

DEVELOPMENT OF DENSITY CUMULANT FUNCTIONAL THEORY

by

ALEXANDER YURYEVICH SOKOLOV

(Under the direction of Henry F. Schaefer III)

ABSTRACT

Ab initio quantum chemistry has been dominated by the wavefunction-based approaches to electron correlation. Alternatively, methods that seek approximate solutions in terms of the reduced density matrices (RDMs) have been developed. RDM methods circumvent the computation of the many-electron wavefunction, but have been shown to incorrectly describe the separability into non-interacting molecular fragments and scaling of the energy with the number of electrons. Recently, a new approach to electron correlation called *density cumulant functional theory* (DCFT) was proposed that eliminates these shortcomings. Herein, we present new developments of DCFT extending the applicability of the theory. We begin by deriving and implementing the analytic energy gradients for the original DCFT method. We then formulate a new DCFT method with the improved description of the one-particle density matrix. Finally, a new orbital-optimized DCFT formulation is proposed and implemented. We offer a perturbative analysis of the new methods and benchmark their performance for a variety of chemical systems.

INDEX WORDS: computational quantum chemistry, electronic structure theory, electron correlation, reduced density matrix methods, density cumulants, N-representability, extensivity, size-consistency, density cumulant functional theory, analytic gradients, orbital optimization

DEVELOPMENT OF DENSITY CUMULANT FUNCTIONAL THEORY

by

ALEXANDER YURYEVICH SOKOLOV

Specialist (M.S.), Saint-Petersburg State University, Russia, 2009

A Dissertation Submitted to the Graduate Faculty
of The University of Georgia in Partial Fulfillment
of the
Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2014

©2014 All Rights Reserved
Alexander Yuryevich Sokolov

DEVELOPMENT OF DENSITY CUMULANT FUNCTIONAL THEORY

by

ALEXANDER YURYEVICH SOKOLOV

Approved:

Major Professors: Henry F. Schaefer III

Committee: Paul v. R. Schleyer
Gary E. Douberly

Electronic Version Approved:

Dr. Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2014

ACKNOWLEDGMENTS

When I look back at the five years I have spent in the Graduate School I realize how fortunate I am in my life. Throughout these years I have been receiving help and support from many wonderful people and I am thankful to all of them for that.

Nothing I accomplished would have been ever possible without the help of my dear parents and my family. Thank you for all of your support, encouragement and never-ending belief in me. Thanks to my Mom for raising me the way I am. Thanks to my Dad for instilling in me a love of science and nature. To my Sister, thanks for encouragement and helping Mom and Dad while I was away.

To my wife, Victoria, thank you for everything that you gave me. Thank you for going with me to US and leaving your life in St. Petersburg behind. You are one of the most wonderful, kind and honest people I've ever met, and I can't imagine how I would live without you. Thank you for raising our kids with love and care, and many late nights we went through together. Thank you for your ideas, encouragement and help. To my children, Kira and Leo. I know you are too small to read this, but I will always be thankful to you for showing me that life is a miracle. Thanks to you, I learned a lot about myself. Thank you for all of your love you give me every day.

To my advisor, Professor Schaefer, thank you for all of the freedom you gave me and for teaching me how to be a scientist. For all of your faith in me, even when the things went

wrong, it was very important for me to feel it. For creating a place to work, where everyone is part of a family.

I would not end up where I am without my dear friend and undergraduate research advisor Professor Olga Sizova who showed me the beauty of theoretical chemistry. It is with her help that I made my first “baby steps” in science and I will never forget that. With a heavy heart, I realize that Prof. Sizova is no longer with us.

Many thanks to Andy Simmonett and Jeremy Wilke for helping me understand theory. To Andy, special thanks for all of the help with DCFT and three fun papers we wrote together. Thanks to Dr. Yukio Yamaguchi for guidance with programming projects and allowing me to TA his course. To Justin Turney, for all of the help with programming and (importantly) debugging. To Francesco Evangelista, for giving important advices at the very early stage and interest in my work.

Thanks to all of the CCQC members I have met, past and present. Special thanks to Stefan Vogt, Brandon Magers and Judy Wu for many discussions and fun collaborations we had together. To David Hollman, for talks about theory and methods. Thanks to Jay Agarwal for bringing a lot of excellent chemistry to CCQC. To Prof. Allen and Prof. Schleyer, for everything I learned in their courses and engagement in research together. To Prof. Schleyer and Prof. Douberly, for being members of my Advisory Committee.

Finally, I’d like to thank many people I met outside CCQC for very insightful discussions that inevitably influenced this work. My trip to Germany in Summer 2013 would have not been possible without the help and hospitality of Prof. Kutzelnigg. Thank you for the wonderful five weeks I spent in Bochum and sharing your deep understanding of quantum chemistry with me. Special thanks to Prof. Marcel Nooijen, Dr. Christian Kollmar, Prof. Frank Neese, Prof. Christian Ochsenfeld, Prof. Takeshi Yanai and Prof. Uğur Bozkaya for a lot of interest in my work.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
1 INTRODUCTION AND LITERATURE REVIEW	1
1.1 ELECTRON CORRELATION	1
1.2 WAVEFUNCTION-BASED APPROACHES TO MANY-ELECTRON PROBLEM . .	3
1.3 EXTENSIVITY AND SIZE-CONSISTENCY	6
1.4 REDUCED DENSITY MATRIX METHODS	8
1.5 DENSITY CUMULANTS	11
1.6 PROSPECTUS	13
2 ANALYTIC GRADIENTS FOR DENSITY CUMULANT FUNCTIONAL THEORY:	
THE DCFT-06 MODEL	15
2.1 ABSTRACT	16
2.2 INTRODUCTION	16
2.3 GENERAL THEORY	18
2.4 COMPUTATIONAL METHODS	26
2.5 RESULTS	26
2.6 CONCLUSIONS	31
2.7 ACKNOWLEDGMENTS	32
2.8 SUPPLEMENTARY MATERIAL	32

3	DENSITY CUMULANT FUNCTIONAL THEORY: THE DC-12 METHOD, AN IMPROVED DESCRIPTION OF THE ONE-PARTICLE DENSITY MATRIX	33
3.1	ABSTRACT	34
3.2	INTRODUCTION	34
3.3	THEORY	36
3.4	COMPUTATIONAL METHODS	45
3.5	RESULTS	46
3.6	CONCLUSIONS	54
3.7	ACKNOWLEDGMENTS	56
3.8	APPENDIX: SIMPLIFICATION OF THE EXPRESSIONS IN THE NATURAL SPIN- ORBITAL (NSO) BASIS	56
4	ORBITAL-OPTIMIZED DENSITY CUMULANT FUNCTIONAL THEORY	58
4.1	ABSTRACT	59
4.2	INTRODUCTION	59
4.3	THEORY	62
4.4	COMPUTATIONAL DETAILS	71
4.5	RESULTS AND DISCUSSION	72
4.6	CONCLUSIONS	80
4.7	ACKNOWLEDGMENTS	81
4.8	SUPPLEMENTARY MATERIAL	81
5	CONCLUSIONS	82
	BIBLIOGRAPHY	85

