are indicated by the vertical dotted lines: $s = 0$ [i.e., CASSCF(2,2)] and $s = 1\ E_h^{-2}$ (used in Fig. 3.2).

for $s < 0.04\ E_h^{-2}$ and then start to increase for $0.04 < s < 1\ E_h^{-2}$. When s keeps growing from 1 to 10 E_h^{-2}, small yet noticeable changes are observed for the $C_5{-}C_6$ bond (≤ 0.4 pm) and $\angle H_1C_2C_3$ ($< 0.4°$) of the singlet.

The bottom panel of Fig. 3.3 also reports the adiabatic singlet–triplet splittings of *p*-benzyne computed using various multireference methods. For DSRG-MRPT2, we again see a quick convergence of ΔE_{ST} near $s = 1\ E_h^{-2}$, while further increase of s leads to only 0.1 kcal mol^{-1} difference in ΔE_{ST}. The DSRG-MRPT2 ($s = 1$)

prediction of ΔE_{ST} is 2.7 kcal mol^{-1}, which is 0.6, 1.8, 1.6 kcal mol^{-1} smaller than that of pc-NEVPT2, CASPT2, and Mk-MRCCSD(T), respectively. This underestimation has been observed previously, even when ΔE_{ST} is computed using optimized Mk-MRCCSD/cc-pVTZ geometries.[74] However, we note that the ΔE_{ST} of DSRG-MRPT2 can be improved via reference relaxation.[72] Specifically, using the corresponding unrelaxed DSRG-MRPT2 geometries reported in Fig. 3.2, the partially relaxed and the relaxed versions of DSRG-MRPT2 ($s = 1$) predict the ΔE_{ST} to be 3.57 and 3.76 kcal mol^{-1}, respectively. These values fall in between the pc-NEVPT2 and Mk-MRCCSD(T) results.

Finally, we investigate the effects of neglecting the 3-body cumulants in DSRG-MRPT2, which avoids the computation of the reference 3-pRDM without introducing artificial intruders.[38,74] Without three-body cumulants, the active space dependence of the DSRG-MRPT2 gradients can be reduced to $\mathcal{O}(N_{\mathrm{A}}^5 N_{\mathrm{V}})$ for the correlation energy terms and $\mathcal{O}(N_{\mathrm{A}}^4 N_{\mathrm{det}})$ for the CASCI contribution to the response equations. To test the accuracy of this approximation, we re-optimized the singlet and triplet geometries of *p*-benzyne using $s = 1$ E_{h}^{-2}. Comparing to the geometries optimized using the complete DSRG-MRPT2 theory (i.e., with 3-body cumulant contributions), the equilibrium bond lengths and bond angles deviate by at maximum 0.4 pm (C_5−C_6 of the singlet) and 0.5° ($\angle H_1C_2C_3$ of the singlet), respectively. We also employed these geometries yet computed the ΔE_{ST} using the complete DSRG-MRPT2 theory. The resulting spin gap is 2.67 kcal mol^{-1}, only 0.03 kcal mol^{-1} smaller than the exact answer (see Fig. 3.3). Such negligible deviation indicates that the pruned DSRG-MRPT2 scheme may be used to optimize the geometry without significant degradation of the accuracy.

3.4 Conclusions

We have derived and implemented the analytic energy gradients of the unrelaxed DSRG-MRPT2 approach. Our derivation uses the method of Lagrange multipliers to impose constraints on the first-order modified integrals and cluster amplitudes, the orbitals, and CI coefficients. Despite the complexity of the final Lagrangian, analytic expressions for the DSRG-MRPT2 gradient could still be derived by hand (albeit via a laborious procedure). Inheriting the numerical robustness of the DSRG equations, the corresponding Lagrangian is similarly well-behaved even when small energy denominators arise, circumventing the intruder state-problem in linear-response computations.

We have used the DSRG-MRPT2 analytic gradients to optimize the equilibrium structures of *p*-benzyne and study the dependence of the optimized geometry on the flow parameter s. The optimized geometries of both the singlet and triplet states show very good agreement with those computed with CASPT2, pc-NEVPT2, and Mk-MRCCSD(T). Comparing geometries between the two states, those of the singlet are more sensitive to the value of s. The s-dependency plot also shows that sufficient correlation contributions are captured with s equal to 1.0 E_h^{-2}. Finally, we investigate the singlet-triplet splittings of *p*-benzyne, which are underestimated using the DSRG-MRPT2 formalism compared against other MRPT2 approaches.

The current results motivate us for further developments of the theory. On the algorithmic side, an important bottleneck is the high memory cost required to store the two-electron integrals, which may be alleviated by applying resolution the identity techniques (i.e., density fitting). Another important limitation of the current implementation is the explicit storage of derivatives of the reduced density matrices required to solve the coupled equations for the orbital and CI multipliers. From the theory perspective, it would highly desirable to develop the gradient theory for the state-averaged DSRG-MRPT2[97] to allow optimizations of both ground- and excited-

state PESs, in particular near conical intersections. A more challenging future extension is the development of analytic gradients of higher-order MR-DSRG theories, including third-order perturbation theory and nonperturbative MR-DSRG methods. These extensions would require significant human effort and the use of automatic implementation techniques.[1,58,64]

3.5 Acknowledgements

The authors would like to acknowledge helpful discussions with Toru Shiozaki, Yuanzhe Xi, Haruyuki Nakano, and Alexander Sokolov. This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004. F.A.E. acknowledges support from a Camille Dreyfus Teacher-Scholar Award.

Appendix

3.A DSRG-MRPT2 Energy Expression

The DSRG-MRPT2 energy contributions are reported in Table 3.A1. In our previous work,[74,77] these terms are written in terms of cluster amplitudes, modified first-order integrals, 1-pRDM, one-hole RDM (1-hRDM), and two- and three-body density cumulants. In contrast, here we expand the 1-hRDM and all density cumulants in terms of 1-, 2-, and 3-pRDMs for the purpose of deriving the CI response terms (see Appendix 3.C.4).

Table 3.A1: DSRG-MRPT2 energy contributions expressed in terms of the modified first-order integrals ($\tilde{h}$), the first-order cluster amplitudes (t), and the reference 1-, 2-, and 3-pRDMs (γ). More compact expressions can be found in Ref. 74.

Term	Energy Expression
A1	$+\sum_{a}^{\mathbb{P}}\sum_{m}^{\mathbb{C}}\tilde{h}_{m}^{a}t_{a}^{m}$
A2	$+\sum_{e}^{\mathbb{V}}\sum_{uv}^{\mathbb{A}}\tilde{h}_{v}^{e}t_{e}^{u}\gamma_{u}^{v}-\sum_{m}^{\mathbb{C}}\sum_{uv}^{\mathbb{A}}\tilde{h}_{m}^{u}t_{v}^{m}\gamma_{u}^{v}$
B1	$+\frac{1}{2}\sum_{e}^{\mathbb{V}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{x}^{e}t_{ey}^{uv}\gamma_{uv}^{xy}-\frac{1}{2}\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{m}^{v}t_{xy}^{um}\gamma_{uv}^{xy}$
B2	$-\sum_{e}^{\mathbb{V}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{x}^{e}t_{ey}^{uv}\gamma_{u}^{x}\gamma_{v}^{y}+\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{m}^{v}t_{xy}^{um}\gamma_{u}^{x}\gamma_{v}^{y}$
C1	$+\frac{1}{2}\sum_{e}^{\mathbb{V}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{xy}^{ev}t_{e}^{u}\gamma_{uv}^{xy}-\frac{1}{2}\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{my}^{uv}t_{x}^{m}\gamma_{uv}^{xy}$
C2	$-\sum_{e}^{\mathbb{V}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{xy}^{ev}t_{e}^{u}\gamma_{u}^{x}\gamma_{v}^{y}+\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{my}^{uv}t_{x}^{m}\gamma_{u}^{x}\gamma_{v}^{y}$
D1	$+\frac{1}{4}\sum_{ab}^{\mathbb{P}}\sum_{mn}^{\mathbb{C}}\tilde{h}_{mn}^{ab}t_{ab}^{mn}$
D2	$+\frac{1}{2}\sum_{ab}^{\mathbb{P}}\sum_{m}^{\mathbb{C}}\sum_{uv}^{\mathbb{A}}\tilde{h}_{mu}^{ab}t_{ab}^{mv}\gamma_{v}^{u}-\frac{1}{2}\sum_{a}^{\mathbb{P}}\sum_{mn}^{\mathbb{C}}\sum_{uv}^{\mathbb{A}}\tilde{h}_{mn}^{av}t_{au}^{mn}\gamma_{v}^{u}$
D3	$+\frac{1}{8}\sum_{ab}^{\mathbb{P}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{xy}^{ab}t_{ab}^{uv}\gamma_{uv}^{xy}+\frac{1}{8}\sum_{mn}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{mn}^{uv}t_{xy}^{mn}\gamma_{uv}^{xy}+\sum_{a}^{\mathbb{P}}\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{mx}^{au}t_{ay}^{mv}\gamma_{uv}^{xy}$
D4	$-\sum_{a}^{\mathbb{P}}\sum_{m}^{\mathbb{C}}\sum_{uvxy}^{\mathbb{A}}\tilde{h}_{mx}^{au}t_{ay}^{mv}\gamma_{u}^{x}\gamma_{v}^{y}$
D5	$-\frac{1}{4}\sum_{e}^{\mathbb{V}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{xy}^{ew}t_{ez}^{uv}\gamma_{uvw}^{xyz}+\frac{1}{4}\sum_{m}^{\mathbb{C}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{mz}^{uv}t_{xy}^{mw}\gamma_{uvw}^{xyz}$
D6	$+\frac{1}{2}\sum_{e}^{\mathbb{V}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{xy}^{ew}t_{ez}^{uv}\gamma_{w}^{x}\gamma_{uv}^{yz}+\frac{1}{2}\sum_{e}^{\mathbb{V}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{xy}^{ew}t_{ez}^{uv}\gamma_{u}^{z}\gamma_{vw}^{xy}$
D6	$-\frac{1}{2}\sum_{m}^{\mathbb{C}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{mz}^{uv}t_{xy}^{mw}\gamma_{w}^{x}\gamma_{uv}^{yz}-\frac{1}{2}\sum_{m}^{\mathbb{C}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{mz}^{uv}t_{xy}^{mw}\gamma_{u}^{z}\gamma_{vw}^{xy}$
D7	$-\sum_{e}^{\mathbb{V}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{xy}^{ew}t_{ez}^{uv}\gamma_{u}^{y}\gamma_{v}^{z}\gamma_{w}^{x}+\sum_{m}^{\mathbb{C}}\sum_{uvwxyz}^{\mathbb{A}}\tilde{h}_{mz}^{uv}t_{xy}^{mw}\gamma_{u}^{y}\gamma_{v}^{z}\gamma_{w}^{x}$

3.B Contributions to the Relaxed Density Matrices

We first focus on the one-body relaxed density matrix elements Γ^q_p. Assuming the use of real orbitals, the nonzero elements of Γ^q_p are given by

$$\Gamma^e_f = \zeta^e_f, \qquad \forall\, e, f \in \mathbb{V} \tag{3.70}$$

$$\Gamma^m_n = \delta^m_n + \zeta^m_n, \qquad \forall\, m, n \in \mathbb{C} \tag{3.71}$$

$$\Gamma^u_v = \gamma^v_u + \tilde{\gamma}^v_u + \zeta^u_v, \qquad \forall\, u, v \in \mathbb{A} \tag{3.72}$$

$$\Gamma^e_m = \alpha^e_m + 2\zeta^e_m, \qquad \forall\, e \in \mathbb{V}, \forall\, m \in \mathbb{C} \tag{3.73}$$

$$\Gamma^u_m = \alpha^u_m + 2\sum_v^{\mathbb{A}} \zeta^v_m \eta^v_u, \qquad \forall\, u \in \mathbb{A}, \forall\, m \in \mathbb{C} \tag{3.74}$$

$$\Gamma^e_u = \alpha^e_u - 2\sum_v^{\mathbb{A}} \zeta^v_e \gamma^v_u, \qquad \forall\, e \in \mathbb{V}, \forall\, u \in \mathbb{A} \tag{3.75}$$

where $\eta^a_b = \delta^a_b - \gamma^a_b$ are the 1-hRDM elements and we also introduce an intermediate $\alpha^{ab\cdots}_{ij\cdots} = 2\kappa^{ab\cdots}_{ij\cdots} + \tau^{ij\cdots}_{ab\cdots}\mathcal{R}_s(\Delta^{ij\cdots}_{ab\cdots})$ for convenience. In deriving Eqs. (3.70)–(3.75), we have used the fact that $\boldsymbol{\zeta}$ is symmetric, that is, $\zeta^q_p = \zeta^p_q$.

The two-body relaxed density matrix elements can be summarized as follows. There are seven terms involving α^{ab}_{ij}:

$$\Gamma^{ef}_{mn} = \frac{1}{4}\alpha^{ef}_{mn}, \tag{3.76}$$

$$\Gamma^{ef}_{uv} = \frac{1}{4}\alpha^{ef}_{uv}, \tag{3.77}$$

$$\Gamma^{uv}_{mn} = \frac{1}{4}\alpha^{ef}_{uv}, \tag{3.78}$$

$$\Gamma^{eu}_{mn} = \frac{1}{2}\alpha^{eu}_{mn}, \tag{3.79}$$

$$\Gamma^{ev}_{mu} = \alpha^{ev}_{mu} + \big(\alpha^e_m + 2\zeta^e_m\big)\gamma^u_v, \tag{3.80}$$

$$\Gamma^{ey}_{ux} = \frac{1}{2}\alpha^{ey}_{ux} + \alpha^e_u\gamma^x_y - \sum_v^{\mathbb{A}} \zeta^e_v \gamma^{vy}_{ux}, \tag{3.81}$$

$$\Gamma^{ux}_{my} = \frac{1}{2}\alpha^{ux}_{my} + \big(\alpha^u_m + 2\zeta^u_m\big)\gamma^y_x - \sum_v^{\mathbb{A}} \zeta^v_m \gamma^{vy}_{ux}. \tag{3.82}$$

The remaining eight terms are expressed as:

$$\Gamma^{em}_{fm} = \Gamma^{e}_{f}, \tag{3.83}$$

$$\Gamma^{en}_{mn} = \Gamma^{e}_{m}, \tag{3.84}$$

$$\Gamma^{un}_{mn} = \Gamma^{u}_{m}, \tag{3.85}$$

$$\Gamma^{em}_{um} = \Gamma^{e}_{u}, \tag{3.86}$$

$$\Gamma^{mo}_{no} = \frac{1}{2}\delta^{m}_{n} + \zeta^{m}_{n}, \tag{3.87}$$

$$\Gamma^{um}_{vn} = \zeta^{m}_{n}\gamma^{v}_{u} + \delta^{m}_{n}\Gamma^{u}_{v}, \tag{3.88}$$

$$\Gamma^{eu}_{fv} = \zeta^{e}_{f}\gamma^{v}_{u}, \tag{3.89}$$

$$\Gamma^{uv}_{xy} = \frac{1}{4}\left(\gamma^{uv}_{xy} + \tilde{\gamma}^{uv}_{xy}\right) + \zeta^{u}_{x}\gamma^{y}_{v}. \tag{3.90}$$

Here, we follow the index convention in Table 3.1 and thus omit the orbital type in the above equations for brevity.

3.C Derivatives in Multiplier Equations

3.C.1 Modified Integrals and Cluster Amplitudes

In this appendix, we derive the expressions for the derivatives given by Eqs. (3.48), (3.49), (3.52) and (3.53). Because $\widetilde{H}$ and $\hat{T}$ include only one- and two-body operators, we only need to derive the fully connected terms from the following three types of commutators $[\{\hat{a}^{i}_{a}\}, \{\hat{a}^{b}_{j}\}]$, $[\{\hat{a}^{i}_{a}\}, \{\hat{a}^{cd}_{kl}\}]$, and $[\{\hat{a}^{ij}_{ab}\}, \{\hat{a}^{cd}_{kl}\}]$. These commutators are

evaluated to be

$$[\{\hat{a}^i_a\},\{\hat{a}^b_j\}]_0 = \gamma^i_j\eta^b_a - \eta^i_j\gamma^b_a, \tag{3.91}$$

$$[\{\hat{a}^i_a\},\{\hat{a}^{cd}_{kl}\}]_0 = \mathcal{P}(cd)\delta^c_a\lambda^{id}_{kl} - \mathcal{P}(kl)\delta^i_k\lambda^{cd}_{al}, \tag{3.92}$$

$$\begin{aligned}[\{\hat{a}^{ij}_{ab}\},\{\hat{a}^{cd}_{kl}\}]_0 =& \mathcal{P}(kl)\mathcal{P}(cd)(\gamma^i_k\gamma^j_l\eta^c_a\eta^d_b - \eta^i_k\eta^j_l\gamma^c_a\gamma^d_b)\\ &+ \mathcal{P}(kl)(\gamma^i_k\gamma^j_l - \eta^i_k\eta^j_l)\lambda^{cd}_{ab}\\ &+ \mathcal{P}(cd)(\eta^c_a\eta^d_b - \gamma^c_a\gamma^d_b)\lambda^{ij}_{kl}\\ &+ \mathcal{P}(ij)\mathcal{P}(ab)\mathcal{P}(kl)\mathcal{P}(cd)\gamma^i_k\eta^c_a\lambda^{jd}_{bl}\\ &- \mathcal{P}(ij)\mathcal{P}(ab)\mathcal{P}(kl)\mathcal{P}(cd)\eta^i_k\gamma^c_a\lambda^{jd}_{bl}\\ &+ \mathcal{P}(ab)\mathcal{P}(cd)\delta^c_a\lambda^{ijd}_{kbl}\\ &- \mathcal{P}(ij)\mathcal{P}(kl)\delta^i_k\lambda^{jcd}_{bal}.\end{aligned} \tag{3.93}$$

Here, we have introduced the two- and three-body density cumulants defined by:

$$\lambda^{xy}_{uv} = \gamma^{xy}_{uv} - \gamma^x_u\gamma^y_v + \gamma^y_u\gamma^x_v, \tag{3.94}$$

$$\lambda^{xyz}_{uvw} = \gamma^{xyz}_{uvw} - \sum_\pi(-1)^{\mathcal{N}(\pi)}\gamma^x_u\lambda^{yz}_{vw} - \det(\gamma^x_u\gamma^y_v\gamma^z_w). \tag{3.95}$$

In Eq. (3.95), $\det(\cdot)$ indicates the sum of all permutations of lower (or upper) labels with a sign factor corresponding to the parity of permutations and $\sum_\pi(-1)^{\mathcal{N}(\pi)}$ indicates a sum over all permutations of the lower and upper labels with a sign factor given by the number of inversions in π [$\mathcal{N}(\pi)$].[31]

It is now easy to check that

$$\begin{aligned}\frac{\partial E^{(2)}}{\partial \tilde{h}_i^a} &= \sum_j^{\mathbb{H}} \sum_b^{\mathbb{P}} t_b^j \gamma_j^i \eta_a^b \\ &\quad + \frac{1}{2} \sum_{kl}^{\mathbb{H}} \sum_d^{\mathbb{P}} t_{ad}^{kl} \lambda_{kl}^{id} - \frac{1}{2} \sum_{cd}^{\mathbb{P}} \sum_l^{\mathbb{H}} t_{cd}^{il} \lambda_{al}^{cd}, \end{aligned} \tag{3.96}$$

$$\begin{aligned} 4\frac{\partial E^{(2)}}{\partial \tilde{h}_{ij}^{ab}} &= \mathcal{P}(ab) \sum_k^{\mathbb{H}} t_a^k \lambda_{kb}^{ij} - \mathcal{P}(ij) \sum_c^{\mathbb{P}} t_c^i \lambda_{ab}^{cj} \\ &\quad + \sum_{kl}^{\mathbb{H}} \sum_{cd}^{\mathbb{P}} t_{cd}^{kl} \left[\gamma_k^i \gamma_l^j \eta_a^c \eta_b^d + \mathcal{P}(ij)\mathcal{P}(ab) \gamma_k^i \eta_a^c \lambda_{bl}^{jd} \right] \\ &\quad + \frac{1}{2} \sum_{kl}^{\mathbb{H}} \sum_{cd}^{\mathbb{P}} t_{cd}^{kl} \left(\gamma_k^i \gamma_l^j \lambda_{ab}^{cd} + \eta_a^c \eta_b^d \lambda_{kl}^{ij} \right) \\ &\quad + \frac{1}{2} \mathcal{P}(ab) \sum_{kl}^{\mathbb{H}} \sum_d^{\mathbb{P}} t_{ad}^{kl} \lambda_{kbl}^{ijd} \\ &\quad - \frac{1}{2} \mathcal{P}(ij) \sum_l^{\mathbb{H}} \sum_{cd}^{\mathbb{P}} t_{cd}^{il} \lambda_{bal}^{jcd}. \end{aligned} \tag{3.97}$$

The derivatives of $E^{(2)}$ with respect to amplitudes can be obtained by making the replacements of $t_a^i \to \tilde{h}_i^a$ and $t_{ab}^{ij} \to \tilde{h}_{ij}^{ab}$ in Eqs. (3.96) and (3.97).

Several properties can be used to further simply Eqs. (3.96) and (3.97). The 1-pRDMs γ_q^p and 1-hRDMs η_q^p possess very simple structures:

$$\gamma_q^p = \begin{cases} \delta_q^p & \forall\, p, q \in \mathbb{C} \\ \gamma_q^p & \forall\, p, q \in \mathbb{A}\,, \\ 0 & \text{otherwise} \end{cases} \quad \eta_q^p = \begin{cases} \delta_q^p & \forall\, p, q \in \mathbb{V} \\ \delta_q^p - \gamma_q^p & \forall\, p, q \in \mathbb{A}\,. \\ 0 & \text{otherwise} \end{cases} \tag{3.98}$$

Density cumulants are only nonzero when all indices are active orbitals, that is, $\lambda_{rs\cdots}^{pq\cdots} = 0$ if any of the indices $p, q, r, s, \cdots$ is not active.

3.C.2 Orbital Energies

For convenience, we first evaluate the derivatives of $\check{f}_i^a$ and $\mathcal{R}_s(\Delta_{ab\cdots}^{ij\cdots})$ with respect to semicanonical orbital energies:

$$\frac{\partial \check{f}_i^a}{\partial \epsilon_p} = \sum_j^{\mathbb{H}} \sum_b^{\mathbb{P}} t_{ab}^{ij} (\gamma_j^p \delta_p^b - \gamma_p^b \delta_j^p), \tag{3.99}$$

$$\frac{\partial \mathcal{R}_s(\Delta_{ab\cdots}^{ij\cdots})}{\partial \epsilon_p} = \left[2se^{-s(\Delta_{ab\cdots}^{ij\cdots})^2} - \mathcal{R}_s(\Delta_{ab\cdots}^{ij\cdots})/\Delta_{ab\cdots}^{ij\cdots}\right] \mathcal{D}_{ab\cdots}^{ij\cdots}, \tag{3.100}$$

where $\mathcal{D}_{ab\cdots}^{ij\cdots} = \partial \Delta_{ab\cdots}^{ij\cdots}/\partial \epsilon_p = \delta_p^i + \delta_p^j + \cdots - \delta_a^p - \delta_b^p - \cdots$. Using these expressions, the partial derivatives in Eq. (3.57) are calculated as follows:

$$\frac{\partial \mathcal{T}_1}{\partial \epsilon_p} = \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \tau_a^i \left[\mathcal{R}_s(\Delta_a^i) \frac{\partial \check{f}_i^a}{\partial \epsilon_p} + \check{f}_i^a \frac{\partial \mathcal{R}_s(\Delta_a^i)}{\partial \epsilon_p}\right], \tag{3.101}$$

$$\frac{\partial \mathcal{T}_2}{\partial \epsilon_p} = \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \tau_{ab}^{ij} v_{ij}^{ab} \frac{\partial \mathcal{R}_s(\Delta_{ab}^{ij})}{\partial \epsilon_p}, \tag{3.102}$$

$$\frac{\partial \tilde{\mathcal{H}}_1}{\partial \epsilon_p} = \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \kappa_i^a (\partial \check{f}_i^a/\partial \epsilon_p) - \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \kappa_i^a t_a^i \mathcal{D}_a^i, \tag{3.103}$$

$$\frac{\partial \tilde{\mathcal{H}}_2}{\partial \epsilon_p} = -\frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \kappa_{ij}^{ab} t_{ab}^{ij} \mathcal{D}_{ab}^{ij}. \tag{3.104}$$

3.C.3 Orbital Rotations

The derivatives of the Lagrangian amplitude contributions with respect to orbital rotations are given by

$$\frac{\partial \mathcal{T}_1}{\partial \vartheta_p^q} = \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \tilde{\tau}_a^i \frac{\partial f_i^a}{\partial \vartheta_p^q}, \tag{3.105}$$

$$\frac{\partial \mathcal{T}_2}{\partial \vartheta_p^q} = \frac{1}{4} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \tilde{\tau}_{ab}^{ij} \frac{\partial v_{ij}^{ab}}{\partial \vartheta_p^q}, \tag{3.106}$$

where $\tilde{\tau}_a^i = \tau_a^i \mathcal{R}_s(\Delta_a^i)$ and $\tilde{\tau}_{ab}^{ij} = \tau_{ab}^{ij} \mathcal{R}_s(\Delta_{ab}^{ij})$. The orbital response from modified integrals also appears transparent:

$$\frac{\partial \tilde{\mathcal{H}}_1}{\partial \vartheta_p^q} = 2 \sum_i^{\mathbb{H}} \sum_a^{\mathbb{P}} \kappa_i^a \frac{\partial f_i^a}{\partial \vartheta_p^q}, \tag{3.107}$$

$$\frac{\partial \tilde{\mathcal{H}}_2}{\partial \vartheta_p^q} = \frac{1}{2} \sum_{ij}^{\mathbb{H}} \sum_{ab}^{\mathbb{P}} \kappa_{ij}^{ab} \frac{\partial v_{ij}^{ab}}{\partial \vartheta_p^q}. \tag{3.108}$$

The derivatives of the bare integrals with respect to orbital rotations are shown to be:

$$\begin{aligned} \frac{\partial f_r^s}{\partial \vartheta_p^q} &= f_p^s \delta_r^q + f_r^p \delta_q^s + \delta_{q\in\mathbb{C}} \big(v_{rp}^{sq} + v_{rq}^{sp} \big) \\ &\quad + \delta_{q\in\mathbb{A}} \Big[\sum_x^{\mathbb{A}} \big(v_{rp}^{sx} \gamma_x^q + v_{rx}^{sp} \gamma_q^x \big) \Big], \end{aligned} \tag{3.109}$$

$$\frac{\partial v_{ij}^{ab}}{\partial \vartheta_p^q} = v_{pj}^{ab} \delta_i^q + v_{ip}^{ab} \delta_j^q + v_{ij}^{pb} \delta_q^a + v_{ij}^{ap} \delta_q^b. \tag{3.110}$$

In Eq. (3.109), we have introduced the indicator function:

$$\delta_{p\in\mathbb{O}} := \begin{cases} 1 & \text{if } p \in \mathbb{O}, \\ 0 & \text{otherwise.} \end{cases} \tag{3.111}$$

The orbital response term from the reference energy E_0 is well-known from CASSCF orbital conditions:

$$\frac{\partial E_0}{\partial \vartheta_p^q} = 2 \big[\delta_{q\in\mathbb{C}} f_p^q + \delta_{q\in\mathbb{A}} \tilde{f}_p^q \big]. \tag{3.112}$$

Similar equations can be obtained for the CI term in $\mathcal{L}$:

$$\frac{\partial \mathcal{X}}{\partial \vartheta_p^q} = 2 \Big[\delta_{q\in\mathbb{C}} \sum_{uv}^{\mathbb{A}} v_{pu}^{qv} \tilde{\gamma}_v^u + \delta_{q\in\mathbb{A}} \tilde{f}_p^q(\gamma \to \tilde{\gamma}) \Big], \tag{3.113}$$

where $\tilde{f}_p^q(\gamma \to \tilde{\gamma})$ holds a similar form of Eq. (3.28) with the pRDMs γ [Eq. (3.5)] replaced by the corresponding mRDMs $\tilde{\gamma}$ [Eq.(3.44)].

Finally, for the Lagrangian term on orbital constraints $\mathcal{F}$, we have

$$\frac{\partial \mathcal{F}}{\partial \vartheta_p^q} = \sum_{rs}^{\mathbb{G}} \zeta_r^s \Big(\big[1 - \overline{\delta_{r\in\mathbb{A}} \delta_{s\in\mathbb{V}}} \big] \frac{\partial f_r^s}{\partial \vartheta_p^q} - \overline{\delta_{r\in\mathbb{A}} \delta_{s\in\mathbb{D}}} \frac{\partial \tilde{f}_r^s}{\partial \vartheta_p^q} \Big), \tag{3.114}$$

where we adopt the short-hand notation $\overline{\delta_{r\in\mathbb{A}}\delta_{s\in\mathbb{V}}} = \delta_{r\in\mathbb{A}}\delta_{s\in\mathbb{V}}+\delta_{s\in\mathbb{A}}\delta_{r\in\mathbb{V}}$ and $\mathbb{D} \equiv \mathbb{C}\cup\mathbb{V}$. The derivatives $\partial\tilde{f}^u_r/\partial\vartheta^q_p$ in Eq. (3.114) are worked out to be

$$\begin{aligned}\frac{\partial\tilde{f}^u_r}{\partial\vartheta^q_p} = \tilde{f}^u_p\delta^q_r + \delta_{q\in\mathbb{C}}\Big[\sum_v^{\mathbb{A}}(v^{vq}_{rp}+v^{vp}_{rq})\gamma^u_v\Big] \\ + \delta_{q\in\mathbb{A}}\Big[\bar{f}^p_r\gamma^u_q + \sum_{xy}^{\mathbb{A}}\Big(\frac{1}{2}v^{xy}_{rp}\gamma^{uq}_{xy} + v^{py}_{rx}\gamma^{ux}_{qy}\Big)\Big].\end{aligned} \tag{3.115}$$

3.C.4 CI Coefficients

We now evaluate all terms in Eq. (3.63). It is straightforward to see that the $\partial\mathcal{X}/\partial c_I$ term yields:

$$\frac{\partial\mathcal{X}}{\partial c_I} = \sum_J^{\mathcal{M}_0}\xi_J\left\langle\Phi_J|\hat{H}^{\mathrm{a}}|\Phi_I\right\rangle, \tag{3.116}$$

where the active part of the bare Hamiltonian is defined by

$$\hat{H}^{\mathrm{a}} = \sum_{uv}^{\mathbb{A}}\bar{f}^v_u\hat{a}^u_v + \frac{1}{4}\sum_{uvxy}^{\mathbb{A}}v^{xy}_{uv}\hat{a}^{uv}_{xy} \tag{3.117}$$

The derivatives of $\mathcal{E}+\mathcal{T}_1+\tilde{\mathcal{H}}_1+\mathcal{F}$ with respect to c_I can generally be written as

$$\begin{aligned}\frac{\partial}{\partial c_I}(\mathcal{E}+\mathcal{T}_1+\tilde{\mathcal{H}}_1+\mathcal{F}) = \sum_{uv}^{\mathbb{A}}g^u_v\frac{\partial\gamma^v_u}{\partial c_I} + \sum_{uvxy}^{\mathbb{A}}g^{uv}_{xy}\frac{\partial\gamma^{xy}_{uv}}{\partial c_I} \\ + \sum_{uvwxyz}^{\mathbb{A}}g^{uvw}_{xyz}\frac{\partial\gamma^{xyz}_{uvw}}{\partial c_I}.\end{aligned} \tag{3.118}$$

Here, the effective integrals are given by

$$\begin{aligned}g^u_v = \bar{f}^u_v + \frac{\partial E^{(2)}}{\partial\gamma^v_u} + \sum_i^{\mathbb{H}}\sum_a^{\mathbb{P}}\left[\alpha^a_i v^{au}_{iv} + (\tilde{\tau}^i_a+\kappa^a_i)\Delta^v_u t^{iu}_{av}\right] \\ + \sum_{rs}^{\mathbb{G}}\left[1-\overline{\delta_{r\in\mathbb{A}}\delta_{s\in\mathbb{V}}}\right]\zeta^s_r v^{su}_{rv} - 2\sum_r^{\mathbb{D}}\zeta^u_r\bar{f}^r_v,\end{aligned} \tag{3.119}$$

$$g^{uv}_{xy} = \frac{1}{4}v^{uv}_{xy} + \frac{\partial E^{(2)}}{\partial\gamma^{xy}_{uv}} - \sum_r^{\mathbb{D}}\zeta^u_r v^{rv}_{xy}, \tag{3.120}$$

$$g^{uvw}_{xyz} = \frac{1}{4}\sum_m^{\mathbb{C}}\tilde{h}^{uv}_{mz}t^{mw}_{xy} - \frac{1}{4}\sum_e^{\mathbb{V}}\tilde{h}^{ew}_{xy}t^{uv}_{ez}. \tag{3.121}$$