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 ELECTRON CORRELATION

Reliable predictions of the molecular properties require accurate description of the instantaneous electron-electron interactions. It is now widely agreed that accounting for these interactions is very important if chemical accuracy (1 kcal mol⁻¹) is desired and often is crucial to obtain even just qualitatively correct results.^{1,2} In 1934 Wigner and Seitz used the term “electron correlation” to describe the effect of these instantaneous interactions on the probability distributions of the individual electrons in the metallic sodium.³ Ever since, electron correlation has been often referred to as the central problem of quantum chemistry and still remains its focus after more than five decades of active development.^{1,2}

Despite its significance, the term “electron correlation” does not have a unique definition in quantum chemistry.⁴ It is common to refer to the electron correlation^{1,5} as the difference between the Hartree-Fock description of the electronic wavefunction (Φ_{HF}) and the exact solution of the nonrelativistic time-independent Schrödinger equation⁶

$$E = \langle \Psi | \hat{H} | \Psi \rangle . \quad (1.1)$$

The Hamiltonian $\hat{H}$ contains the electronic kinetic energy operator, electron-nuclear and electron-electron potential.⁶ The Born-Oppenheimer approximation is typically assumed.⁷ Taking the difference between the exact energy and the energy obtained from the Hartree-Fock approximation (E_{HF}), one can obtain a measure of the electron correlation, called the correlation energy:^{1,5}

$$\Delta E_{corr} = E - E_{HF} . \quad (1.2)$$

Alternative but less frequently used definition of the electron correlation^{4,8,9} has its origin in probability theory. Let us define the one- and two-particle probability densities. These can be obtained by integrating the N -electron wavefunction over the coordinates of $(N - 1)$ and $(N - 2)$ electrons, respectively:

$$\begin{aligned} \gamma(\mathbf{x}_1; \mathbf{x}'_1) &= \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi(\mathbf{x}'_1, \mathbf{x}_2, \dots, \mathbf{x}_N) d\mathbf{x}_2 \dots d\mathbf{x}_N , \\ \Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) &= \int \Psi^*(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) \Psi(\mathbf{x}'_1, \mathbf{x}'_2, \dots, \mathbf{x}_N) d\mathbf{x}_3 \dots d\mathbf{x}_N . \end{aligned} \quad (1.3)$$

Two electrons are called correlated if their pair probability density cannot be written as the product of two one-particle probabilities:

$$\Gamma(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) \neq \gamma(\mathbf{x}_1; \mathbf{x}'_1) \gamma(\mathbf{x}_2; \mathbf{x}'_2) . \quad (1.4)$$

In the case of the two independent electrons, the inequality (1.4) becomes an equality, the corresponding approximation is called a mean-field (or an independent-particle) model.⁹ Using the definition (1.4), two types of the electron correlation can be distinguished: the correlation between the electrons with the same and the opposite spins, referred to as Fermi and Coulomb correlation, respectively.^{5,9} Importantly, the pair probability density in the Hartree-Fock approximation can be written as the antisymmetrized product of the one-

particle probabilities^{1,9}

$$\Gamma_{\text{HF}}(\mathbf{x}_1, \mathbf{x}_2; \mathbf{x}'_1, \mathbf{x}'_2) = \gamma(\mathbf{x}_1; \mathbf{x}'_1)\gamma(\mathbf{x}_2; \mathbf{x}'_2) - \gamma(\mathbf{x}_1; \mathbf{x}'_2)\gamma(\mathbf{x}_2; \mathbf{x}'_1) , \quad (1.5)$$

which ensures the validity of the Pauli principle and entirely accounts for the Fermi correlation. Thus, the correlation energy (1.2) in the first definition of the electron correlation provides solely an energetic measure of the Coulomb correlation. Deviation of the pair probability density (1.3) from that of the Hartree-Fock approximation (1.5) provides an alternative measure of the Coulomb correlation based on the second definition of the electron correlation. Usually, the correlation is classified as *dynamic* or *static (non-dynamic)*, which corresponds to the weak or strong electron correlation limits, respectively.¹ While the difference between these two limiting cases is not strictly defined, the correlation is considered to be dynamic if the Hartree-Fock model (1.5) provides a good zeroth-order approximation to the pair probability density (and the corresponding wavefunction) and is referred to static otherwise. The latter case is common for the electronic states with unpaired electrons, bond breaking or degenerate electronic configurations. In the following sections of this Chapter two types of approaches to the electron correlation problem will be briefly outlined. We will begin by describing the electron correlation methods based on the expansion of the electronic wavefunction. We will then proceed by introducing important concepts of extensivity and size-consistency. Finally, the reduced density matrix approach to the electron correlation problem will be presented and the concept of the density cumulant will be introduced.

1.2 WAVEFUNCTION-BASED APPROACHES TO MANY-ELECTRON PROBLEM

Let us consider an N -electron molecular wavefunction expanded in the finite one-particle basis $\{\psi_p\}_{p=1}^M$. In the simplest case, the wavefunction can be written as the antisymmetrized

product of N spin-orbitals ψ_i , known as Slater determinant:^{10,11}

$$\Phi_{\text{HF}} = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_2(1) & \cdots & \psi_N(1) \\ \psi_1(2) & \psi_2(2) & \cdots & \psi_N(2) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_1(N) & \psi_2(N) & \cdots & \psi_N(N) \end{vmatrix} = \hat{\mathcal{A}}\psi_1\psi_2\cdots\psi_N. \quad (1.6)$$

In the Hartree-Fock approximation^{1,12} a single Slater determinant wavefunction Φ_{HF} is used and the molecular orbitals ψ_p are variationally optimized to yield the lowest energy possible. The N spin-orbitals that appear in the Slater determinant are usually referred to as occupied, while the extra $(M - N)$ spin-orbitals are called unoccupied or virtual.

The Hartree-Fock model has been remarkably successful in quantum chemistry. It recovers up to 99% of the total electronic energy and is capable of providing a reasonable (sometimes even accurate) description of the molecular species with closed electronic shells near equilibrium geometries.^{1,2,12} While the errors in the electronic energy become even smaller if only the relative energies are of interest, the missing correlation energy contributions are still chemically important.¹ To go beyond the Hartree-Fock approximation, in the wavefunction-based electron correlation methods the molecular wavefunction is expanded in the basis of more than one Slater determinant.^{1,12,13} In order to do this, the Hartree-Fock wavefunction is usually obtained first, known as the reference determinant. Then, a set of the “excited” Slater determinants is generated by replacing the occupied spin-orbitals in the reference determinant by the virtual spin-orbitals.

In a simple expansion, known as *configuration interaction* (CI) approach,^{14–16} the N -electron wavefunction Ψ is written as the linear combination of the reference and excited determinants:

$$\Psi = \Phi_{\text{HF}} + \sum_{ia} C_i^a \Phi_i^a + \sum_{i<j, a<b} C_{ij}^{ab} \Phi_{ij}^{ab} + \dots \quad (1.7)$$

In Eq. (1.7) indices $i, j, \dots$ and $a, b, \dots$ denote occupied and virtual spin-orbitals, respectively. Determinants Φ_i^a are called *singly excited* and are formed by replacing one occupied spin-orbital ψ_i by the virtual orbital ψ_a . Similarly, Φ_{ij}^{ab} are *doubly excited* determinants. In the case when expansion (1.7) is complete, i.e. includes up to all possible N -fold excitations, the obtained energy is the exact solution of the Schrödinger equation in a given finite basis, known as full CI. Full CI solutions are very expensive and in practice the linear expansion (1.7) is truncated at a low excitation rank (two- or three-fold excitations).

While the truncated CI methods provide a rather compact description of the electron correlation problem, they have been rarely used in the past decade due to the incorrect scaling of the correlation energy with the number of electrons (see Sec. 1.3 for details).^{1,13} This problem can be avoided if one abandons the linear expansion (1.7) and, instead, constructs the correlated wavefunction using an exponential excitation operator:

$$\Psi = e^{\hat{T}} \Phi_{\text{HF}} . \quad (1.8)$$

The exponential ansatz (1.8) is known as *coupled-cluster* (CC) theory.^{13,17-23} CC theory has been one of the most successful approaches to the electron correlation problem due to its high accuracy and extensivity (see Sec. 1.3).¹³ The cluster operator $\hat{T}$ in general includes up to N -fold excitations, which is equivalent to full CI. For practical applications, however, the $\hat{T}$ operator is usually truncated at the single and double excitation level ($\hat{T} = \hat{T}_1 + \hat{T}_2$, known as CCSD)¹³ or includes up to triple (in CCSDT)²⁴⁻²⁷ or quadrupole (in CCSDTQ)^{28,29} excitations, if high accuracy is desired. Whereas CC theory offers a well-defined hierarchy of methods that converge to the full CI limit, the evaluation of the molecular properties using CC methods is complicated due to the non-linear nature of the amplitude equations and the asymmetric energy functional.³⁰⁻³³ A less expensive alternative to CC theory is Møller-Plesset perturbation theory (MPPT),^{34,35} which allows to approach the electron correlation

problem based on an order-by-order expansion of the wavefunction and the energy. While MPPT can be formulated to any given order, only the second-, third-, and fourth-order MPPT (MP2,^{36–38} MP3,^{14,39,40} and MP4,^{40–42} respectively) found a wide application. It has been shown that MPPT provides poor results and may not converge to full CI in situations when the Hartree-Fock reference is inadequate zeroth-order approximation.⁴³

1.3 EXTENSIVITY AND SIZE-CONSISTENCY

Solving the electronic Schrödinger equation approximately requires a careful choice of approximations to be made. Of particular importance are the approximate solutions that retain as many important properties of the exact wavefunction as possible. Among some of these properties are antisymmetry, square-integrability, invariance with respect to orbital transformations, spatial and spin symmetry, extensivity and size-consistency.^{1,21} While some properties are trivially satisfied if one expands the approximate wavefunction in the basis of Slater determinants, the fulfillment of other properties is not guaranteed. In particular, extensivity and size-consistency,¹² very important for the quantum-mechanical prediction of thermochemical properties, can be easily violated and need to be taken care of.

A quantum-mechanical model is said to be *extensive* (or *size-extensive*)^{13,20,21,44} if its energy scales linearly with the number of electrons N , just as the exact energy does. A precise mathematical definition of extensivity for arbitrary systems is rather complicated.⁴⁵ It can be shown that a truncated CI model is not size-extensive.^{20,22} For example, in the case of a system composed of the N noninteracting He atoms¹³ the correlation energy of the CI method restricted to double excitations (CID) is proportional to $\sqrt{N}$. Thus, as the number of the noninteracting subunits of the system increases (increasing the number of electrons), the CID correlation energy will begin to deviate from that of full CI, producing large extensivity errors for big values of N . In practice, the truncated CI methods provide poor quality results for the molecular systems with more than ten correlated electrons.

A closely related concept to extensivity is *size-consistency* (or *separability*).^{14,39,46} Let us consider a super system containing two noninteracting subsystems, which we label A and B . The exact wavefunction of the super system can be written as a product of the wavefunctions describing each subsystem independently:

$$\Psi_{AB} = \Psi_A \Psi_B . \quad (1.9)$$

Similarly, the exact energy of the super system can be expressed as a sum of the energies of the two subunits:

$$E_{AB} = E_A + E_B . \quad (1.10)$$

As in the case with size-extensivity, the truncated CI methods are known to violate size-consistency, which prohibits their use for the accurate prediction of thermochemical properties. The lack of separability of the CI wavefunction can be easily demonstrated²² if we write the CI ansatz (1.7) in the following form:

$$\Psi_{\text{CI}} = (1 + \hat{C}) \Phi_{\text{HF}} , \quad (1.11)$$

where the excitation operator $\hat{C}$ generates all possible excited Slater determinants out of the reference Φ_{HF} , similarly to the $\hat{T}$ operator in CC theory. Choosing a localized orbital representation,^{*} the excitation operator $\hat{C}$ can be separated into components involving the intrafragment excitations only, i.e. $\hat{C} = \hat{C}_A + \hat{C}_B$. As a result, for the total CI wavefunction we get:

$$\Psi_{\text{CI}} = (1 + \hat{C}_A + \hat{C}_B) \Phi_{\text{HF}} . \quad (1.12)$$

^{*}The localized orbital representation is chosen for convenience and does not lead to a loss of generality in the discussion.

From Eq. (1.12) one can see that the linear CI ansatz does not guarantee the multiplicative separability (1.9) of the total wavefunction Ψ_{CI} . Indeed, the CI wavefunction (1.12) is rigorously size-consistent only in the limit of the complete excitation operator $\hat{C}$, i.e. full CI wavefunction.