In Eqs. (3.119) and (3.120), the partial derivatives of $E^{(2)}$ with respect to γ_u^v and γ_{uv}^{xy} are given by:

$$\begin{aligned}
\frac{\partial E^{(2)}}{\partial \gamma_u^v} = &+ \sum_{e}^{\mathbb{V}} \tilde{h}_v^e t_e^u - \sum_{m}^{\mathbb{C}} \tilde{h}_m^u t_v^m \\
&+ \frac{1}{2} \sum_{ab}^{\mathbb{P}} \sum_{m}^{\mathbb{C}} \tilde{h}_{mv}^{ab} t_{ab}^{mu} - \frac{1}{2} \sum_{a}^{\mathbb{P}} \sum_{mn}^{\mathbb{C}} \tilde{h}_{mn}^{au} t_{av}^{mn} \\
&+ \sum_{xy}^{\mathbb{A}} \gamma_x^y \Big[\mathcal{P}(ux) \sum_{m}^{\mathbb{C}} \tilde{h}_m^x t_{vy}^{um} - \mathcal{P}(vy) \sum_{e}^{\mathbb{V}} \tilde{h}_v^e t_{ey}^{ux} \Big] \\
&+ \sum_{xy}^{\mathbb{A}} \gamma_x^y \Big[\mathcal{P}(vy) \sum_{m}^{\mathbb{C}} \tilde{v}_{my}^{ux} t_v^m - \mathcal{P}(ux) \sum_{e}^{\mathbb{V}} \tilde{h}_{vy}^{ex} t_e^u \Big] \\
&- \sum_{xy}^{\mathbb{A}} \gamma_x^y \sum_{a}^{\mathbb{P}} \sum_{m}^{\mathbb{C}} (\tilde{h}_{mv}^{au} t_{ay}^{mx} + \tilde{h}_{my}^{ax} t_{av}^{mu}) \\
&+ \frac{1}{2} \sum_{wxyz}^{\mathbb{A}} \lambda_{wx}^{yz} \sum_{e}^{\mathbb{V}} (\tilde{h}_{yv}^{eu} t_{ez}^{xw} + \tilde{h}_{yz}^{ex} t_{ev}^{uw}) \\
&- \frac{1}{2} \sum_{wxyz}^{\mathbb{A}} \lambda_{wx}^{yz} \sum_{m}^{\mathbb{C}} (\tilde{h}_{mz}^{xw} t_{yv}^{mu} + \tilde{h}_{mv}^{uw} t_{yz}^{mx}) \\
&+ \sum_{wxyz}^{\mathbb{A}} \gamma_y^z \gamma_w^x \Big(\sum_{m}^{\mathbb{C}} \tilde{h}_{mz}^{uy} t_{xv}^{mw} - \sum_{e}^{\mathbb{V}} \tilde{h}_{xv}^{ew} t_{ez}^{uy} \Big),
\end{aligned} \tag{3.122}$$

$$\begin{aligned}
\frac{\partial E^{(2)}}{\partial \gamma_{uv}^{xy}} = &+ \frac{1}{2} \sum_{e}^{\mathbb{V}} \tilde{h}_x^e t_{ey}^{uv} - \frac{1}{2} \sum_{m}^{\mathbb{C}} \tilde{h}_m^v t_{xy}^{um} \\
&+ \frac{1}{2} \sum_{e}^{\mathbb{V}} \tilde{h}_{xy}^{ev} t_e^u - \frac{1}{2} \sum_{m}^{\mathbb{C}} \tilde{v}_{my}^{uv} t_x^m \\
&+ \frac{1}{8} \sum_{ab}^{\mathbb{P}} \tilde{h}_{xy}^{ab} t_{ab}^{uv} + \frac{1}{8} \sum_{mn}^{\mathbb{C}} \tilde{h}_{mn}^{uv} t_{xy}^{mn} + \sum_{a}^{\mathbb{P}} \sum_{m}^{\mathbb{C}} \tilde{h}_{mx}^{au} t_{ay}^{mv} \\
&+ \frac{1}{2} \sum_{wz}^{\mathbb{A}} \gamma_w^z \sum_{e}^{\mathbb{V}} (\tilde{h}_{yz}^{ew} t_{ex}^{uv} + \tilde{h}_{xy}^{eu} t_{ez}^{vw}) \\
&- \frac{1}{2} \sum_{wz}^{\mathbb{A}} \gamma_w^z \sum_{m}^{\mathbb{C}} (\tilde{h}_{mx}^{uv} t_{yz}^{mw} + \tilde{h}_{mz}^{vw} t_{xy}^{mu}).
\end{aligned} \tag{3.123}$$

The RDM derivatives in Eq. (3.118) are generically written as

$$\frac{\partial \gamma_{xy\ldots}^{uv\ldots}}{\partial c_I} = \langle \Phi_I | \hat{a}_{xy\ldots}^{uv\ldots} + \hat{a}_{uv\ldots}^{xy\ldots} | \Psi_0 \rangle \,. \tag{3.124}$$

As such, the one- and two-body terms of Eq. (3.118) can be computed using a standard CI sigma build with the revised integrals g^u_v and g^{uv}_{xy}. In this work, we store the derivatives $\partial\gamma^{xyz}_{uvw}/\partial c_I$ appeared in Eq. (3.118), which requires further optimizations in the future.

Finally, for the $\mathcal{Y}$ term, we have

$$\partial\mathcal{Y}/\partial c_I = -2\iota c_I. \tag{3.125}$$

The multiplier ι can be easily obtained from Eq. (3.65) using Eq. (3.118):

$$\iota = \sum_{uv}^{\mathbb{A}} g^u_v \gamma^v_u + \sum_{uvxy}^{\mathbb{A}} g^{uv}_{xy}\gamma^{xy}_{uv} + \sum_{uvwxyz}^{\mathbb{A}} g^{uvw}_{xyz}\gamma^{xyz}_{uvw}, \tag{3.126}$$

using the fact that $\gamma^{uv\cdots}_{xy\cdots} = \frac{1}{2}\sum_I^{\mathcal{M}_0} c_I \partial\gamma^{uv\cdots}_{xy\cdots}/\partial c_I$. Note that ι depends on the orbital multipliers ζ^q_p and changes every iteration of solving the linear system [Eq. (3.66)].

3.D The Response Equation for the Orbital and CI Coefficients

We are now equipped to show all the components of the coupled linear system [Eq. (3.66)]. The four blocks of the coefficient matrix $\mathbf{A}$ can be written in the partial derivative form:

$$A^{\mathrm{oo}}_{pq,rs} = \frac{\partial}{\partial\zeta^s_r}\mathcal{P}(pq)\left(\frac{\partial\mathcal{F}}{\partial\vartheta^q_p}\right) = \frac{\partial F^s_r}{\partial\vartheta^q_p} - \frac{\partial F^s_r}{\partial\vartheta^p_q}, \tag{3.127}$$

$$A^{\mathrm{oc}}_{pq,J} = \frac{\partial}{\partial\xi_J}\mathcal{P}(pq)\left(\frac{\partial\mathcal{X}}{\partial\vartheta^q_p}\right) = \frac{\partial X_J}{\partial\vartheta^q_p} - \frac{\partial X_J}{\partial\vartheta^p_q}, \tag{3.128}$$

$$A^{\mathrm{co}}_{I,rs} = \frac{\partial}{\partial\zeta^s_r}\frac{\partial}{\partial c_I}\mathcal{F} = \frac{\partial F^s_r}{\partial c_I}, \tag{3.129}$$

$$A^{\mathrm{cc}}_{I,J} = \frac{\partial}{\partial\xi_J}\frac{\partial}{\partial c_I}\mathcal{X} = \frac{\partial X_J}{\partial c_I}. \tag{3.130}$$

As such, the orbital [Eq. (3.62)] and CI [Eq. (3.64)] response equations can be written as

$$\sum_{rs}^{\mathbb{G}} A^{\mathrm{oo}}_{pq,rs}\zeta^s_r + \sum_{J}^{\mathcal{M}_0} A^{\mathrm{oc}}_{pq,J}\xi_J = b^{\mathrm{o}}_{pq}, \tag{3.131}$$

$$\sum_{rs}^{\mathbb{G}} A^{\mathrm{co}}_{I,rs}\zeta^s_r + \sum_{J}^{\mathcal{M}_0} A^{\mathrm{cc}}_{I,J}\xi_J = b^{\mathrm{c}}_{I}. \tag{3.132}$$

The block elements of the $\mathbf{b}$ vector on the right-hand-side of Eqs. (3.131) and (3.132) are given by

$$b^{\mathrm{o}}_{pq} = -\mathcal{P}(pq)\Big(\frac{\partial}{\partial\vartheta^q_p}\big[E_0 + \sum_{n=1}^{2}(\mathcal{T}_n + \mathcal{H}_n)\big]\Big), \tag{3.133}$$

$$b^{\mathrm{c}}_{I} = -\frac{\partial}{\partial c_I}(\mathcal{E} + \mathcal{T}_1 + \mathcal{H}_1 + \mathcal{Y}). \tag{3.134}$$

For b^{o}_{pq}, all components of Eq. (3.133) have been reported in Sec. 3.C.3, specifically Eqs. (3.105)–(3.112). The expression of b^{c}_{I} can be obtained using Eqs. (3.118)–(3.121) by omitting the ζ^s_r contributions in Eqs. (3.119) and (3.120).

We may further express $A^{\mathrm{oo}}_{pq,rs}$ [Eq. (3.127)] in terms of f^s_r and $\tilde{f}^s_r$, resulting in the following cases:

$$\begin{aligned} A^{\mathrm{oo}}_{pq,rs} =& \mathcal{P}(pq)\big(\big[1 - \overline{\delta_{r\in\mathbb{V}}\delta_{s\in\mathbb{A}}}\big](\partial f^s_r/\partial\vartheta^q_p)\big) \\ &- \mathcal{P}(pq)\big[\overline{\delta_{r\in\mathbb{D}}\delta_{s\in\mathbb{A}}}(\partial\tilde{f}^s_r/\partial\vartheta^q_p)\big]. \end{aligned} \tag{3.135}$$

The partial derivatives appeared in Eq. (3.135) are reported in Eqs. (3.109) and (3.115). Simplifications may be achieved by utilizing the CASSCF semicanonical orbital constraint [Eqs. (3.27) and (3.32)]. For example, when all $p, q, r, s \in \mathbb{C}$, the expression of $A^{\mathrm{oo}}_{pq,rs}$ is simply

$$A^{\mathrm{oo}}_{pq,rs} = \Delta^p_q\big(\delta^q_r\delta^s_p + \delta^p_r\delta^s_q\big), \quad \forall\, p, q, r, s \in \mathbb{C}. \tag{3.136}$$

In Eq. (3.128), the partial derivatives $\partial X_J/\partial\vartheta_p^q$ yield:

$$\Big(\frac{\partial X_J}{\partial\vartheta_p^q}\Big)^{\perp} = \delta_{q\in\mathbb{A}}\Big(2\sum_{v}^{\mathbb{A}}\bar{f}_p^v\frac{\partial\tilde{\gamma}_v^q}{\partial\xi_J} + \sum_{vxy}^{\mathbb{A}}v_{pv}^{xy}\frac{\partial\tilde{\gamma}_{xy}^{qv}}{\partial\xi_J}\Big) + \delta_{q\in\mathbb{C}}\Big(2\sum_{uv}^{\mathbb{A}}v_{pu}^{qv}\frac{\partial\tilde{\gamma}_v^u}{\partial\xi_J}\Big), \tag{3.137}$$

where $\partial\tilde{\gamma}_{xy\cdots}^{uv\cdots}/\partial\xi_J = \langle\Phi_J|\hat{a}_{xy\cdots}^{uv\cdots}|\Psi_0\rangle$ and we only keep those components that are perpendicular to the CI vector $\mathbf{c}$. For the $A_{I,rs}^{\mathrm{co}}$ term [Eq. (3.129)], we write

$$A_{I,rs}^{\mathrm{co}} = \big[1-\overline{\delta_{r\in\mathbb{V}}\delta_{s\in\mathbb{A}}}\big]\Big(\frac{\partial f_r^s}{\partial c_I}\Big) - \overline{\delta_{r\in\mathbb{D}}\delta_{s\in\mathbb{A}}}\Big(\frac{\partial\tilde{f}_r^s}{\partial c_I}\Big). \tag{3.138}$$

Here, the derivatives of f_r^s and $\tilde{f}_r^s$ with respect to c_I are evaluated to be

$$\frac{\partial f_r^s}{\partial c_I} = \sum_{uv}^{\mathbb{A}}v_{rv}^{su}\frac{\partial\gamma_u^v}{\partial c_I}, \tag{3.139}$$

$$\frac{\partial\tilde{f}_r^s}{\partial c_I} = \sum_{v}^{\mathbb{A}}\bar{f}_p^v\frac{\partial\gamma_v^s}{\partial c_I} + \frac{1}{2}\sum_{vxy}^{\mathbb{A}}v_{pv}^{xy}\frac{\partial\gamma_{xy}^{qv}}{\partial c_I}. \tag{3.140}$$

Lastly, the $A_{I,J}^{\mathrm{cc}}$ term [Eq. (3.130)] can be easily derived using Eq. (3.116):

$$A_{I,J}^{\mathrm{cc}} = \langle\Phi_I|\hat{H}^{\mathrm{a}}|\Phi_J\rangle\,. \tag{3.141}$$

Bibliography

[1] Abbott, A. S.; Abbott, B. Z.; Turney, J. M.; Schaefer, H. F. Arbitrary-Order Derivatives of Quantum Chemical Methods via Automatic Differentiation. *J. Phys. Chem. Lett.* **2021**, *12*, 3232–3239.

[2] Iftimie, R.; Minary, P.; Tuckerman, M. E. Ab initio molecular dynamics: Concepts, recent developments, and future trends. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6654–6659.

[3] Park, J. W.; Shiozaki, T. On-the-Fly CASPT2 Surface-Hopping Dynamics. *J. Chem. Theory Comput.* **2017**, *13*, 3676–3683.

[4] Curchod, B. F. E.; Martínez, T. J. Ab Initio Nonadiabatic Quantum Molecular Dynamics. *Chem. Rev.* **2018**, *118*, 3305–3336.

[5] Pinski, P.; Neese, F. Communication: Exact analytical derivatives for the domain-based local pair natural orbital MP2 method (DLPNO-MP2). *J. Chem. Phys.* **2018**, *148*, 31101.

[6] Dornbach, M.; Werner, H.-J. Analytical energy gradients for local second-order Møller-Plesset perturbation theory using intrinsic bond orbitals. *Mol. Phys.* **2019**, *117*, 1252–1263.

[7] Ni, Z.; Wang, Y.; Li, W.; Pulay, P.; Li, S. Analytical Energy Gradients for the Cluster-in-Molecule MP2 Method and Its Application to Geometry Optimizations of Large Systems. *J. Chem. Theory Comput.* **2019**, *15*, 3623–3634.

[8] Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev.* **2012**, *112*, 182–243.

[9] Köhn, A.; Hanauer, M.; Mück, L. A.; Jagau, T.-C.; Gauss, J. State-specific multireference coupled-cluster theory. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 176–197.

[10] Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.

[11] Szalay, P. G.; Müller, T.; Gidofalvi, G.; Lischka, H.; Shepard, R. Multiconfiguration Self-Consistent Field and Multireference Configuration Interaction Methods and Applications. *Chem. Rev.* **2012**, *112*, 108–181.

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

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

[14] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

[15] Kozlowski, P. M.; Davidson, E. R. Considerations in constructing a multireference second-order perturbation theory. *J. Chem. Phys.* **1994**, *100*, 3672–3682.

[16] Werner, H.-J. Third-order multireference perturbation theory The CASPT3 method. *Mol. Phys.* **1996**, *89*, 645–661.

[17] Mahapatra, U. S.; Datta, B.; Mukherjee, D. Development of a size-consistent state-specific multireference perturbation theory with relaxed model-space coefficients. *Chem. Phys. Lett.* **1999**, *299*, 42–50.

[18] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of *n*-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

[19] Khait, Y. G.; Song, J.; Hoffmann, M. R. Explication and revision of generalized Van Vleck perturbation theory for molecular electronic structure. *J. Chem. Phys.* **2002**, *117*, 4133–4145.

[20] Szabados, Á.; Rolik, Z.; Tóth, G.; Surján, P. R. Multiconfiguration perturbation theory: Size consistency at second order. *J. Chem. Phys.* **2005**, *122*, 114104.

[21] others,, et al. Comparison of low-order multireference many-body perturbation theories. *J. Chem. Phys.* **2005**, *122*, 134105.

[22] Hoffmann, M. R.; Datta, D.; Das, S.; Mukherjee, D.; Szabados, A.; Rolik, Z.;

Surján, P. R. Comparative study of multireference perturbative theories for ground and excited states. *J. Chem. Phys.* **2009**, *131*, 204104.

[23] Evangelista, F. A.; Simmonett, A. C.; Schaefer, H. F.; Mukherjee, D.; Allen, W. D. A companion perturbation theory for state-specific multireference coupled cluster methods. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4728–4741.

[24] Sokolov, A. Y.; Guo, S.; Ronca, E.; Chan, G. K.-L. Time-dependent N-electron valence perturbation theory with matrix product state reference wavefunctions for large active spaces and basis sets: Applications to the chromium dimer and all-trans polyenes. *J. Chem. Phys.* **2017**, *146*, 244102.

[25] Giner, E.; Angeli, C.; Garniron, Y.; Scemama, A.; Malrieu, J.-P. A Jeziorski-Monkhorst fully uncontracted multi-reference perturbative treatment. I. Principles, second-order versions, and tests on ground state potential energy curves. *J. Chem. Phys.* **2017**, *146*, 224108.

[26] Roos, B. O.; Andersson, K. Multiconfigurational perturbation theory with level shift—the Cr_2 potential revisited. *Chem. Phys. Lett.* **1995**, *245*, 215–223.

[27] Forsberg, N.; Malmqvist, P.-Å. Multiconfiguration perturbation theory with imaginary level shift. *Chem. Phys. Lett.* **1997**, *274*, 196–204.

[28] Dyall, K. G. The choice of a zeroth-order Hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.* **1995**, *102*, 4909–4918.

[29] Valdemoro, C. Approximating the second-order reduced density matrix in terms of the first-order one. *Phys. Rev. A* **1992**, *45*, 4462–4467.

[30] Colmenero, F.; Valdemoro, C. Approximating q-order reduced density matrices

in terms of the lower-order ones. II. Applications. *Phys. Rev. A* **1993**, *47*, 979–985.

[31] Kutzelnigg, W.; Mukherjee, D. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.* **1997**, *107*, 432–449.

[32] Mazziotti, D. A. Contracted Schrödinger equation: Determining quantum energies and two-particle density matrices without wave functions. *Phys. Rev. A* **1998**, *57*, 4219–4234.

[33] Mazziotti, D. A. Variational reduced-density-matrix method using three-particle N-representability conditions with application to many-electron molecules. *Phys. Rev. A* **2006**, *74*, 032501.

[34] Shamasundar, K. R. Cumulant decomposition of reduced density matrices, multireference normal ordering, and Wicks theorem: A spin-free approach. *J. Chem. Phys.* **2009**, *131*, 174109.

[35] Misiewicz, J. P.; Turney, J. M.; Schaefer, H. F. Reduced Density Matrix Cumulants: The Combinatorics of Size-Consistency and Generalized Normal Ordering. *J. Chem. Theory Comput.* **2020**, *16*, 6150–6164.

[36] Kurashige, Y.; Chalupsky, J.; Lan, T. N.; Yanai, T. Complete active space second-order perturbation theory with cumulant approximation for extended active-space wavefunction from density matrix renormalization group. *J. Chem. Phys.* **2014**, *141*, 174111.

[37] Phung, Q. M.; Wouters, S.; Pierloot, K. Cumulant Approximated Second-Order Perturbation Theory Based on the Density Matrix Renormalization Group for Transition Metal Complexes: A Benchmark Study. *J. Chem. Theory Comput.* **2016**, *12*, 4352–4361.

[38] Zgid, D.; Ghosh, D.; Neuscamman, E.; Chan, G. K.-L. A study of cumulant approximations to *n*-electron valence multireference perturbation theory. *J. Chem. Phys.* **2009**, *130*, 194107.

[39] Guo, Y.; Sivalingam, K.; Neese, F. Approximations of density matrices in N-electron valence state second-order perturbation theory (NEVPT2). I. Revisiting the NEVPT2 construction. *J. Chem. Phys.* **2021**, *154*, 214111.

[40] Freitag, L.; Knecht, S.; Angeli, C.; Reiher, M. Multireference Perturbation Theory with Cholesky Decomposition for the Density Matrix Renormalization Group. *J. Chem. Theory Comput.* **2017**, *13*, 451–459.

[41] Sharma, S.; Chan, G. K.-L. Communication: A flexible multi-reference perturbation theory by minimizing the Hylleraas functional with matrix product states. *J. Chem. Phys.* **2014**, *141*, 111101.

[42] Sharma, S.; Knizia, G.; Guo, S.; Alavi, A. Combining internally contracted states and matrix product states to perform multireference perturbation theory. *J. Chem. Theory Comput.* **2017**, *13*, 488–498.

[43] Sokolov, A. Y.; Chan, G. K.-L. A time-dependent formulation of multi-reference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 064102.

[44] Sokolov, A. Y. Multi-reference algebraic diagrammatic construction theory for excited states: General formulation and first-order implementation. *J. Chem. Phys.* **2018**, *149*, 204113.

[45] Chatterjee, K.; Sokolov, A. Y. Extended Second-Order Multireference Algebraic Diagrammatic Construction Theory for Charged Excitations. *J. Chem. Theory Comput.* **2020**, *16*, 6343–6357.

[46] Nakano, H.; Hirao, K.; Gordon, M. S. Analytic energy gradients for multiconfigurational self-consistent field second-order quasidegenerate perturbation theory (MC-QDPT). *J. Chem. Phys.* **1998**, *108*, 5660–5669.

[47] Nakano, H. Quasidegenerate perturbation theory with multiconfigurational self-consistent-field reference functions. *J. Chem. Phys.* **1993**, *99*, 7983–7992.

[48] Nakano, H.; Otsuka, N.; Hirao, K. *Recent Advances in Multireference Methods*; World Scientific, 1999; pp 131–160.

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

[50] Park, J. W. Analytical First-Order Derivatives of Second-Order Extended Multiconfiguration Quasi-Degenerate Perturbation Theory (XMCQDPT2): Implementation and Application. *J. Chem. Theory Comput.* **2020**, *16*, 5562–5571.

[51] Park, J. W. Analytical Gradient Theory for Resolvent-Fitted Second-Order Extended Multiconfiguration Perturbation Theory (XMCQDPT2). *J. Chem. Theory Comput.* **2021**, *17*, 6122–6133.

[52] Celani, P.; Werner, H.-J. Analytical energy gradients for internally contracted second-order multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 5044–5057.

[53] Dudley, T. J.; Khait, Y. G.; Hoffmann, M. R. Molecular gradients for the second-order generalized Van Vleck variant of multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 651–660.

[54] Theis, D.; Khait, Y. G.; Hoffmann, M. R. GVVPT2 energy gradient using a Lagrangian formulation. *J. Chem. Phys.* **2011**, *135*, 044117.

[55] Khait, Y. G.; Theis, D.; Hoffmann, M. R. Nonadiabatic coupling terms for the GVVPT2 variant of multireference perturbation theory. *Chem. Phys.* **2012**, *401*, 88–94.

[56] Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **2011**, *135*, 081106.

[57] Győrffy, W.; Shiozaki, T.; Knizia, G.; Werner, H.-J. Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* **2013**, *138*, 104104.

[58] MacLeod, M. K.; Shiozaki, T. Communication: Automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory. *J. Chem. Phys.* **2015**, *142*, 051103.

[59] Park, J. W.; Shiozaki, T. Analytical Derivative Coupling for Multistate CASPT2 Theory. *J. Chem. Theory Comput.* **2017**, *13*, 2561–2570.

[60] Vlaisavljevich, B.; Shiozaki, T. Nuclear Energy Gradients for Internally Contracted Complete Active Space Second-Order Perturbation Theory: Multistate Extensions. *J. Chem. Theory Comput.* **2016**, *12*, 3781–3787.

[61] Park, J. W.; Al-Saadon, R.; Strand, N. E.; Shiozaki, T. Imaginary Shift in CASPT2 Nuclear Gradient and Derivative Coupling Theory. *J. Chem. Theory Comput.* **2019**, *15*, 4088–4098.

[62] Shiozaki, T. BAGEL: Brilliantly Advanced General Electronic-structure Library. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1331.

[63] Song, C.; Neaton, J. B.; Martínez, T. J. Reduced scaling formulation of CASPT2

analytical gradients using the supporting subspace method. *J. Chem. Phys.* **2021**, *154*, 014103.

[64] Song, C.; Martínez, T. J.; Neaton, J. B. A diagrammatic approach for automatically deriving analytical gradients of tensor hyper-contracted electronic structure methods. *J. Chem. Phys.* **2021**, *155*, 024108.

[65] Park, J. W. Analytical Gradient Theory for Strongly Contracted (SC) and Partially Contracted (PC) N-Electron Valence State Perturbation Theory (NEVPT2). *J. Chem. Theory Comput.* **2019**, *15*, 5417–5425.

[66] Park, J. W. Analytical Gradient Theory for Quasidegenerate N-Electron Valence State Perturbation Theory (QD-NEVPT2). *J. Chem. Theory Comput.* **2020**, *16*, 326–339.

[67] Nishimoto, Y. Analytic first-order derivatives of partially contracted *n*-electron valence state second-order perturbation theory (PC-NEVPT2). *J. Chem. Phys.* **2019**, *151*, 114103.

[68] Nishimoto, Y. Locating conical intersections using the quasidegenerate partially and strongly contracted NEVPT2 methods. *Chem. Phys. Lett.* **2020**, *744*, 137219.

[69] Nishimoto, Y. Analytic gradients for restricted active space second-order perturbation theory (RASPT2). *J. Chem. Phys.* **2021**, *154*.

[70] Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.

[71] Li, C.; Evangelista, F. A. Multireference Theories of Electron Correlation Based on the Driven Similarity Renormalization Group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245–273.

[72] Li, C.; Evangelista, F. A. Driven similarity renormalization group: Third-order multireference perturbation theory. *J. Chem. Phys.* **2017**, *146*, 124132.

[73] Zhang, T.; Li, C.; Evangelista, F. A. Improving the Efficiency of the Multireference Driven Similarity Renormalization Group via Sequential Transformation, Density Fitting, and the Noninteracting Virtual Orbital Approximation. *J. Chem. Theory Comput.* **2019**, *15*, 4399–4414.

[74] Li, C.; Evangelista, F. A. Multireference driven similarity renormalization group: A second-order perturbative analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.

[75] (a) Li, C.; Evangelista, F. A. Towards numerically robust multireference theories: The driven similarity renormalization group truncated to one- and two-body operators. *J. Chem. Phys.* **2016**, *144*, 164114.

[76] Li, C.; Evangelista, F. A. Spin-free formulation of the multireference driven similarity renormalization group: A benchmark study of first-row diatomic molecules and spin-crossover energetics. *J. Chem. Phys.* **2021**, *155*, 114111.

[77] Hannon, K. P.; Li, C.; Evangelista, F. A. An integral-factorized implementation of the driven similarity renormalization group second-order multireference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 204111.

[78] Schriber, J. B.; Hannon, K. P.; Li, C.; Evangelista, F. A. A Combined Selected Configuration Interaction and Many-Body Treatment of Static and Dynamical Correlation in Oligoacenes. *J. Chem. Theory Comput.* **2018**, *14*, 6295–6305.

[79] Khokhlov, D.; Belov, A. Toward an Accurate Ab Initio Description of Low-Lying Singlet Excited States of Polyenes. *J. Chem. Theory Comput.* **2021**, *17*, 4301–4315.

[80] Helgaker, T.; Jørgensen, P. Analytical calculation of geometrical derivatives in molecular electronic structure theory. *Adv. Quantum Chem.* **1988**, *19*, 183–245.

[81] Helgaker, T. In *Encyclopedia of Computational Chemistry*; Schleyer, P. v. R., Allinger, N. L., Clark, T., Gasteiger, J., Kollman, P. A., Schaefer, H. F., Schreiner, P. R., Eds.; John Wiley & Sons, Ltd: Chichester, UK, 1998; pp 1157–1169.

[82] Wang, S.; Li, C.; Evangelista, F. A. Analytic gradients for the single-reference driven similarity renormalization group second-order perturbation theory. *J. Chem. Phys.* **2019**, *151*, 044118.

[83] Handy, N. C.; Schaefer, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[84] Yamaguchi, Y.; Osamura, Y.; Goddard, J. D.; Schaefer, H. F. *A New Dimension to Quantum Chemistry: Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory*; Oxford University Press, 1994.

[85] Osamura, Y.; Yamaguchi, Y.; Schaefer, H. F. Generalization of analytic configuration interaction (CI) gradient techniques for potential energy hypersurfaces, including a solution to the coupled perturbed Hartree–Fock equations for multiconfiguration SCF molecular wave functions. *J. Chem. Phys.* **1982**, *77*, 383–390.

[86] Rice, J. E.; Amos, R. D. On the efficient evaluation of analytic energy gradients. *Chem. Phys. Lett.* **1985**, *122*, 585–590.

[87] Levchenko, S. V.; Wang, T.; Krylov, A. I. Analytic gradients for the spin-conserving and spin-flipping equation-of-motion coupled-cluster models with single and double substitutions. *J. Chem. Phys.* **2005**, *122*, 224106.

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

[89] Werner, H.-J.; Knowles, P. J. A comparison of variational and non-variational internally contracted multiconfiguration-reference configuration interaction calculations. *Theor. Chim. Acta.* **1990**, *78*, 175–187.

[90] Forte, a suite of quantum chemistry methods for strongly correlated electrons. For current version see https://github.com/evangelistalab/forte, 2021.

[91] Smith, D. G. A. et al. Psi4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.

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

[93] Woon, D. E.; Dunning, T. H. Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon. *J. Chem. Phys.* **1995**, *103*, 4572–4585.

[94] Mahapatra, U. S.; Datta, B.; Mukherjee, D. A state-specific multi-reference coupled cluster formalism with molecular applications. *Mol. Phys.* **1998**, *94*, 157–171.

[95] Evangelista, F. A.; Prochnow, E.; Gauss, J.; Schaefer, H. F. Perturbative triples corrections in state-specific multireference coupled cluster theory. *J. Chem. Phys.* **2010**, *132*, 074107.

[96] Werner, H.-J. et al. MOLPRO, version 2015.1, a package of *ab initio* programs. 2015; see http://www.molpro.net.

[97] Li, C.; Evangelista, F. A. Driven similarity renormalization group for excited states: A state-averaged perturbation theory. *J. Chem. Phys.* **2018**, *148*, 124106.

Chapter 4 Analytic Gradients for DF-DSRG-MRPT2

Chapter Abstract

We report an efficient implementation of the analytic energy gradients for the driven similarity renormalization group (DSRG) multireference second-order perturbation theory (MRPT2) using density fitting (DF). In our implementation, the four-index electron repulsion integrals (ERIs) are factorized with multiple three- and two-index integrals and the backtransformation of derivative-associated densities is changed correspondingly. Whilst the density fitting technique has negligible effects on electronic energies and optimized geometries, the viability of the analytic gradients has been improved for large systems with up to a thousand basis functions. We employed the new gradient regime to investigate the equilibrium geometry, harmonic vibrational frequencies, and infrared (IR) intensities of *p*-benzyne, for both singlet and triplet states. It is shown that the results are in good agreement with experimental data and those obtained from other electronic structure theories. Interestingly, we noticed that certain vibrational modes of *p*-benzyne using DSRG are numerically sensitive to the choice of basis sets and the approximation applied to the three-particle density cumulants.

4.1 INTRODUCTION

The analytic energy derivative theory has been a preoccupation in quantum chemistry for decades. Analytic gradients, first-order energy derivatives concerning nuclear coordinates, serve as necessary quantities for first-principle geometry optimization and ab initio molecular dynamics. The gradients for single reference (SR) methods

were well investigated and efficiently implemented in the last century. Despite the success achieved by SR electronic structure theories, there are certain strongly correlated chemical problems for which the SR approaches qualitatively fail, such as bond dissociations, organic diradicals, electronically excited states, and transition metal complexes. A multireference (MR) description is thus entailed to approximate both static and dynamic correlations.[1–3] Among various MR formalisms, the MR second-order perturbation theories (MRPT2) have gained great popularity ascribed to reliable accuracy and efficiency.[4–16] Prominent examples include complete active space second-order perturbation theory (CASPT2) and n-electron valence state second-order perturbation theory (NEVPT2), proven effective for strongly correlated closed-shell or high-spin open-shell systems.[6,10] Developing the analytic energy gradient theory for an MR theory is typically challenging, considering esoteric mathematical derivation and software implementations. The analytic derivative theories for CASPT2 and NEVPT2 and associated extensions were recently developed. Celani and Werner successfully employed the analytic energy gradients for the partially internally contracted CASPT2 to study the equilibrium geometries of Pyrrole with low-lying valence and Rydberg states.[17] Subsequently, the analytic gradient theory for the fully internally contracted CASPT2 was realized by MacLeod and Shiozaki using automatic code generation and was used to evaluate the vertical and adiabatic ionization potentials of the porphin molecule.[18] Shiozaki and co-workers further developed the analytic gradients for CASPT2 with density fitting or imaginary shifts,[19,20] multi-state (MS) CASPT2 and extend multi-state (XMS) CASPT2,[21,22] and explicitly correlated CASPT2 (CASPT2-F12).[23] These gradient methods are implemented in an open-source BAGEL package.[24] Park and Nishimoto contributed to the gradient theory for NEVPT2, including strongly contracted, partially contracted, and quasidegenerate NEVPT2 extensions.[25–28]

Aside from these remarkable achievements, the CASPT2 and NEVPT2, along with

other widely used MRPT theories, concomitantly suffer from a well-known intruder state problem.[29,30] The intruders are introduced when excitations excluded from the reference become near-degenerate with the determinants that lie within the reference state. This numerical issue could be improved by applying level shifts to energy denominators obtained from zeroth-order Hamiltonian,[31,32] but noticeable deviations in yielded spectroscopic parameters would be featured.[33,34] As discussed by Dyall,[35] another effective intruder-removal technique, which has gained success in MRPT and NEVPT variants, is to define a partially bielectronic zeroth-order Hamiltonian.[10,36] However, Zgid and co-workers found the intruders would be reintroduced if particle density cumulants are approximated.[37]

To address the intruder state problem, the driven similarity renormalization group (DSRG) formulated within the MR formalism was developed.[38–41] Within the unitary DSRG ansatz, a continuous unitary transformation applies to the bare Hamiltonian to decouple the reference and excited states, in terms of exponentially parametrized unitary operators tuned by a DSRG flow parameter s. Importantly, the DSRG many-body cluster amplitudes are numerically finite even with vanishing Møller-Plesset energy denominators. Among all MR-DSRG regimes, the DSRG-MRPT2 approach has gained the most attention ascribed to its relatively low computational cost. The accuracy of DSRG-MRPT2 is comparable to those of CASPT2 and NEVPT2, as reported in our previous benchmark studies. Aside from intrinsic immunity to intruder states, another significant advantage of DSRG-MRPT2 is the circumvention of four-body quantities, including four-body reduced density matrices (4-RDMs) and four-particle density cumulants as required in CASPT2 and NEVPT2. With factorized electron-repulsion integrals and approximate CASCI reference wave functions, DSRG-MRPT2 could be applied to large systems with up to 2000 basis functions and a large active space of CAS$(30e, 30o)$.[42,43]

In recent work, we have developed the analytic energy gradients for DSRG-

MRPT2 in the FORTE package.[44–46] It is also worth mentioning that the analytic gradient theory for the spin-free state-averaged DSRG-MRPT2 variant was independently proposed by Park.[47] Due to our implementations, the equilibrium geometries of the singlet and triplet state $p-$benzyne are similar to those obtained via other MRPT2 theories, and the adiabatic singlet-triplet splittings are slightly underestimated. An s-dependency study implies that correlation effects are appropriately captured when the flow parameter is selected around $1.0E_{\mathrm{h}}^{-2}$, in a good agreement with previous studies. The exclusion of three-body density cumulants only introduces negligible errors to energies and optimized structures of $p-$benzyne.