Let us now approach the description of the super system AB from the standpoint of CC theory. In the localized representation we can write

$$\Psi_{\text{CC}} = e^{\hat{T}_A + \hat{T}_B} \Phi_{\text{HF}}^A \Phi_{\text{HF}}^B = (e^{\hat{T}_A} \Phi_{\text{HF}}^A) (e^{\hat{T}_B} \Phi_{\text{HF}}^B) = \Psi_{\text{CC}}^A \Psi_{\text{CC}}^B, \quad (1.13)$$

where we used the separability of the HF reference in the localized basis. Eq. (1.13) demonstrates that the CC wavefunction is multiplicatively separable for an arbitrary $\hat{T}$ operator. As a result, the computed CC energies are size-consistent even when approximations are introduced. The exponential ansatz also guarantees size-extensivity of the CC theory that manifests itself in the *connected* amplitude equations, i.e. those that only contain terms, in which the Hamiltonian and the cluster operator $\hat{T}$ have common indices.^{13,21,47–51}

1.4 REDUCED DENSITY MATRIX METHODS

In Sec. 1.2 we explored several approaches to the electron correlation based on the expansion of the molecular wavefunction in the basis of Slater determinants. Truncating this expansion at a certain excitation rank (e.g., single, double, triple ... excitations), one obtains a hierarchy of methods that converges to the full CI limit. As the truncation rank increases the complexity of the approximate wavefunction scales rapidly, approaching the exponential complexity of the exact wavefunction and making the *ab initio* electron correlation methods prohibitively expensive for large molecular systems. For example, the CC method incorporating the single, double and triple excitations (CCSDT) scales as $\mathcal{O}(M^8)$ with the size of the basis set M .

It has long been known that the exact energy can be computed from the reduced one- and two-particle density matrices.^{8,52–56} Ignoring the normalization factor, these matrices can be defined as the one- and two-particle probability densities (γ and Γ , Eq. (1.3)) given in Sec. 1.1. We will instead use a different notation here and define the reduced density matrices using the language of second quantization:^{22,57,58}

$$\begin{aligned}\gamma_p^q &= \langle \Psi | a_p^q | \Psi \rangle , & \gamma_{pq}^{rs} &= \langle \Psi | a_{pq}^{rs} | \Psi \rangle , \\ a_p^q &\equiv a_q^\dagger a_p , & a_{pq}^{rs} &\equiv a_r^\dagger a_s^\dagger a_q a_p ,\end{aligned}\tag{1.14}$$

where $a_p^\dagger$ and a_p are the creation and annihilation second-quantized operators, respectively; γ_p^q and γ_{pq}^{rs} are the elements of the one- and two-particle reduced density matrices ($\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$). Taking the expectation value of the second-quantized Hamiltonian⁵⁷

$$\begin{aligned}\hat{H} &= h_p^q a_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} a_{rs}^{pq} , & h_p^q &= \langle \psi_p(1) | \hat{h} | \psi_q(1) \rangle , \\ \bar{g}_{pq}^{rs} &= g_{pq}^{rs} - g_{pq}^{sr} , & g_{pq}^{rs} &= \langle \psi_p(1) \psi_q(2) | \frac{1}{r_{12}} | \psi_r(1) \psi_s(2) \rangle ,\end{aligned}\tag{1.15}$$

we can write the exact energy as a functional of $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$, as well as the one- and antisymmetrized two-body integrals (h_p^q and $\bar{g}_{pq}^{rs}$):

$$E = \langle \Psi | \hat{H} | \Psi \rangle = h_p^q \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \gamma_{rs}^{pq} .\tag{1.16}$$

In Eqs. (1.15) and (1.16) the summation over the repeated indices is implied.⁵⁷ The simplicity of the energy functional (1.16) is a direct result of the one- and two-body nature of the molecular Hamiltonian. The fact that the exact (or sufficiently accurate) energy can in principle be obtained from the one- and two-particle reduced density matrices bypassing the computation of the many-dimensional wavefunction was realized in the early 1960's,^{59–62} which led to the development of the reduced density matrix (RDM) theory.^{63,64}

In the RDM theory the energy expression (1.16) is simplified using the partial trace relation between γ_1 and γ_2 :

$$\sum_r \gamma_{pr}^{qr} = (N-1)\gamma_p^q . \quad (1.17)$$

Expressing γ_1 in terms of γ_2 via Eq. (1.17), one can write the energy (1.16) solely in terms of γ_2 :

$$E = \frac{1}{2} \tilde{H}_{pq}^{rs} \gamma_{rs}^{pq} , \quad \tilde{H}_{pq}^{rs} \equiv \frac{1}{N-1} (h_p^r \delta_q^s + h_q^s \delta_p^r) + g_{pq}^{rs} , \quad (1.18)$$

where the elements of the two-particle reduced Hamiltonian $\tilde{H}_{pq}^{rs}$ were defined. Unfortunately, direct variational minimization^{60,65–67} of the functional (1.18) yields the energies that are much lower than that of full CI and needs to be constrained to ensure that γ_1 and γ_2 correspond to an antisymmetric N -electron wavefunction, *i.e.* N -representable.^{59,60,62,68–75}

N -representability problem is the crux of the RDM-based methods. Unfortunately, no simple set of the necessary and sufficient conditions that enforce N -representability are known and some of the known conditions are not easily imposed.⁷⁶ Among the known necessary conditions are the positive semidefiniteness of the diagonal elements of the two-particle (γ_2 , Eq. (1.14)), particle-hole and two-hole density matrices (β_2 and η_2 , respectively):^{60,65,77}

$$\gamma_{pq}^{pq} \geq 0 , \quad \beta_{pq}^{qp} \geq 0 , \quad \eta_{pq}^{pq} \geq 0 , \quad (1.19)$$

where the elements of β_2 and η_2 are defined as:⁷⁷

$$\beta_{pr}^{qs} = \langle \Psi | a_q^\dagger a_p a_s^\dagger a_r | \Psi \rangle , \quad \eta_{rs}^{pq} = \langle \Psi | a_r a_s a_q^\dagger a_p^\dagger | \Psi \rangle . \quad (1.20)$$

The conditions for the off-diagonal elements can be obtained using the Cauchy–Schwarz inequality (no summation is implied):

$$\gamma_{rs}^{pq} \gamma_{pq}^{rs} \leq \gamma_{rs}^{rs} \gamma_{pq}^{pq} , \quad \beta_{qs}^{pr} \beta_{rp}^{sq} \leq \beta_{qp}^{pq} \beta_{rs}^{sr} , \quad \eta_{rs}^{pq} \eta_{pq}^{rs} \leq \eta_{rs}^{rs} \eta_{pq}^{pq} . \quad (1.21)$$