However, those preliminary results may be less convincing since a minimal active space of CAS$(2e, 2o)$ and a small cc-pCVDZ basis set were employed, only considering the correlation effects between unpaired electrons of p-benzyne. The memory bottleneck originates from the explicit storage and associated backtransformation of four-index quantities. Another concern is that it is difficult to determine whether an optimized geometry represents a stable or a transition state, due to the lack of Hessians in FORTE.

Herein, we report an efficient implementation of the analytic energy gradients for DSRG-MRPT2 with density fitting (DF) to address these issues. The DF technique has been rigorously studied and implemented to improve the performance of analytic gradients for both SR and MR formalisms.[20,48–54] Anticipated from previous studies, DF has negligible effects on the accuracy of equilibrium geometries and electronic energies. Combined with the functionality of computing numerical Hessians and analytic dipole moments, DSRG-MRPT2 could be used for spectroscopic analysis. Our DF gradient implementation is also generalized, and compatible with the potential gradient extensions for DSRG-MRPT2: a. a different reference wave function, such as the one from a generalized active space self-consistent-field (GASSCF) computation;[55,56] b. an alternative CASCI approximation formalism, such as the adaptive

CI method;[42,57] c. a different integral factorization technique, such as the Cholesky decomposition.[58] This provides great convenience for the future development of the analytic gradient theory for DSRG.

This paper proceeds as: In Sec.4.2, we briefly review the DSRG-MRPT2 formalism and introduce the formulation of DF analytic energy gradients. We conclude the theory section by discussing the computational scaling. In Sec.4.3, we report the vibrational frequencies and IR intensities at equilibrium structures of $p-$benzyne for both singlet and triplet states. Finally, we summarize this work in Sec.4.4 and discuss the prospects of DSRG-MRPT2.

4.2 THEORY

We first introduce the orbital notation in this work. A molecular spin orbital (MSO) ψ_p is formulated as a product of a spin function $\sigma_p(\omega)$ and a molecular orbital (MO) $\phi_p(\mathbf{r})$, where $\mathbf{r}$ and ω respectively denotes spatial and spin coordinates. Each MO can be linearly expanded via a set of nonorthogonal atomic orbitals (AOs) $\{\chi_\mu(\mathbf{r}), \mu = 1, 2, \ldots, N\}$,

$$\phi_p(\mathbf{r}) = \sum_{\mu}^{\mathrm{AO}} \chi_\mu(\mathbf{r})\, C_{\mu p} \tag{4.1}$$

where $C_{\mu p}$ denotes MO orbital coefficients.

The MSOs are partitioned into core ($\mathbb{C}$, doubly occupied), active ($\mathbb{A}$, partially occupied), and virtual ($\mathbb{V}$, unoccupied) subspaces by their occupation in the reference wavefunction. For convenience, we further formulate the composite hole ($\mathbb{H} = \mathbb{C} \cup \mathbb{A}$) and particle ($\mathbb{P} = \mathbb{A} \cup \mathbb{V}$) spaces. The index labeling of MSOs is summarized in Table 4.21. Distinctively, we denote AOs using Greek letters $\mu, \nu, \rho, \tau \cdots$.

Table 4.21: Index labeling of partitioned MSOs.

Space	Symbol	Size	Indices	Description
Core	$\mathbb{C}$	N_C	m, n, o	Occupied
Active	$\mathbb{A}$	N_A	u, v, w, x, y, z	Partially Occupied
Virtual	$\mathbb{V}$	N_V	e, f	Unoccupied
Hole	$\mathbb{H}$	N_H	i, j, k, l	$\mathbb{C} \cup \mathbb{A}$
Particle	$\mathbb{P}$	N_P	a, b, c, d	$\mathbb{A} \cup \mathbb{V}$
General	$\mathbb{G}$	N_G	p, q, r, s	$\mathbb{C} \cup \mathbb{A} \cup \mathbb{V}$

4.2.1 CASSCF reference

We use a CASSCF reference wave function Ψ_0,

$$|\Psi_0\rangle = \sum_I^{\mathcal{M}_0} c_I \, |\Phi_I\rangle \tag{4.2}$$

where ϕ_I denotes Slater determinants and c_I are associated CI coefficients. The n-particle reduced density matrix (n-pRDM) is formulated in terms of the CASSCF reference wave function and a series of fermionic creation ($\hat{a}_p^\dagger$) and annihilation ($\hat{a}_p$) operators,

$$\gamma_{ij\cdots}^{kl\cdots} = \langle \Psi_0 | \hat{a}_k^\dagger \hat{a}_l^\dagger \cdots \hat{a}_j \hat{a}_i | \Psi_0 \rangle \tag{4.3}$$

Given the 1- and 2-pRDMs, the one-electron integral h, and the antisymmetrized two-electron integrals v, the reference energy $E_0 = \langle \Psi_0 | \hat{H} | \Psi_0 \rangle$ possesses the form[Eq. (4.4)],

$$E_0 = \sum_m^{\mathbb{C}} h_m^m + \frac{1}{2} \sum_{mn}^{\mathbb{C}} v_{mn}^{mn} + \sum_{uv}^{\mathbb{A}} \left(h_v^u + \sum_m^{\mathbb{C}} v_{vm}^{um} \right) \gamma_v^u + \frac{1}{4} \sum_{uvxy}^{\mathbb{A}} v_{uv}^{xy} \gamma_{xy}^{uv} \tag{4.4}$$

4.2.2 DSRG-MRPT2 energy

Herein, we review the electronic energy within unrelaxed DSRG-MRPT2.[39] For convenience, all tensors and integrals are summarized in Table 4.22.

The total energy is the sum of a reference energy E_0, same to the CASSCF energy in this study, and a second-order correlation contribution $E^{(2)}(s)$,

$$E(s) = E_0 + E^{(2)}(s), \tag{4.5}$$

Table 4.22: Parameters within DSRG-MRPT2.

Name	Expression	Description
h^q_p	$\langle\psi_p\vert\hat{h}\vert\psi_q\rangle$	one-electron integrals
v^{rs}_{pq}	$\langle\psi_p\psi_q\Vert\psi_r\psi_s\rangle$	antisymmetrized electron repulsion integrals
f^q_p	Eq. (4.14)	generalized Fock matrix
$\gamma^{k\cdots}_{i\cdots}$	Eq. (4.3)	n-particle reduced density matrices
η^c_a	$\delta^c_a - \gamma^c_a$	one-hole reduced density matrices (1-hRDMs)
$\Delta^{i\cdots}_{a\cdots}$	Eq. (4.15)	Møller–Plesset denominators
$t^{i\cdots}_{a\cdots}(s)$	Eqs. (4.10) & (4.11)	first-order cluster amplitudes
$\tilde{h}^a_i(s)$	Eq. (4.12)	modified one-electron integrals
$\tilde{v}^{rs}_{pq}(s)$	Eq. (4.13)	antisymmetrized modified electron repulsion integrals

where $s \in [0, +\infty)$ is a DSRG flow parameter that regularizes DSRG quantities. The flow parameter controls the continuous unitary transformation applied to a bare Hamiltonian ($\hat{H}$), such that the DSRG effective Hamiltonian could be formulated,

$$\bar{H}(s) = \hat{U}^\dagger(s)\hat{H}\hat{U}(s) \tag{4.6}$$

Specifically, when s equals 0, $\hat{U}(0) = \hat{1}$ so that the DSRG Hamiltonian is identical to the bare Hamiltonian; and as s becomes larger, the reference is gradually decoupled from excited states.

The perturbative energy correction $E^{(2)}(s)$ can be expressed as a fully internal contraction,

$$E^{(2)}(s) = \langle\Psi_0|[\widetilde{H}^{(1)}(s), \hat{T}^{(1)}(s)]|\Psi_0\rangle \tag{4.7}$$

where $\widetilde{H}(s)$ and $\hat{T}(s)$ respectively denote the first-order effective Hamiltonian and first-order cluster operators, in terms of normal-ordered creation and annihilation operators $\{\hat{a}\}$,

$$\widetilde{H}^{(1)}(s) = \sum_i^{\mathbb{H}}\sum_a^{\mathbb{P}} \tilde{h}^{a,(1)}_i(s)\{\hat{a}^i_a\} + \frac{1}{4}\sum_{ij}^{\mathbb{H}}\sum_{ab}^{\mathbb{P}} \tilde{v}^{ab,(1)}_{ij}(s)\{\hat{a}^{ij}_{ab}\} \tag{4.8}$$

$$\hat{T}^{(1)}(s) = \sum_i^{\mathbb{H}}\sum_a^{\mathbb{P}} t^{i,(1)}_a(s)\{\hat{a}^a_i\} + \frac{1}{4}\sum_{ij}^{\mathbb{H}}\sum_{ab}^{\mathbb{P}} t^{ij,(1)}_{ab}(s)\{\hat{a}^{ab}_{ij}\} \tag{4.9}$$

For concision, the superscript "(1)" for the first-order quantities and the label "(s)" for s-dependent parameters are omitted in the following context. The DSRG cluster

amplitudes and the modified first-order integrals are expressed as [Eq (4.10) - (4.13)],

$$t_a^i = \left(f_a^i + \Delta_u^x t_{ax}^{iu} \gamma_u^x\right) \mathcal{R}_s(\Delta_a^i) \tag{4.10}$$

$$t_{ab}^{ij} = v_{ij}^{ab} \mathcal{R}_s(\Delta_{ab}^{ij}) \tag{4.11}$$

$$\tilde{h}_i^a = f_i^a \mathcal{P}_s(\Delta_a^i) + \Delta_u^x t_{ax}^{iu} \gamma_u^x \left[\mathcal{P}_s(\Delta_a^i) - 1\right] \tag{4.12}$$

$$\tilde{v}_{ij}^{ab} = v_{ij}^{ab} \mathcal{P}_s(\Delta_{ab}^{ij}) \tag{4.13}$$

where $\Delta_{ab\cdots}^{ij\cdots}$ are Møller-Plesset denominators in terms of MO energies ϵ_p, and f_q^p represents components of a generalized Fock matrix,

$$f_q^p = h_q^p + \sum_m^{\mathbb{C}} v_{qm}^{pm} + \sum_{uv}^{\mathbb{A}} v_{qv}^{pu} \gamma_u^v \tag{4.14}$$

$$\Delta_{ab\cdots}^{ij\cdots} = \epsilon_i + \epsilon_j + \cdots - \epsilon_a - \epsilon_b - \cdots \tag{4.15}$$

In DSRG, the cluster amplitudes ($t_{ab\cdots}^{ij\cdots}$) are antisymmetric to respective permutations of superscript and subscript indices, and the internal components are enforced to be zero, that is, $t_{uv\cdots}^{xy\cdots} = 0, \forall\, u, v, x, y, \cdots \in \mathbb{A}$. For convenience, we introduce regularized functions $\mathcal{R} = (1 - e^{-s\Delta^2})/\Delta$ and $\mathcal{P} = 1 + e^{-s\Delta^2}$. Given an arbitrary value of s, both $\mathcal{R}$ and $\mathcal{P}$ are numerically bounded even if the Møller-Plesset denominators are zero or near-zero, thus the associated DSRG parameters are robust to intruder states. Appropriate values of s, around $\sim 1\ E_\mathrm{h}^{-2}$ as discussed in previous studies,[39,44] must be chosen to circumvent excessive capture of dynamic electron correlations.

The DSRG-MRPT2 energy expressions are concluded in Table 4.23, in terms of DSRG cluster amplitudes, modified first-order integrals, 1-pRDM, 1-hRDM, and two- and three-particle density cumulants.[59,60]

4.2.3 Analytic gradients

In this section, we briefly discuss the analytic energy gradients for DSRG-MRPT2. Readers are encouraged to consult the original work for details.[44]

In unrelaxed DSRG-MRPT2, the MO coefficients $C_{\mu p}$ [Eq. (4.1)] and the CI coefficients c_I[Eq. (4.2)] optimized by CASSCF and CI are not variationally optimized

Table 4.23: Summary of DSRG-MRPT2 energy terms.

Term	Energy Expression
E_0	$f^q_p\gamma^p_q - \frac{1}{2}v^{rs}_{pq}\gamma^p_r\gamma^q_s + \frac{1}{4}v^{rs}_{pq}\lambda^{pq}_{rs}$
$E^{(2)}$	$+\tilde{h}^b_j t^i_a \gamma^j_i \gamma^a_b + \frac{1}{4}\tilde{v}^{cd}_{kl} t^{ij}_{ab} \gamma^k_i \gamma^l_j \eta^a_c \eta^b_d + \frac{1}{4}\left(\tilde{v}^{uv}_{iz} t^{iw}_{xy} + \tilde{v}^{wa}_{xy} t^{uv}_{az}\right)\lambda^{xyz}_{uvw}$ $+\frac{1}{2}\left(\tilde{v}^{ev}_{xy} t^u_e - \tilde{v}^{uv}_{my} t^m_x + \tilde{h}^e_x t^{uv}_{ey} - \tilde{h}^v_m t^{um}_{xy}\right)\lambda^{xy}_{uv}$ $+\frac{1}{8}\left(\tilde{v}^{cd}_{xy} t^{uv}_{ab} \eta^a_c \eta^b_d + \tilde{v}^{uv}_{kl} t^{ij}_{xy} \gamma^k_i \gamma^l_j + 8\tilde{v}^{vb}_{jx} t^{iu}_{ay} \gamma^j_i \eta^a_b\right)\lambda^{xy}_{uv}$

in DSRG-MRPT2. Therefore, evaluating derivatives of MO and CI coefficients regarding nuclear coordinates ($\partial C_{\mu p}/\partial R$ and $\partial c_I/\partial R$) is indispensable, which could be achieved by solving the coupled-perturbed(CP) CASSCF equations.[61,62] Unfortunately, the number of required CP-CASSCF equations significantly increases if a molecule comprises more atoms, making it intractable for large systems. The Z-vector approach, proposed by Handy and Schaefer, substitutes a set of CP equations with one equivalent system-size-independent response equation.[63,64]

Following the Z-vector routine, we employed the Lagrangian formulation, as carefully discussed by Helgaker and Jørgensen.[65] The Lagrangian for DSRG-MRPT2 could be rigorously formulated by incorporating all energy-relevant quantities, including the n-body cluster amplitudes ($\mathcal{T}_n$), the n-body modified integrals ($\tilde{\mathcal{H}}_n$), the semicanonical CASSCF orbitals ($\mathcal{F}$), the orthonormality of the MSOs ($\mathcal{W}$), and a CASCI reference ($\mathcal{X}$) subject to normalization ($\mathcal{Y}$). In summary, the DSRG-MRPT2 Lagrangian ($\mathcal{L}$) can be simplified as,

$$\mathcal{L} = \mathcal{E} + \sum_{n=1}^{2}(\mathcal{T}_n + \tilde{\mathcal{H}}_n) + \mathcal{F} + \mathcal{W} + \mathcal{X} + \mathcal{Y}, \tag{4.16}$$

where each component is a product of a specific DSRG constraint equation and an affiliated Lagrange multiplier. All these constraints and associated multipliers are conceptually manifested in Table 4.24.