N -representability conditions (1.19) and (1.21) can be enforced using the technique of semidefinite programming.^{76,78–81} However, the computational cost of these methods has limited their widespread adoption. One of the most significant drawbacks of the RDM-based methods is their lack of extensivity and size-consistency, which originate from the nonseparability of the two-particle reduced density matrix. Indeed, variational RDM methods have been shown to produce significant errors in size-extensivity and size-consistency,^{82,83} as well as the incorrect description of the dissociation limits with fractional charges on the well separated atoms.^{83,84}

1.5 DENSITY CUMULANTS

In Sec. 1.1 it has been demonstrated that in the absence of the Coulomb correlation the two-particle probability density Γ can be entirely written in terms of the one-particle density (Eq. (1.5)). Analogous expression for the two-particle reduced density matrix γ_2 in the tensor notation has the following form:

$$\gamma_{pq}^{rs} = \gamma_p^r \gamma_q^s - \gamma_p^s \gamma_q^r . \quad (1.22)$$

Importantly, Eq. (1.22) demonstrates that in the mean-field approximation γ_2 is redundant, i.e. can be fully expressed in terms of γ_1 . In the case of the correlated system, however, Eq. (1.22) does not hold. In 1997, Kutzelnigg and Mukherjee derived the relationship between γ_1 and γ_2 for an arbitrary molecular system:⁵⁸

$$\gamma_{pq}^{rs} = \gamma_p^r \gamma_q^s - \gamma_p^s \gamma_q^r + \lambda_{pq}^{rs} , \quad (1.23)$$

where λ_{pq}^{rs} are the elements of the two-particle density cumulant λ_2 ,^{58,77,85–91} which represents the irreducible part of γ_2 that cannot be expressed through γ_1 . Comparing

Eqs. (1.22) and (1.23), we realize that in the Hartree-Fock approximation (i.e. for a single determinant wavefunction) the density cumulant λ_2 is zero.

It may seem somewhat surprising that until 1997 cumulants have not been used in the density matrix theory applied to electron correlation. Cumulants naturally arise in probability theory and statistics in the analysis of the correlated variables.⁸⁷ Similarly, in the expansion of the two-particle density matrix (1.23) λ_2 describes the pairwise correlation between two electrons.⁸⁷ More specifically, the two-body cumulant is a measure of the correlation between the occupation numbers of two spin-orbitals.^{90,91} General definition of the density cumulants of an arbitrary order (λ_k , $k \geq 1$) can be given via a generating function.^{85,87} For $k = 1$ the one-particle cumulants are simply $\lambda_1 = \gamma_1$. The density cumulants λ_2 , like γ_2 , are Hermitian and antisymmetric, i.e.

$$\lambda_{pq}^{rs} = (\lambda_{rs}^{pq})^* , \quad \lambda_{pq}^{rs} = -\lambda_{qp}^{rs} = -\lambda_{pq}^{sr} = \lambda_{qp}^{sr} . \quad (1.24)$$

The most important property of density cumulants λ_k is that they are extensive and additively separable, unlike their γ_k counterparts (for $k > 1$).^{58,87} The extensivity of λ_k implies that in a graphical representation they can only contain fully connected diagrams.⁸⁷ Thus, formulating the theory in terms of the density cumulants guarantees size-extensivity and size-consistency, which cannot be rigorously preserved in the RDM theories.

An interesting property of the density cumulants is that for an N -electron system the elements of λ_k with $k > N$ do not vanish.^{87,92} Instead, λ_k are non-zero and can be entirely expressed in terms of the cumulants of the lower order.⁹² For example, for a two-electron

system the elements of λ_3 can be obtained from the requirement that γ_3 is zero:

$$\begin{aligned} \gamma_{pqr}^{stu} = & \lambda_{pqr}^{stu} + \gamma_p^s \lambda_{qr}^{tu} - \gamma_p^t \lambda_{qr}^{su} + \gamma_p^u \lambda_{qr}^{st} - \gamma_q^s \lambda_{pr}^{tu} + \gamma_q^t \lambda_{pr}^{su} - \gamma_q^u \lambda_{pr}^{st} + \gamma_r^s \lambda_{pq}^{tu} - \gamma_r^t \lambda_{pq}^{su} + \gamma_r^u \lambda_{pq}^{st} \\ & + \gamma_p^s \gamma_q^t \gamma_r^u - \gamma_p^s \gamma_r^t \gamma_q^u + \gamma_q^s \gamma_r^t \gamma_p^u - \gamma_q^s \gamma_p^t \gamma_r^u + \gamma_r^s \gamma_p^t \gamma_q^u - \gamma_r^s \gamma_q^t \gamma_p^u = 0 . \end{aligned} \quad (1.25)$$

It has been suggested to consider λ_k for $k > N$ as disconnected.⁹³ Recently, Hanauer and Köhn⁹¹ argued that in the basis of M spin-orbitals λ_k with $k > N$ have physical significance, since the density cumulants describe the correlation between the N particles and $(M - N)$ holes on equal footing and do not necessarily truncate at either rank.

1.6 PROSPECTUS

After its introduction into electronic structure theory in 1997, density cumulants have been used in the framework of the generalized normal ordering⁵⁸ and as a tool for the reconstruction of the many-particle reduced density matrices in the contracted Schrödinger equation theory.^{67,86,94,95} In 2006, Kutzelnigg demonstrated that the exact electronic energy can be entirely expressed in terms of the density cumulants and developed a cumulant-based approach called *density cumulant functional theory* (DCFT).⁹⁶ After introducing several approximations he arrived at the size-extensive energy functional and a simple set of stationarity conditions that resembled that of linearized coupled-cluster theory. In 2010, Simmonett and co-workers performed the first implementation of the Kutzelnigg's DCFT formulation.⁹⁷ Their preliminary benchmark results demonstrated that the quality of the method (denoted as DCFT-06) ranged between that of CCSD and CCSD with perturbative triple excitations [CCSD(T)], while its efficiency was much higher, compared to CCSD.

Herein, we present several contributions to the development of DCFT, aiming for extending the applicability of this theory and improving its accuracy. In Chapter 2, we develop the

analytic gradient theory for the DCFT-06 method. The use of the analytic gradients greatly simplifies the computation of the equilibrium geometries and first-order molecular properties. In this work, we demonstrate that in order to obtain the DCFT-06 analytic gradient a system of coupled response equations needs to be solved. We demonstrate the solution of these equations and present a benchmark study of the DCFT-06 molecular properties for a set of chemical systems. In Chapter 3, we present a new formulation of DCFT called DC-12, where the N -representability of the one-particle density matrix is greatly improved by deriving the correlation contribution to γ_1 from λ_2 exactly. To assess the importance of this effect, the perturbative analysis and the benchmark studies of the new method are demonstrated. Finally, in Chapter 4, a new orbital-optimized variant of DCFT is developed, where the molecular orbitals are variationally optimized. This leads to significant simplification of the analytic gradient theory and the improvement of the accuracy, which we demonstrate using perturbative analysis and numerical results.

CHAPTER 2

ANALYTIC GRADIENTS FOR DENSITY CUMULANT FUNCTIONAL THEORY: THE DCFT-06 MODEL*

*A. Yu. Sokolov, J. J. Wilke, A. C. Simmonett, and H. F. Schaefer, J. Chem. Phys. **137**, 054105 (2012).
Reproduced with permission of the American Institute of Physics.

2.1 ABSTRACT

Density cumulant functional theory (DCFT) is one of a number of nascent electron correlation methods that are derived from reduced density matrices and cumulants thereof, instead of the wavefunction. Deriving properties from the density cumulant naturally yields methods that are size-extensive and size-consistent. In this work we derive expressions for the analytic gradient, with respect to an external perturbation, for the DCFT-06 variant of density cumulant functional theory. Despite the fact that the DCFT-06 energy functional is stationary with respect to the density cumulant, the analytic gradients of the energy require the solution of perturbation-independent equations for both orbital and cumulant response. These two sets of linear response equations are coupled in nature and are solved iteratively with the solution of orbital and cumulant response equations each macroiteration, exhibiting rapid convergence. The gradients are implemented and benchmarked against coupled cluster theory with single and double excitations (CCSD) and CCSD with perturbative triple excitations [CCSD(T)], as well as accurate empirically-corrected experimental data, for a test set comprising 15 small molecules. For most of the test cases, results from DCFT-06 are closer to CCSD(T) and empirical data than those from CCSD. Although the total energy and analytic gradient have the same asymptotic scaling, the present experience shows that the computational cost of the gradient is significantly lower.

2.2 INTRODUCTION

The chemical problems that can be accurately studied by coupled cluster theory^{13,22,23} are many and varied. The pursuit of similarly accurate *ab initio* tools for even larger chemical systems has seen the development of density based methods such as the density matrix renormalization group (DMRG)^{98–101} for treating static correlation, and its canonical transformation theory^{102–104} extension, as well as parametric reduced density matrix meth-

ods^{64,105–108} and density cumulant functional theory (DCFT).^{96,97} When developing density-based approaches, care must be taken to ensure that the resulting density matrices at least approximately correspond to an antisymmetric n -particle wavefunction, *i.e.*, that they are n -representable. Wavefunction approaches that seek similar computational savings can be motivated by reducing the nonlinearity of the coupled-cluster equations,¹⁰⁹ or introducing size-consistency and extensivity to configuration interaction theory.^{110–115}

In DCFT the energy is expressed exactly in terms of the cumulant of the two-particle reduced density matrix,^{58,77,86,87,90} a Hermitian and extensive quantity, which leads to size consistent and extensive theories. Working equations can be derived by making the functional stationary with respect to the cumulant, as long as the corresponding density is constrained to be at least approximately n -representable. In the DCFT-06 method, approximate n -representability conditions derived from perturbation theory by Kutzelnigg^{92,96} in 2006 are used; these take the form of equations relating all non-zero cumulant elements to a minimal number of independent parameters. We have recently reported the first implementation and numerical results for DCFT-06.⁹⁷ In the present research, we extend the applicability of the DCFT-06 model by deriving and implementing analytic gradients of the energy, with respect to nuclear motion.

The development of analytic energy derivatives has been one of the major achievements of quantum chemistry in the past few decades. Analytic gradients have been derived for variational correlation methods such as configuration interaction (CI),^{116,117} as well as non-variational methods, *e.g.*, second-order Møller-Plesset perturbation theory (MP2)^{31,118} and coupled cluster theory.^{30,32,119–121} More recently, gradients have been derived for coupled cluster models based on non-canonical reference functions, such as projectively optimized Brueckner orbitals¹²² and MP2 natural orbitals.¹²³ In the case of non-variational methods the computation of analytic gradients is complicated by the requisite equations that describe the response of the wavefunction parameters to the perturbation. As we will show, the

DCFT-06 energy functional is stationary, resembling CI methods, but response equations must be solved in order to compute its derivatives, as is the case in coupled cluster.

2.3 GENERAL THEORY

2.3.1 EXACT DENSITY CUMULANT FUNCTIONAL THEORY

The exact energy of a many-electron system can be written as an expectation value

$$E = \langle \Psi | \hat{H} | \Psi \rangle = \langle \Psi | h_p^q a_q^p + \frac{1}{2} g_{pq}^{rs} a_{rs}^{pq} | \Psi \rangle = h_p^q \gamma_q^p + \frac{1}{2} g_{pq}^{rs} \gamma_{rs}^{pq} \quad (2.1)$$

where $\hat{H}$ is the molecular Hamiltonian in the second-quantized form, a_p^q and a_{pq}^{rs} are excitation operators^{13,22}

$$a_p^q = a_q^\dagger a_p, \quad a_{pq}^{rs} = a_r^\dagger a_s^\dagger a_q a_p, \quad (2.2)$$

h_p^q and g_{pq}^{rs} are the usual one- and two-electron integrals

$$h_p^q = \langle \psi_p(1) | \hat{h} | \psi_q(1) \rangle, \quad g_{pq}^{rs} = \langle \psi_p(1) \psi_q(2) | \frac{1}{r_{12}} | \psi_r(1) \psi_s(2) \rangle, \quad (2.3)$$

and γ_p^q and γ_{pq}^{rs} are the elements of the reduced one- and two-particle density matrices

$$\gamma_p^q = \langle \Psi | a_p^q | \Psi \rangle, \quad \gamma_{pq}^{rs} = \langle \Psi | a_{pq}^{rs} | \Psi \rangle. \quad (2.4)$$

In density cumulant functional theory^{96,97} (DCFT) the two-particle density matrix γ_2 is eliminated in favor of the two-particle density cumulant λ_2 ^{58,77,86,87,90}

$$\lambda_{pq}^{rs} = \gamma_{pq}^{rs} - \gamma_p^r \gamma_q^s + \gamma_p^s \gamma_q^r. \quad (2.5)$$

The one-particle density matrix γ_1 is subsequently decomposed into two contributions; its best idempotent approximation¹²⁴ κ and a non-idempotent correction⁹⁶ τ :

$$\gamma_p^q = \kappa_p^q + \tau_p^q, \quad \kappa_p^r \kappa_r^q = \kappa_p^q. \quad (2.6)$$