The analytic gradients for DSRG-MRPT2 are thus equivalent to the first-order derivatives of the Lagrangian with respect to the nuclear displacements (R). Mean-

Table 4.24: Summary for the DSRG-MRPT2 Lagrangian constraints.

Term	Constraints	Multipliers	Description
$\mathcal{T}_n$	$T^{ij\cdots}_{ab\cdots}$	$\tau^{ij\cdots}_{ab\cdots}$	n-body cluster amplitudes
$\tilde{\mathcal{H}}_n$	$\tilde{H}^{ab\cdots}_{ij\cdots}$	$\kappa^{ab\cdots}_{ij\cdots}$	n-body modified integrals
$\mathcal{F}$	F^q_p	ζ^q_p	CASSCF semicanonical orbitals
$\mathcal{W}$	W^q_p	ω^q_p	orthonormal orbitals
$\mathcal{X}$	X_I	ξ_I	CI coefficients from CASCI
$\mathcal{Y}$	Y	ι	normalized CI coefficients

while the Lagrangian derivatives with respect to all other parameters are assumed zero.

$$\frac{\mathrm{d}E}{\mathrm{d}R} = \left.\frac{\partial \mathcal{L}}{\partial R}\right|_{\partial\mathcal{L}/\partial\mathcal{M}=0} = \sum_{pq}^{\mathbb{G}} \Gamma^p_q (h^p_q)^x + \sum_{pqrs}^{\mathbb{G}} \Gamma^{pq}_{rs} (v^{rs}_{pq})^x + \sum_{pq}^{\mathbb{G}} \omega^p_q (S^q_p)^x \tag{4.17}$$

where $(h^p_q)^x$, $(v^{rs}_{pq})^x$, $(S^q_p)^x$ are derivatives of one-electron integrals, antisymmetrized ERIs, and the MSO overlap integral. The associated terms Γ^p_q, Γ^{pq}_{rs}, ω^p_q denote the one-body, the two-body and the energy-weighted density matrices, respectively.

In addition, we discuss the analytic dipole moment, whose derivatives could further be used to investigate IR intensities. As showed by the Hellmann-Feyman theorem,[66] the energy derivatives regarding a perturbative parameter λ could be treated as the expectation value of the λ-dependent Hamiltonian derivatives,

$$\frac{dE_\lambda}{d\lambda} = \langle \psi_\lambda | \frac{d\hat{H}_\lambda}{d\lambda} | \psi_\lambda \rangle \tag{4.18}$$

We assume that λ_i represents the electrostatic field strength along the axis i, thus the analytic dipole moment could be written as a product of an one-body density matrix Γ^ρ_ν and an axis-dependent permanent dipole matrix $\bar{\mu}^i_{\nu\rho}$,

$$\mu_i = \frac{dE_\lambda}{d\lambda_i} = \langle \psi_\lambda | \frac{d(\hat{H} + \lambda_i \bar{\mu}^i_{\nu\rho} \hat{a}^\dagger_\rho \hat{a}_\nu)}{d\lambda_i} | \psi_\lambda \rangle = \sum_{\rho\nu}^{\mathrm{AO}} \Gamma^\rho_\nu \bar{\mu}^i_{\nu\rho} \tag{4.19}$$

It is worth mentioning that Γ^ρ_ν is the backtransformed density in AO basis,

$$\Gamma^\rho_\tau = \sum_{pq}^{\mathbb{G}} C_{\rho p} \Gamma^p_q C_{\tau q} \tag{4.20}$$

The analytic energy gradients and dipole moments for CASSCF could also be obtained, though inefficiently, by setting the DSRG flow parameter s to zero. The scheme to compute analytic gradients is conceptually shown in Fig. 4.21.

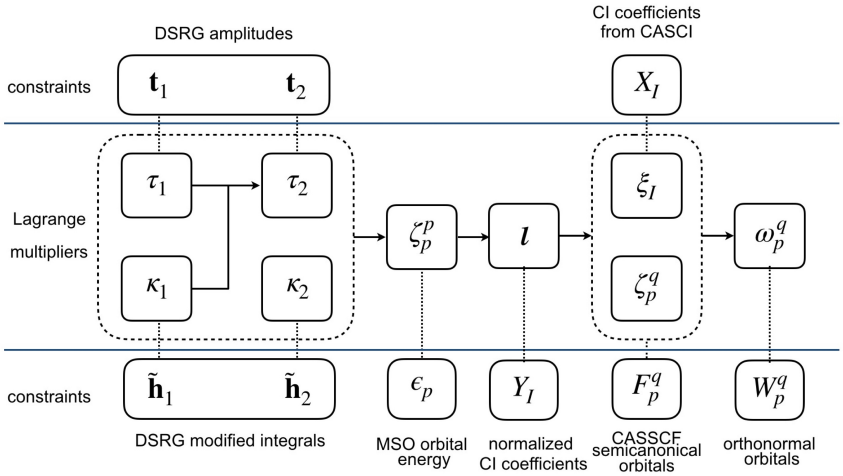

Figure 4.21: Schematic analytic gradient routine within DSRG-MRPT2.

4.2.4 Density-fitting

Density fitting (DF), also named the resolution of identity, is a technique to approximate conventional two-electron integrals as the repulsion between generalized electron densities ρ_{pq} using an auxiliary basis set $\bar{P}$,

$$(pq|rs) = (\psi_p\psi_q|\psi_r\psi_s) = (\rho_{pq}|\rho_{rs}) \tag{4.21}$$

$$\rho_{pq}(\mathbf{r}) \approx \sum_{\bar{P}} \bar{C}_{\bar{P}}^{pq} \chi_{\bar{P}}(\mathbf{r}) \tag{4.22}$$

where $\bar{C}_{\bar{P}}^{pq}$ are density fitting coefficients, and the Mulliken notation is used for ERIs. The advantage of DF approximation is twofold: a. explicit storage of large four-index quantities could be avoided; b. ERI-associated integral transformations or backtransformations become significantly cheaper, as they can be performed on three-index integrals. Specifically, using the DF regime proposed by Dunlap *et al.*,[67,68] we have,

$$(pq|rs) \approx \sum_{\bar{P}\bar{Q}} (pq|\bar{P})(\bar{P}|\bar{Q})^{-1}(\bar{Q}|rs) \tag{4.23}$$

When applying density fitting to formulate Z-vector equations, consistently treating all ERI-associated quantities and procedures is especially important. Among the changes, reformulating the DSRG-MRPT2 energy is straightforward. Specifically, all ERI terms involved in the reference energy (E_0) [Eq. (4.4)] and the DSRG parameters [$f^p_q, t^{ij}_{ab}, \tilde{v}^{ab}_{ij}$] [Eq. (4.14), (4.11), (4.13)] are factorized. In FORTE, the factorization is accomplished using distinctive three-index factor tensors $B^{\bar{Q}}_{pq}$,

$$(pq|rs) \approx \sum_{\bar{Q}}^{\mathcal{M}} B^{\bar{Q}}_{pq} B^{\bar{Q}}_{rs} \tag{4.24}$$

$$B^{\bar{Q}}_{pq} = \sum_{\bar{P}}^{\mathcal{M}} (pq|\bar{P})[\mathbf{J}^{-1/2}]_{\bar{P}\bar{Q}} \tag{4.25}$$

where $\mathcal{M}$ denotes the size of the auxiliary basis set $\chi_{\bar{P}}(\mathbf{r})$, $(pq|\bar{P})$ and $\mathbf{J}_{\bar{P}\bar{Q}} = (\bar{P}|\bar{Q})$ represent the three- and two-center integrals, respectively. Based on how the two-electron integrals are factorized [Eq. (4.24)], current DF gradient implementation in FORTE is intrinsically compatible to the Cholesky decomposition approximation.[58]

Aside from the substitution of ERIs, the integral backtransformation entails additional modifications. The second term in Eq. (4.17) could be separated into coulomb and exchange contributions,

$$\Gamma^{pq}_{rs}(v^{rs}_{pq})^x = \Gamma^{pq}_{rs}(pr|qs)^x - \Gamma^{pq}_{rs}(ps|qr)^x \tag{4.26}$$

For brevity, only the coulomb parts are discussed, and the exchange complement could be analogously analyzed.

$$\Gamma^{pq}_{rs}(pr|qs)^x = 2\widetilde{\Gamma}^{\bar{P}}_{pr}(pr|\bar{P})^x - \hat{\Gamma}^{\bar{P}\bar{Q}}\mathbf{J}^x_{\bar{P}\bar{Q}} \tag{4.27}$$

$$\widetilde{\Gamma}^{\bar{P}}_{pr} = [\mathbf{J}^{-1}]^x_{\bar{P}\bar{Q}}(\bar{Q}|qs)\Gamma^{pq}_{rs} \tag{4.28}$$

$$\hat{\Gamma}^{\bar{P}\bar{Q}} = (pr|\bar{P})\Gamma^{pq}_{rs}(\bar{Q}|qs) \tag{4.29}$$

where $\widetilde{\Gamma}$ and $\hat{\Gamma}$ denote the three- and two-center densities, respectively. Note that these densities are always represented in the AO basis, but the three-center densities solved from Z-vector equations are originally formulated with MO quantities.

Therefore, a backtransformation from MO to AO is required, shown as,

$$\widetilde{\Gamma}^{\bar{P}}_{\mu\nu} = \sum_{pr}^{\mathbb{G}} C_{\mu p} C_{\nu r} \widetilde{\Gamma}^{\bar{P}}_{pr} \tag{4.30}$$

Although density-fitting modifies the quantities and processes associated with Z-vector equations, we want to clarify that such changes have negligible accuracy effects on the resulting energy gradients. Importantly, the immunity to the intruder states is still guaranteed for all DSRG parameters, thus both the electronic energies and the energy gradients are guaranteed numerically robust.

4.3 RESULTS

In this work, various RI basis sets of Weigend were tested,[69] including the cc-pVDZ/JKFIT, the cc-pVTZ/JKFIT, and the cc-pVQZ/JKFIT. For brevity, they are denoted as DZ/JKFIT, TZ/JKFIT, and QZ/JKFIT in the following context. A consistent RI basis set is employed for both reference and correlation components in a single computation. We note that the relative errors by comparing the total energies using DSRG-MRPT2 with those computed from DF-DSRG-MRPT2 are mostly around 10^{-3} kcal·mol^{-1}. In addition, the relative errors for optimized bond lengths and bond angles are normally at the magnitude of 10^{-4} Åand 10^{-2} degrees, respectively. Importantly, these error comparison tests were conducted for systems for which the non-DF computations are affordable, such as the *p*-benzyne with a small active space CAS(2,2), and we lack such comparison for most large systems investigated in this study due to the memory limit. However, as expected from other studies,[20,43] the system-size-dependent DF relative errors are normally negligible.

The analytic gradients and the analytic dipole moments for unrelaxed DF-DSRG-MRPT2 were implemented in Forte, and the integral backtransformation of three-center densities [Eq. (4.28)] was implemented in the open-source quantum chemistry software Psi4 1.5.[70] The accuracy of the analytic gradients and dipole moments

were verified using the numerical complements got from five-point finite-difference computations. Specifically, a nuclear displacement of 0.005 a.u. and a perturbed electric field strength displacement of 0.001 were used, respectively. These validation tests were conducted on HF, H_4 with a C_1 symmetry, N_2, CO_2 and *p*-benzyne. Whilst high accuracy was achieved, we employed a timing test on the pentacene molecule with CAS(2,2) and DZ/JKFIT basis. It took a wall time of 2.3 hours to complete a single-point DF-DSRG-MRPT2 gradient computation, while the five-point finite-difference approach entailed a wall time of 157.4 hours under the same conditions.

For spectroscopic analysis, the Hessians and the derivatives of dipole moments with respect to nuclear coordinates are necessary, and they could be computed based on the finite difference method using energies at distorted geometries. However, such a routine would introduce three major issues: a. Computing Hessians from energies is extremely slow, especially for large systems; b. The Hessians are computed using appropriate gradients, downgrading the accuracy; 3, Distorting a reference geometry may break symmetries of associated orbitals, for example, normal MOs within an irreducible representation of a D_{2h} point group may become a lower-symmetry C_{2v} point group. To address these issues, we compute the Hessians from analytic gradients. Similarly, the derivatives of the dipole moments are obtained from the analytic dipole moments. The correctness of the resulting vibrational frequencies and IR intensities was validated using CO and CHFO molecules, with relative errors as around 0.1 cm^{-1} and 0.1 $km{\cdot}mol^{-1}$, respectively.

We particularly investigate the *p*-benzyne. The *p*-benzyne molecule is a biradical molecule with high stereoselectivity that could be used for drug designs.[71] A distinctive aspect of such a molecule is that its ground state is a singlet state, thus it could hardly be analyzed by experimental measurements.[72] Because of the multiconfigurational character of the biradical system, single-reference methods fail to yield accurate descriptions.

We first discuss the equilibrium structures of *p*-benzyne for singlet and triplet states, computed with DSRG-MRPT2, CASSCF, and CASPT2. The CASPT2 results were distinctively obtained via Molpro 2015.1,[73] and the DSRG or CASSCF results were yielded using Forte. The active space was set to CAS(8,8), including eight electrons in two *sp*2 and six *p* orbitals, and the basis was set to QZ/JKFIT. For DSRG-MRPT2, the DZ/JKFIT and TZ/JKFIT basis sets were also employed. We assumed that the equilibrium geometry was reached if the maximum component of the gradients was less than 2×10^{-6} a.u. The optimized structures of *p*-benzyne for singlet and triplet states are shown in Fig. 4.31.

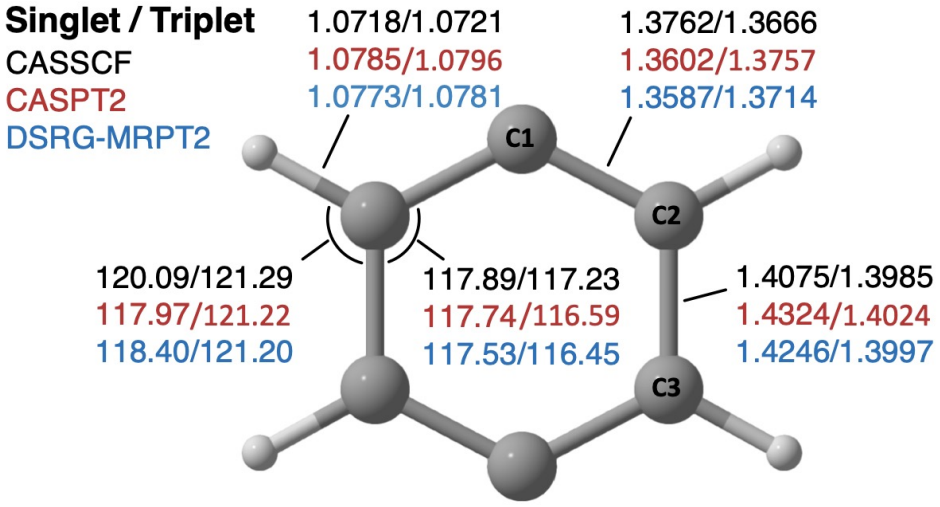

Figure 4.31: Equilibrium structures of *p*-benzyne for singlet and triplet states computed from DSRG-MRPT2, CASSCF, and CASPT2. The DSRG flow parameter was set to 1.0 E_{h}^{-2}.

For *p*-benzyne, it has been well-acknowledged that there are through-bond couplings between unpaired electrons, such as the spin polarization effects.[72] These interactions shorten the four CC(H) bonds and elongate the two (H)CC(H) bonds simultaneously. The bond length difference $\Delta r = r[(\mathrm{H})\mathrm{CC}(\mathrm{H})] - r[\mathrm{CC}(\mathrm{H})]$ could be treated as a biradical character indicator. The DSRG-MRPT2 optimization yields a small Δr value as 0.07 Å for the singlet *p*-benzyne, showing that the biradical properties are properly captured other than erroneously described as a closed-shell

singlet. In contrast, such biradical quantity Δr is reduced to 0.03Å for the triplet *p*-benzyne. Consistent with our previous study, the equilibrium geometries for both singlet and triplet states optimized by DSRG-MRPT2 are consistent with those of CASPT2 given a small basis DZ/JKFIT and a minimal active space CAS(2,2). The maximum bond length difference is 1.0 pm, for the (H)CC(H) bond of the singlet *p*-benzyne. Interestingly, if a larger active space is used while the small basis remains, the geometry changes are negligible for the singlet *p*-benzyne but become noticeable for triplet *p*-benzyne. When large basis and large active space are concomitantly used, the resulting geometries for both singlet and triplet *p*-benzyne greatly deviate from those using small basis and small active space. Under the new condition, two MRPT2 methods are still comparable for both singlet and triplet *p*-benzyne, though DSRG-MRPT2 slightly underestimates all bond lengths. The CASSCF results also imply that considering dynamic correlations would elongate the (H)CC and the (H)CC(H) bond, meanwhile shortening the (H)CC bond for both singlet and triplet *p*-benzyne. We may conclude that the DSRG-MRPT2 results are in good agreement with those of CASPT2, and the equilibrium geometries for the singlet state *p*-benzyne are comparatively more sensitive.