The non-idempotent contribution depends exclusively on the density cumulant. The energy is then written as a functional of the idempotent contribution to the one-particle density matrix and the density cumulant:

$$E = h_p^q \gamma_q^p + \frac{1}{2} \bar{g}_{pq}^{rs} \gamma_r^p \gamma_s^q + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq} = \frac{1}{2} (h_p^q + f_p^q) \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq}, \quad (2.7)$$

where we define antisymmetrized two-electron integrals

$$\bar{g}_{pq}^{rs} = g_{pq}^{rs} - g_{pq}^{sr} \quad (2.8)$$

and the generalized Fock operator matrix

$$f_p^q = h_p^q + \bar{g}_{pr}^{qs} \gamma_s^r. \quad (2.9)$$

It is convenient to decompose f_p^q into two contributions

$$f_p^q = (f_0)_p^q + (f')_p^q, \quad (f_0)_p^q = h_p^q + \bar{g}_{pr}^{qs} \kappa_s^r, \quad (f')_p^q = \bar{g}_{pr}^{qs} \tau_s^r. \quad (2.10)$$

The first contribution $(f_0)_p^q$ corresponds to the Fock operator used in Hartree-Fock theory, whereas $(f')_p^q$ is the correction that depends on the density cumulant and describes the effects of electron correlation on the orbitals.

2.3.2 DCFT-06 ENERGY

In the DCFT-06 method^{96,97} one takes care of the n -representability of the density cumulant by using approximate conditions derived from second-order perturbation theory.⁹² Inserting these conditions into Eq. (2.7), we obtain the DCFT-06 energy expression:

$$E = \frac{1}{2}(h_p^q + f_p^q)\gamma_q^p + \frac{1}{2}\bar{g}_{ij}^{ab}\lambda_{ab}^{ij} + \frac{1}{8}\lambda_{kl}^{ab}\bar{g}_{ij}^{kl}\lambda_{ab}^{ij} + \frac{1}{8}\lambda_{ij}^{cd}\bar{g}_{cd}^{ab}\lambda_{ab}^{ij} - \lambda_{ik}^{ac}\bar{g}_{jc}^{kb}\lambda_{ab}^{ij}. \quad (2.11)$$

All five terms of Eq. (2.11) depend on the density cumulant. The dependence of the first term on λ_2 is implicit by means of Eqs. (2.6) and (2.10), and the approximate relation^{92,96} between the non-idempotent correction τ and the cumulant λ_2 :

$$\tau_i^j \approx -\frac{1}{2}\lambda_{ik}^{ab}\lambda_{ab}^{jk}, \quad \tau_a^b \approx \frac{1}{2}\lambda_{ac}^{ij}\lambda_{ij}^{bc}. \quad (2.12)$$

The stationarity conditions of the energy (2.11) with respect to variation of the idempotent part of the one-particle density matrix (κ) are¹²⁵

$$[\mathbf{f}, \kappa] = 0 \quad \Rightarrow \quad f_i^a = 0, \quad (2.13)$$

which can be satisfied by self-consistently diagonalizing the generalized Fock operator matrix $\mathbf{f}$. The resulting orbitals yield the best single determinant approximation to the density, unlike Brueckner orbitals, which have maximal overlap with the exact wave function, or the Hartree-Fock orbitals, which yield the lowest energy for a single-determinant wave function.¹²⁴ The equations for the density cumulant are obtained by making the energy stationary with respect to the variation of λ_2 :⁹⁷

$$\frac{\partial E}{\partial \lambda_{ij}^{ab}} = R_{ab}^{ij} = 0, \quad (2.14)$$

where the residual R_{ab}^{ij} can be written as:

$$R_{ab}^{ij} = P_{-}(ab) [f_a^c \lambda_{cb}^{ij}] - P_{-}(ij) [f_k^i \lambda_{ab}^{kj}] + \bar{g}_{ab}^{ij} + \frac{1}{2} \bar{g}_{kl}^{ij} \lambda_{ab}^{kl} + \frac{1}{2} \bar{g}_{ab}^{cd} \lambda_{cd}^{ij} - P_{-}(ij) P_{-}(ab) [\bar{g}_{kb}^{jc} \lambda_{ac}^{ik}]. \quad (2.15)$$

In Eq. (2.15) we introduced the antisymmetric permutation operator $P_{-}(ij)f(i, j) \equiv f(i, j) - f(j, i)$ that acts only on the terms taken in square brackets. Equation (2.14) is non-linear as a result of the implicit dependence of $\mathbf{f}$ on $\boldsymbol{\lambda}_2$ through $\boldsymbol{\tau}$ (Eqs. (2.10) and (2.12)). By applying the stationarity conditions from Eq. (2.14) to the energy expression (2.11), we obtain a simple equation for the energy at convergence:

$$E = \frac{1}{2} (h_p^q + (f_0)_p^q) \kappa_q^p - \frac{1}{2} \bar{g}_{pr}^{qs} \tau_q^p \tau_s^r + \frac{1}{4} \bar{g}_{ij}^{ab} \lambda_{ab}^{ij}. \quad (2.16)$$

The first term of this equation can be related to the Hartree-Fock energy contribution, whereas the last two terms describe electron correlation effects.

2.3.3 DCFT-06 LAGRANGIAN

To derive expressions for the analytic gradient of the DCFT-06 energy it is convenient to use the method of Lagrange multipliers.^{126–128} We start by defining the energy functional

$$\mathcal{L} = E + \frac{1}{4} Z_{ij}^{ab} R_{ab}^{ij} + z_i^a f_a^i + \omega_p^q (S_p^q - \delta_p^q), \quad (2.17)$$

where S_p^q are the overlap integrals and δ_p^q is the Kronecker delta. The functional $\mathcal{L}$ is obtained by augmenting the DCFT-06 energy expression E (2.16) with a set of constraints corresponding to the energy stationarity conditions (Eqs. (2.14) and (2.13)) and the orthonormality condition for the orbitals. The functional $\mathcal{L}$ is often referred to as the Lagrangian, and the coefficients Z_{ij}^{ab} , z_i^a and ω_p^q are the corresponding Lagrange multipliers.