We further employed the approximate DSRG-MRPT2 formalism named cu-DSRG-MRPT2, following the terminology proposed by Zgid and co-workers,[37] in which the three-particle density cumulant terms are completely neglected. The equilibrium structures of singlet *p*-benzyne using DSRG-MRPT2 and cu-DSRG-MRPT2 are shown in Fig. 4.32. The cu-DSRG-MRPT2 optimized bond angles and bond lengths are in excellent agreement with those of the original DSRG-MRPT2, given various basis and active space settings. For example, the bond lengths deviate by maximally 0.8 pm for (H)CC(H), and the bond angles deviate by at most 0.8° for ∠HCC(H). Such negligible deviation indicates that cu-DSRG-MRPT2 could serve as a reliable DSRG-MRPT2 alternative for geometry optimizations, even with large active space

and a large basis. Interestingly, we also noticed that cu-DSRG-MRPT2 always overestimates the bond lengths for the CH and (H)CC bonds, meanwhile underestimates the (H)CC(H) bond.

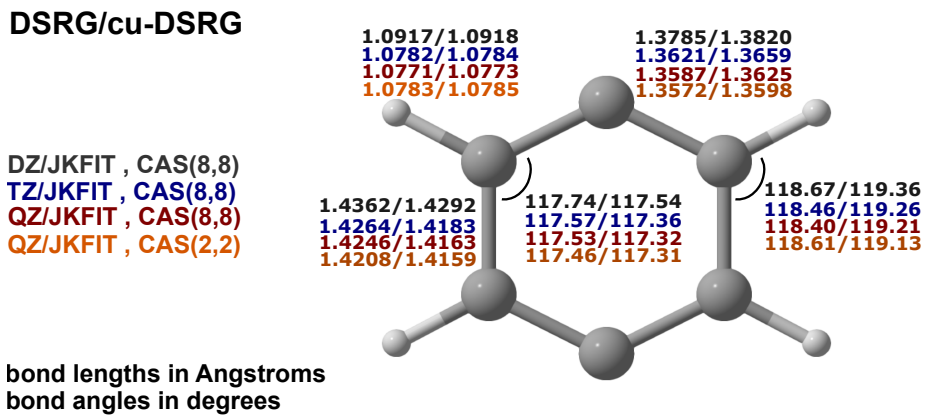

Figure 4.32: Equilibrium geometries of singlet p-benzyne optimized using DSRG-MRPT2 and cu-DSRG-MRPT2. The DSRG flow parameter was set to 1.0 E_{h}^{-2}.

The vibrational frequencies of the singlet p-benzyne with equilibrium geometries are shown in Table 4.31. For comparison, we listed the experimental data and those obtained from various approaches, including RHF, CCSD(T) with RHF or UHF orbitals, CASSCF, DSRG-MRPT2, and four-reference reduced multireference CCSD (4R RMR CCSD). Note that the RHF and CCSD(T) with the RHF reference both introduce imaginary frequencies, implying that the associated geometries are at transition states. If a UHF reference is employed instead, the CCSD(T) formalism would give reliable results, though it may be affected by spin contaminations from the triplet states. Meanwhile, the MR methods all yield real frequencies for all vibrational modes, meaning the associated geometries are stable. Comparing DSRG-MRPT2 with CASSCF, it could be seen that the dynamic correlations significantly affect certain modes with b_{3g}, a_g, b_{1u} and b_{2u} symmetries. For example, $\omega_2(a_g)$ differs by nearly 200 cm^{-1}. The DSRG-MRPT2 results are in good agreement with the experimental data, though $\omega_{20}(b_{2u})$ are overestimated by over 100 cm^{-1}. It could be seen that 4R RMR CCSD yields a better result for this vibrational mode, but its

Table 4.31: Vibrational frequencies (in cm^{-1}) of singlet *p*-benzyne obtained with various approaches, and experiment. The flow parameter of DSRG is set to 1.0 E_h^{-2}.

Frequency	Exp.[a]	RHF[b]	CCSD(T)[b]/RHF	CCSD(T)[b]/UHF	CASSCF[c]	DSRG[c]	4R RMR CCSD[d]
$\omega_1(a_g)$		3414	3264	3258	3349	3266	3236
$\omega_2(a_g)$		1379	1342	1498	1531	1329	1385
$\omega_3(a_g)$		1256	1174	1183	1221	1161	1167
$\omega_4(a_g)$	990	980	1018	1044	1042	1025	1023
$\omega_5(a_g)$	635	775	608	620	648	632	647
$\omega_6(a_u)$		1004	913	938	970	971	924
$\omega_7(a_u)$		489	417	407	435	435	429
$\omega_8(b_{1g})$		961i	349	771	775	709	692
$\omega_9(b_{2g})$		946	832	897	941	927	857
$\omega_{10}(b_{2g})$		310i	5612i	577	663	617	547
$\omega_{11}(b_{3g})$		3396	3250	3242	3332	3253	3217
$\omega_{12}(b_{3g})$		1869	1657	1683	1692	1670	1665
$\omega_{13}(b_{3g})$		1335	1301	1309	1379	1278	1290
$\omega_{14}(b_{3g})$		592	586	584	626	566	573
$\omega_{15}(b_{1u})$		3391	3238	3241	3332	3251	3215
$\omega_{16}(b_{1u})$	1403	1504	1482	1489	1571	1452	1468
$\omega_{17}(b_{1u})$	976	1164	1058	1067	1090	1063	1063
$\omega_{18}(b_{1u})$	918	621i	3739i	953	1042	932	869
$\omega_{19}(b_{2u})$		3415	3263	3257	3348	3266	3230
$\omega_{20}(b_{2u})$	1331	1546	1388	1391	1409	1457	1341
$\omega_{21}(b_{2u})$	1207	1144	1275	1256	1262	1232	1223
$\omega_{22}(b_{2u})$		306	1068	1079	1088	1056	1040
$\omega_{23}(b_{3u})$	721	812	750	766	782	776	758
$\omega_{24}(b_{3u})$	435	519	443	439	463	460	500

[a] References.[72,74]
[b] References.[75]
[c] DSRG-MRPT2 with QZ/JKFIT and CAS(8,8). The flow parameter s is 1.0 E_h^{-2}.
[d] References.[76]

predictions on $\omega_{24}(b_{3u})$ and $\omega_{18}(b_{1u})$ are comparatively less accurate. We further employed cu-DSRG-MRPT2 to investigate the vibrational frequencies and IR intensities of the singlet *p*-benzyne, the results are summarized in Table 4.32. Note that cu-DSRG-MRPT2 still yields good approximations to DSRG-MRPT2, given a different basis and active space. Surprisingly, for certain vibrational modes, such as $\omega_2(a_g)$, $\omega_8(b_{1g})$ and $\omega_{21}(b_{2u})$, the deviations compared to the DSRG-MRPT2 results are larger than 20 cm^{-1}. This implies that we may need other density cumulant approximation techniques for high-accuracy spectroscopic analysis, instead of brutally omitting all three-particle components.[37]

Table 4.32: Vibrational frequencies (in cm^{-1}) and IR intensities (in $km \cdot mol^{-1}$)of singlet *p*-benzyne obtained with DSRG-MRPT2. The flow parameter of DSRG is set to 1.0 E_h^{-2}.

	CAS(8e,8o)						CAS(2e,2o)	
	cc-pCVDZ		cc-pCVTZ		cc-pCVQZ		cc-pCVQZ	
Parameter	DSRG	cu-DSRG	DSRG	cu-DSRG	DSRG	cu-DSRG	DSRG	cu-DSRG
$\omega_1(a_g)$	3269	3266	3266	3262	3266	3263	3253	3251
$\omega_2(a_g)$	1363	1410	1333	1386	1329	1383	1353	1390
$\omega_3(a_g)$	1153	1152	1160	1159	1161	1160	1161	1160
$\omega_4(a_g)$	1023	1027	1023	1029	1025	1031	1035	1038
$\omega_5(a_g)$	624	615	629	619	632	621	628	623
$\omega_6(a_u)$	944	957	965	977	971	982	953	956
$\omega_7(a_u)$	422	417	433	426	435	428	436	430
$\omega_8(b_{1g})$	711	737	712	741	709	739	716	740
$\omega_9(b_{2g})$	900	914	922	936	927	939	913	918
$\omega_{10}(b_{2g})$	577	593	616	632	617	633	612	625
$\omega_{11}(b_{3g})$	3255	3252	3252	3249	3253	3249	3240	3237
$\omega_{12}(b_{3g})$	1664	1658	1664	1658	1670	1663	1687	1682
$\omega_{13}(b_{3g})$	1266	1271	1277	1283	1278	1284	1274	1278
$\omega_{14}(b_{3g})$	565	570	565	570	566	571	566	569
$\omega_{15}(b_{1u})$	3253	3250	3250	3247	3251	3248	3238	3235
$\omega_{16}(b_{1u})$	1448	1452	1451	1456	1452	1457	1452	1455
$\omega_{17}(b_{1u})$	1054	1049	1061	1056	1063	1058	1065	1062
$\omega_{18}(b_{1u})$	928	936	929	938	932	941	951	955
$\omega_{19}(b_{2u})$	3269	3266	3266	3262	3266	3262	3253	3250
$\omega_{20}(b_{2u})$	1450	1447	1453	1449	1457	1453	1456	1450
$\omega_{21}(b_{2u})$	1249	1279	1236	1271	1232	1268	1241	1263
$\omega_{22}(b_{2u})$	1060	1070	1056	1070	1056	1070	1059	1067
$\omega_{23}(b_{3u})$	760	769	776	784	776	784	764	766
$\omega_{24}(b_{3u})$	448	444	460	454	460	455	460	455
$I_{15}(b_{1u})$	0.08	0.22	0.01	0.01	0.07	0.00	0.10	0.03
$I_{16}(b_{1u})$	4.09	5.31	4.38	5.86	3.88	5.34	3.78	4.65
$I_{17}(b_{1u})$	19.74	16.47	19.73	16.01	20.42	16.58	18.61	16.52
$I_{18}(b_{1u})$	16.53	10.39	16.82	9.70	16.40	9.09	9.30	7.13
$I_{19}(b_{2u})$	0.13	0.73	0.01	0.22	0.19	0.03	0.13	0.00
$I_{20}(b_{2u})$	1.50	1.43	1.54	1.51	1.24	1.25	0.79	0.82
$I_{21}(b_{2u})$	5.80	4.98	6.17	5.15	6.11	4.98	5.19	4.64
$I_{22}(b_{2u})$	4.32	4.71	3.15	3.77	3.08	3.72	3.51	3.85
$I_{23}(b_{3u})$	77.86	75.79	83.14	80.94	82.89	80.60	78.68	76.58
$I_{24}(b_{3u})$	14.71	14.57	17.81	17.46	18.06	17.67	19.21	18.93

4.4 CONCLUSION

We have presented the analytic energy gradient theory for unrelaxed DSRG-MRPT2 with density fitting. The DF technique avoids the explicit storage and the time-costly backtransformation of four-index integrals. Instead, the associated large

integrals are factorized using small two- and three-center integrals, and the integral transformation in the Z-vector equation is also correspondingly modified. In principle, these treatments are similar to those of DF-MP2 and DF-CASPT2. While providing significant computational acceleration and alleviating the memory bottlenecks, DF only introduces negligible effects on the accuracy. It is further worth mentioning that the density-fitted DSRG parameters are still numerically robust, making DSRG immune to intruder state problems.

To demonstrate the viability of the DF analytic gradients, we have used DSRG-MRPT2 to optimize the equilibrium geometries of *p*-benzyne for both singlet and triplet states, using large basis and large active space. The DSRG-MRPT2 results are in good agreement with those of CASPT2, though DSRG-MRPT2 slightly underestimates the bond lengths. Comparing geometries between two states, those for the triplet state *p*-benzyne are more sensitive to basis and active space. The structures optimized with a pruned DSRG-MRPT2 formalism, cu-DSRG-MRPT2, perfectly match those computed using DSRG-MRPT2. We also reported the spectroscopic parameters, including vibrational frequencies and IR intensities consistent with those computed with 4R RMR CCSD. The spectroscopic data implies that other density cumulant approximation techniques should be used for DSRG-MRPT2 if high accuracy is required, instead of the simplest cu-DSRG-MRPT2 method.

The current work points out several future directions. Since the DF gradient implementation is generalized, we would expect the development of analytic gradients for DF-DSRG with a GASSCF reference wave function, or with the conventional CASCI substituted with selected CI methods, such as ACI. A challenging extension would be the gradient theory for higher-order MR-DSRG theories, such as the MR-DSRG third-order perturbation theory (DSRG-MRPT3). We would also like to develop a new approximate DSRG-MRPT2 formalism instead of cu-DSRG-MRPT2 for accurate spectroscopic analysis.

4.5 ACKNOWLEDGMENTS

This work was supported by the U.S. Department of Energy under Award No. DE-SC0016004. F.A.E acknowledges support from a Camille Dreyfus Teacher-Scholar Award.

Bibliography

[1] Mok, D. K.; Neumann, R.; Handy, N. C. Dynamical and nondynamical correlation. *J. Phys. Chem.* **1996**, *100*, 6225–6230.

[2] Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. Multireference nature of chemistry: the coupled-cluster view. *Chem. Rev.* **2012**, *112*, 182–243.

[3] Evangelista, F. A. Perspective: Multireference coupled cluster theories of dynamical electron correlation. *J. Chem. Phys.* **2018**, *149*, 030901.

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

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

[6] Andersson, K.; Malmqvist, P.-Å.; Roos, B. O. Second-order perturbation theory with a complete active space self-consistent field reference function. *J. Chem. Phys.* **1992**, *96*, 1218–1226.

[7] Kozlowski, P. M.; Davidson, E. R. Considerations in constructing a multireference second-order perturbation theory. *J. Chem. Phys.* **1994**, *100*, 3672–3682.

[8] Werner, H.-J. Third-order multireference perturbation theory The CASPT3 method. *Mol. Phys.* **1996**, *89*, 645–661.

[9] Mahapatra, U. S.; Datta, B.; Mukherjee, D. Development of a size-consistent state-specific multireference perturbation theory with relaxed model-space coefficients. *Chem. Phys. Lett.* **1999**, *299*, 42–50.

[10] Angeli, C.; Cimiraglia, R.; Evangelisti, S.; Leininger, T.; Malrieu, J.-P. Introduction of n-electron valence states for multireference perturbation theory. *J. Chem. Phys.* **2001**, *114*, 10252–10264.

[11] Khait, Y. G.; Song, J.; Hoffmann, M. R. Explication and revision of generalized Van Vleck perturbation theory for molecular electronic structure. *J. Chem. Phys.* **2002**, *117*, 4133–4145.

[12] Szabados, Á.; Rolik, Z.; Tóth, G.; Surján, P. R. Multiconfiguration perturbation theory: Size consistency at second order. *J. Chem. Phys.* **2005**, *122*, 114104.

[13] Hoffmann, M. R.; Datta, D.; Das, S.; Mukherjee, D.; Szabados, A.; Rolik, Z.; Surján, P. R. Comparative study of multireference perturbative theories for ground and excited states. *J. Chem. Phys.* **2009**, *131*, 204104.

[14] Evangelista, F. A.; Simmonett, A. C.; Schaefer, H. F.; Mukherjee, D.; Allen, W. D. A companion perturbation theory for state-specific multireference coupled cluster methods. *Phys. Chem. Chem. Phys.* **2009**, *11*, 4728–4741.

[15] Sokolov, A. Y.; Guo, S.; Ronca, E.; Chan, G. K.-L. Time-dependent N-electron valence perturbation theory with matrix product state reference wavefunctions for large active spaces and basis sets: Applications to the chromium dimer and all-trans polyenes. *J. Chem. Phys.* **2017**, *146*, 244102.

[16] Giner, E.; Angeli, C.; Garniron, Y.; Scemama, A.; Malrieu, J.-P. A Jeziorski-Monkhorst fully uncontracted multi-reference perturbative treatment. I. Principles, second-order versions, and tests on ground state potential energy curves. *J. Chem. Phys.* **2017**, *146*, 224108.

[17] Celani, P.; Werner, H.-J. Analytical energy gradients for internally contracted second-order multireference perturbation theory. *J. Chem. Phys.* **2003**, *119*, 5044–5057.

[18] MacLeod, M. K.; Shiozaki, T. Communication: Automatic code generation enables nuclear gradient computations for fully internally contracted multireference theory. *J. Chem. Phys.* **2015**, *142*, 051103.

[19] Park, J. W.; Al-Saadon, R.; Strand, N. E.; Shiozaki, T. Imaginary shift in CASPT2 nuclear gradient and derivative coupling theory. *J. Chem. Theory Comput.* **2019**, *15*, 4088–4098.

[20] Győrffy, W.; Shiozaki, T.; Knizia, G.; Werner, H.-J. Analytical energy gradients for second-order multireference perturbation theory using density fitting. *J. Chem. Phys.* **2013**, *138*, 104104.

[21] Vlaisavljevich, B.; Shiozaki, T. Nuclear energy gradients for internally contracted complete active space second-order perturbation theory: Multistate extensions. *J. Chem. Theory Comput.* **2016**, *12*, 3781–3787.

[22] Shiozaki, T.; Győrffy, W.; Celani, P.; Werner, H.-J. Communication: Extended multi-state complete active space second-order perturbation theory: Energy and nuclear gradients. *J. Chem. Phys.* **2011**, *135*, 081106.

[23] Shiozaki, T.; Werner, H.-J. Communication: Second-order multireference perturbation theory with explicit correlation: CASPT2-F12. *J. Chem. Phys.* **2010**, *133*, 141103.

[24] Shiozaki, T. BAGEL: brilliantly advanced general electronic-structure library. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1331.

[25] Park, J. W. Analytical Gradient Theory for Strongly Contracted (SC) and Partially Contracted (PC) N-Electron Valence State Perturbation Theory (NEVPT2). *J. Chem. Theory Comput.* **2019**, *15*, 5417–5425.

[26] Park, J. W. Analytical Gradient Theory for Quasidegenerate N-Electron Valence State Perturbation Theory (QD-NEVPT2). *J. Chem. Theory Comput.* **2020**, *16*, 326–339.

[27] Nishimoto, Y. Analytic first-order derivatives of partially contracted *n*-electron valence state second-order perturbation theory (PC-NEVPT2). *J. Chem. Phys.* **2019**, *151*, 114103.

[28] Nishimoto, Y. Locating conical intersections using the quasidegenerate partially and strongly contracted NEVPT2 methods. *Chem. Phys. Lett.* **2020**, *744*, 137219.

[29] Evangelisti, S.; Daudey, J.; Malrieu, J. Qualitative intruder-state problems in effective Hamiltonian theory and their solution through intermediate Hamiltonians. *Phys. Rev. A* **1987**, *35*, 4930.

[30] Chattopadhyay, S.; Chaudhuri, R. K.; Mahapatra, U. S.; Ghosh, A.; Ray, S. S. State-specific multireference perturbation theory: development and present status. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2016**, *6*, 266–291.

[31] Witek, H. A.; Choe, Y.-K.; Finley, J. P.; Hirao, K. Intruder state avoidance multireference Møller–Plesset perturbation theory. *J. Comput. Chem.* **2002**, *23*, 957–965.

[32] Choe, Y.-K.; Witek, H. A.; Finley, J. P.; Hirao, K. Identifying and removing intruder states in multireference Mo/ller–Plesset perturbation theory. *J. Chem. Phys.* **2001**, *114*, 3913–3918.

[33] Camacho, C.; Witek, H. A.; Yamamoto, S. Intruder states in multireference perturbation theory: The ground state of manganese dimer. *J. Comput. Chem.* **2009**, *30*, 468–478.

[34] Camacho, C.; Cimiraglia, R.; Witek, H. A. Multireference perturbation theory can predict a false ground state. *Phys. Chem. Chem. Phys.* **2010**, *12*, 5058–5060.

[35] Dyall, K. G. The choice of a zeroth-order Hamiltonian for second-order perturbation theory with a complete active space self-consistent-field reference function. *J. Chem. Phys.* **1995**, *102*, 4909–4918.

[36] Angeli, C.; Pastore, M.; Cimiraglia, R. New perspectives in multireference perturbation theory: the n-electron valence state approach. *Theor. Chem. Acc.* **2007**, *117*, 743–754.

[37] Zgid, D.; Ghosh, D.; Neuscamman, E.; Chan, G. K.-L. A study of cumulant approximations to n-electron valence multireference perturbation theory. *J. Chem. Phys.* **2009**, *130*, 194107.

[38] Evangelista, F. A. A driven similarity renormalization group approach to quantum many-body problems. *J. Chem. Phys.* **2014**, *141*, 054109.

[39] Li, C.; Evangelista, F. A. Multireference driven similarity renormalization group: A second-order perturbative analysis. *J. Chem. Theory Comput.* **2015**, *11*, 2097–2108.

[40] Li, C.; Evangelista, F. A. Multireference theories of electron correlation based on the driven similarity renormalization group. *Annu. Rev. Phys. Chem.* **2019**, *70*, 245–273.

[41] Li, C.; Evangelista, F. A. Driven similarity renormalization group: Third-order multireference perturbation theory. *J. Chem. Phys.* **2017**, *146*, 124132.

[42] Schriber, J. B.; Hannon, K. P.; Li, C.; Evangelista, F. A. A combined selected configuration interaction and many-body treatment of static and dynamical correlation in oligoacenes. *J. Chem. Theory Comput.* **2018**, *14*, 6295–6305.

[43] Hannon, K. P.; Li, C.; Evangelista, F. A. An integral-factorized implementation of the driven similarity renormalization group second-order multireference perturbation theory. *J. Chem. Phys.* **2016**, *144*, 204111.

[44] Wang, S.; Li, C.; Evangelista, F. A. Analytic energy gradients for the driven similarity renormalization group multireference second-order perturbation theory. *J. Chem. Theory Comput.* **2021**, *17*, 7666–7681.

[45] Wang, S.; Li, C.; Evangelista, F. A. Analytic gradients for the single-reference driven similarity renormalization group second-order perturbation theory. *J. Chem. Phys.* **2019**, *151*, 044118.

[46] Forte, a suite of quantum chemistry methods for strongly correlated electrons. For current version see https://github.com/evangelistalab/forte, 2021.

[47] Park, J. W. Analytical Gradient Theory for Spin-Free State-Averaged Second-Order Driven Similarity Renormalization Group Perturbation Theory (SA-DSRG-MRPT2) and Its Applications for Conical Intersection Optimizations. *J. Chem. Theory Comput.* **2022**, *18*, 2233–2245.

[48] Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the coupled-cluster singles and doubles method with the density-fitting approximation. *J. Chem. Phys.* **2016**, *144*, 174103.

[49] Bozkaya, U.; Sherrill, C. D. Analytic energy gradients for the coupled-cluster singles and doubles with perturbative triples method with the density-fitting approximation. *J. Chem. Phys.* **2017**, *147*, 044104.

[50] Schütz, M.; Werner, H.-J.; Lindh, R.; Manby, F. R. Analytical energy gradients for local second-order Møller–Plesset perturbation theory using density fitting approximations. *J. Chem. Phys.* **2004**, *121*, 737–750.

[51] Delcey, M. G.; Pedersen, T. B.; Aquilante, F.; Lindh, R. Analytical gradients of the state-average complete active space self-consistent field method with density fitting. *J. Chem. Phys.* **2015**, *143*, 044110.

[52] Bozkaya, U. Analytic energy gradients and spin multiplicities for orbital-optimized second-order perturbation theory with density-fitting approximation: an efficient implementation. *J. Chem. Theory Comput.* **2014**, *10*, 4389–4399.

[53] Bozkaya, U. Analytic energy gradients for orbital-optimized MP3 and MP2. 5 with the density-fitting approximation: An efficient implementation. *J. Comput. Chem.* **2018**, *39*, 351–360.

[54] Distasio Jr, R. A.; Steele, R. P.; Rhee, Y. M.; Shao, Y.; Head-Gordon, M. An improved algorithm for analytical gradient evaluation in resolution-of-the-identity second-order Møller-Plesset perturbation theory: Application to alanine tetrapeptide conformational analysis. *J. Comput. Chem.* **2007**, *28*, 839–856.

[55] Ma, D.; Li Manni, G.; Olsen, J.; Gagliardi, L. Second-order perturbation theory for generalized active space self-consistent-field wave functions. *J. Chem. Theory Comput.* **2016**, *12*, 3208–3213.

[56] Huang, M.; Li, C.; Evangelista, F. A. Theoretical Calculation of Core-Excited States along Dissociative Pathways beyond Second-Order Perturbation Theory. *J. Chem. Theory Comput.* **2021**,

[57] Schriber, J. B.; Evangelista, F. A. Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy. *J. Chem. Phys.* **2016**, *144*, 161106.

[58] Aquilante, F.; Boman, L.; Boström, J.; Koch, H.; Lindh, R.; Merás, A. S. d.; Pedersen, T. B. *Linear-Scaling Techniques in Computational Chemistry and Physics*; Springer, 2011; pp 301–343.

[59] Kutzelnigg, W.; Mukherjee, D. Normal order and extended Wick theorem for a multiconfiguration reference wave function. *J. Chem. Phys.* **1997**, *107*, 432–449.

[60] Kutzelnigg, W.; Mukherjee, D. Cumulant expansion of the reduced density matrices. *J. Chem. Phys.* **1999**, *110*, 2800–2809.

[61] Koch, W. A New Dimension to Quantum Chemistry. Analytic Derivative Methods in Ab Initio Molecular Electronic Structure Theory. Von Y. Yamaguchi, Y. Osamura, JD Goddard und HF Schaefer. Oxford University Press, Oxford, 1994. 471 S., geb. 60.00£.–ISBN 0-19-507028-3. 1995.

[62] Osamura, Y.; Yamaguchi, Y.; Schaefer III, H. F. Generalization of analytic configuration interaction (CI) gradient techniques for potential energy hypersurfaces, including a solution to the coupled perturbed Hartree–Fock equations for multiconfiguration SCF molecular wave functions. *J. Chem. Phys.* **1982**, *77*, 383–390.

[63] Handy, N. C.; Schaefer III, H. F. On the evaluation of analytic energy derivatives for correlated wave functions. *J. Chem. Phys.* **1984**, *81*, 5031–5033.

[64] Amos, R. D.; Rice, J. E. Implementation of analytic derivative methods in quantum chemistry. *Comput. Phys. Rep.* **1989**, *10*, 147–187.

[65] Helgaker, T.; Jørgensen, P. Analytical calculation of geometrical derivatives in molecular electronic structure theory. *Adv. Quantum Chem.* **1988**, *19*, 183–245.

[66] Politzer, P.; Murray, J. S. The Hellmann-Feynman theorem: a perspective. *J. Mol. Model.* **2018**, *24*, 1–7.

[67] Dunlap, B.; Connolly, J.; Sabin, J. On first-row diatomic molecules and local density models. *J. Chem. Phys.* **1979**, *71*, 4993–4999.

[68] Dunlap, B. I.; Connolly, J.; Sabin, J. On some approximations in applications of X α theory. *J. Chem. Phys.* **1979**, *71*, 3396–3402.

[69] Weigend, F. A fully direct RI-HF algorithm: Implementation, optimised auxiliary basis sets, demonstration of accuracy and efficiency. *Phys. Chem. Chem. Phys.* **2002**, *4*, 4285–4291.

[70] others,, et al. PSI4 1.4: Open-source software for high-throughput quantum chemistry. *J. Chem. Phys.* **2020**, *152*, 184108.

[71] Kraka, E.; Cremer, D. Computer design of anticancer drugs. A new enediyne warhead. *J. Am. Chem. Soc.* **2000**, *122*, 8245–8264.

[72] Marquardt, R.; Balster, A.; Sander, W.; Kraka, E.; Cremer, D.; Radziszewski, J. G. p-Benzyne. *Angew. Chem. Int. Ed.* **1998**, *37*, 955–958.

[73] others,, et al. MOLPRO, version 2015.1, a package of ab initio programs. 2015; see http://www. molpro. net.

[74] Wenthold, P. G.; Squires, R. R.; Lineberger, W. Ultraviolet Photoelectron Spectroscopy of the o-, m-, and p-Benzyne Negative Ions. Electron Affinities and Singlet- Triplet Splittings for o-, m-, and p-Benzyne. *J. Am. Chem. Soc.* **1998**, *120*, 5279–5290.

[75] Crawford, T. D.; Kraka, E.; Stanton, J. F.; Cremer, D. Problematic p-benzyne: Orbital instabilities, biradical character, and broken symmetry. *J. Chem. Phys.* **2001**, *114*, 10638–10650.

[76] Li, X.; Paldus, J. Force field of para-and metabenzyne diradicals: A multireference coupled-cluster study. *J. Chem. Phys.* **2010**, *132*, 114103.

Chapter 5 Conclusion and Future outlook

In this dissertation, we have outlined the development of the analytic gradient theory for DSRG-PT2 and its applications to geometry optimization and vibrational spectroscopy. Our efficient implementation of gradients substantially extends the viability of DSRG for investigating larger strongly correlated systems, which may be impractical to compute with conventional finite-difference approaches.

Our initial work was devoted to the analytic energy gradients of DSRG-SRPT2. The major point of this study was to evaluate how the regularized DSRG equations affect the analytic derivation. Such derivation was achieved by solving a set of orbital response equations formulated with the method of Lagrange multipliers, enforcing canonical orbital constraints and the definition of DSRG parameters. The exactness of the analytic derivatives is confirmed by comparing to those computed with the finite-difference method. Importantly, we determined that the gradients are numerically robust to intruder states. We then employed the DSRG-SRPT2 analytic gradients to optimize the geometries of 15 small molecules, and the resulting structures are found to be in decent agreement with those of other single-reference methods. For example, a mean absolute error of 0.0033 Å and a standard deviation of 0.0045 Å is observed for bond lengths, compared to CCSD(T) results. However, such single-reference approach is not adequate for exploring strongly correlated systems.

Therefore, we further investigated the analytic energy derivatives within the unrelaxed DSRG-MRPT2 formalism. Switching to the multireference regime, the Lagrangian was reformulated with relatively comprehensive constraints, imposing definitions of DSRG modified integrals, one- and two-body cluster amplitudes, semi-canonical molecular orbitals, and CI coefficients. Furthermore, an iterative routine, the general minimal residual method (GMRES), is utilized to circumvent the explicit

storage of the large linear system of CI response equations. Two similar properties are retained in our resulting derivative equations: (1) numerical stability is inherited by principle, making the solutions to the response equations immune to intruder states; (2) the computational scaling is the same as a single-point DSRG-MRPT2 energy computation. The analytic gradients are used to optimize the singlet and triplet states of *p*-benzyne with a minimal CAS(2,2) active space. The equilibrium geometries show good agreement with the ones obtained with CASPT2, pc-NEVPT2, and Mk-MRCCSD(T). From the *s*-dependency study, we illustrate that the optimized geometry of the singlet state is more sensitive to s, while for both singlet and triplet states the value 1.0 E_{h}^{-2} represents an optimal choice to capture correlation energies. We also note that the DSRG-MRPT2 method underestimates the adiabatic singlet-triplet gap of *p*-benzyne compared to other MR approaches. It is worth mentioning that two significant issues exist within this implementation: (1) a large basis or a large active space must be avoided, due to the storage and the integral transformation of four-index integrals; (2) explicit storage of six-index intermediate tensors drastically increase the memory cost of the CI response equations. We tentatively employed a pruned DSRG-MRPT2 approach, in which all expensive three-body density cumulant terms are excluded. The negligible difference in both optimized geometry and spin gap suggests that this approximate variant of DSRG-MRPT2 may be useful in applications involving large systems, without substantial degradation of accuracy.

To address the memory bottlenecks discussed in the previous study, we applied the density fitting treatment to the DSRG-MRPT2 analytic gradients. With density fitting, direct storage of large four-index integrals, such as ERIs and dependent two-body Lagrange multipliers, is circumvented. Moreover, the expensive integral back-transformation of four-index ERI-associated derivative densities is converted to multiple affordable back-transformations of three- and two-index density integrals. Meanwhile, the immunity to intruder states is still guaranteed. The DF-DSRG-

MRPT2 gradients allow us to study the troublesome *p*-benzyne, with a much larger basis and active space such as cc-pCVQZ/CAS(8,8). We applied DF-DSRG-MRPT2 gradients to optimize the geometries of singlet and triplet *p*-benzyne, and further performed a vibrational analysis on singlet *p*-benzyne using semi-numerical Hessian and the dipole moment derivatives. These results are compared to experimental data, and found in good agreement with those of other MR approaches. We also employed the pruned DSRG-MRPT2 formalism, named cu-DSRG-MRPT2, which neglects all three-particle density cumulants. While the equilibrium structures show only negligible deviations compared to the original DSRG-MRPT2 method, there exists a noticeable difference in vibrational frequencies for certain modes, as large as 20 cm^{-1}.

In conclusion, the developed analytic gradient theory significantly expands the applicability of DSRG to large molecular systems. Herein, we propose several potential extensions: (1) a straightforward modification would be applying the frozen-orbital approximation, which further alleviates the computational bottleneck for molecules with large core spaces; (2) the reference CASSCF wave function may be substituted with those of other MR approaches, such as restricted (RASSCF) or generalized active space self-consistent-field (GASSCF); (3) the conventional CI routine could be substituted with selected CI approaches that effectively shrink the determinant space by neglecting unimportant configurations; (4) an alternative approximate regime for DSRG-MRPT2 other than cu-DSRG-MRPT2 should be proposed and benchmarked; (5) the analytic gradients considering higher-order perturbative analysis, such as DSRG-MRPT3, and orbital-relaxed DSRG formalism may be challenging projects; (6) motivated by Toru Shiozaki, we envision an automatic gradient generator developed in Forte to mitigate laborious derivations. Overall, we anticipate that the further development of the analytic derivative theory could levitate DSRG to a much higher level.