2.3.4 DCFT-06 RESPONSE EQUATIONS

In the next step, the Lagrangian $\mathcal{L}$ is made stationary with respect to all of its independent variables. Despite the fact that the DCFT-06 energy (2.16) is stationary with respect to the density cumulant, this stationarity condition is no longer satisfied by the Lagrangian (2.17); this is a corollary of the cumulant dependence of the stationarity condition with respect to κ (2.13). Thus, the derivative of this condition with respect to λ_2 is non-zero:

$$\frac{\partial(z_k^c f_c^k)}{\partial \lambda_{ab}^{ij}} = z_k^c \{ P_-(ab) [(\bar{g}_{cd}^{ka} + \bar{g}_{ca}^{kd}) \lambda_{ij}^{db}] - P_-(ij) [(\bar{g}_{cl}^{ki} + \bar{g}_{ci}^{kl}) \lambda_{lj}^{ab}] \} \neq 0. \quad (2.18)$$

By making the functional $\mathcal{L}$ stationary with respect to the density cumulant, we obtain the following system of equations:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda_{ab}^{ij}} = & P_-(ab) [f_c^a Z_{ij}^{cb}] - P_-(ij) [f_i^k Z_{kj}^{ab}] + \bar{g}_{ab}^{ij} + \frac{1}{2} \bar{g}_{ij}^{kl} Z_{kl}^{ab} + \frac{1}{2} \bar{g}_{cd}^{ab} Z_{ij}^{cd} \\ & - P_-(ij) P_-(ab) [\bar{g}_{jc}^{kb} Z_{ik}^{ac}] + 2P_-(ab) [\bar{g}_{dc}^{ea} \lambda_{ij}^{cb}] \Delta \tilde{\tau}_e^d - 2P_-(ij) [\bar{g}_{ci}^{dk} \lambda_{kj}^{ab}] \Delta \tilde{\tau}_d^c \\ & + 2P_-(ab) [\bar{g}_{kc}^{la} \lambda_{ij}^{cb}] \Delta \tilde{\tau}_l^k - 2P_-(ij) [\bar{g}_{ki}^{ml} \lambda_{lj}^{ab}] \Delta \tilde{\tau}_m^k \\ & + z_k^c \{ P_-(ab) [(\bar{g}_{cd}^{ka} + \bar{g}_{ca}^{kd}) \lambda_{ij}^{db}] - P_-(ij) [(\bar{g}_{cl}^{ki} + \bar{g}_{ci}^{kl}) \lambda_{lj}^{ab}] \} = 0, \end{aligned} \quad (2.19)$$

where we introduced $\Delta \tilde{\tau}_p^q$

$$\Delta \tilde{\tau}_p^q = \tilde{\tau}_p^q - \tau_p^q \quad (2.20)$$

and defined the relaxed non-idempotent contribution to the one-particle density matrix

$$\tilde{\tau}_i^j = -\frac{1}{4} P_+(ij) [Z_{ik}^{ab} \lambda_{ab}^{jk}], \quad \tilde{\tau}_a^b = \frac{1}{4} P_+(ab) [Z_{ij}^{ac} \lambda_{bc}^{ij}]. \quad (2.21)$$

In Eq. (2.21) the symmetric permutation operator $P_+(ij)f(i, j) \equiv f(i, j) + f(j, i)$ was used to guarantee the symmetric property of the relaxed one-particle density matrix.

To obtain the stationarity conditions of the Lagrangian with respect to orbital rotations, we will first rewrite Eq. (2.17) in the following form:

$$\mathcal{L} = \tilde{\gamma}_i^j h_j^i + \tilde{\gamma}_a^b h_b^a + \Gamma_{ij}^{kl} \tilde{g}_{kl}^{ij} + \Gamma_{ab}^{cd} \tilde{g}_{cd}^{ab} + \Gamma_{ij}^{ab} \tilde{g}_{ab}^{ij} + \Gamma_{ia}^{jb} \tilde{g}_{jb}^{ia} + z_i^a f_a^i + \omega_p^q (S_q^p - \delta_q^p) \quad (2.22)$$

where $\tilde{\gamma}_p^q$ and Γ_{pq}^{rs} are the elements of the relaxed one- and two-particle density matrices, respectively:

$$\tilde{\gamma}_i^j = \kappa_i^j + \tilde{\tau}_i^j, \quad \tilde{\gamma}_a^b = \tilde{\tau}_a^b, \quad (2.23)$$

$$\begin{aligned} \Gamma_{ij}^{kl} = & \frac{1}{4} P_-(kl) [\kappa_i^k \kappa_j^l] + \frac{1}{4} P_-(ij) P_-(kl) [\gamma_i^k \tilde{\tau}_j^l] \\ & - \frac{1}{4} P_-(kl) [\tau_i^k \tau_j^l] + \frac{1}{16} (Z_{ij}^{ab} \lambda_{ab}^{kl} + Z_{kl}^{ab} \lambda_{ab}^{ij}), \end{aligned} \quad (2.24)$$

$$\Gamma_{ab}^{cd} = \frac{1}{4} P_-(ab) P_-(cd) [\tau_a^c \tilde{\tau}_b^d] - \frac{1}{4} P_-(cd) [\tau_a^c \tau_b^d] + \frac{1}{16} (Z_{ij}^{ab} \lambda_{cd}^{ij} + Z_{ij}^{cd} \lambda_{ab}^{ij}), \quad (2.25)$$

$$\Gamma_{ij}^{ab} = \frac{1}{4} (Z_{ij}^{ab} + \lambda_{ab}^{ij}), \quad (2.26)$$

$$\Gamma_{ia}^{jb} = \gamma_i^j \tilde{\tau}_a^b + \tau_a^b \Delta \tilde{\tau}_i^j - \frac{1}{2} (Z_{ki}^{cb} \lambda_{ca}^{kj} + Z_{kj}^{ca} \lambda_{cb}^{ki}). \quad (2.27)$$

Making the rewritten Lagrangian (2.22) stationary with respect to the rotation between the occupied and virtual orbitals yields the system of equations:

$$z_i^a (\varepsilon_a - \varepsilon_i) + z_k^c [(\bar{g}_{cj}^{ka} + \bar{g}_{ca}^{kj}) \gamma_i^j - (\bar{g}_{cb}^{ki} + \bar{g}_{ci}^{kb}) \gamma_a^b] = 2(X_a^i - X_i^a), \quad (2.28)$$

where ε_p are the eigenvalues of the generalized Fock operator (2.9). The elements X_p^q that appear in the right-hand side of Eq. (2.28) are defined in terms of the relaxed density matrices as

$$X_a^p = \tilde{\gamma}_a^b h_b^p + 2\Gamma_{ab}^{cd} \tilde{g}_{cd}^{pb} + \Gamma_{ij}^{ab} \tilde{g}_{pb}^{ij} + \Gamma_{ia}^{jb} \tilde{g}_{jb}^{ip} \quad (2.29)$$

and

$$X_i^p = \tilde{\gamma}_i^j h_j^p + 2\Gamma_{ij}^{kl} \tilde{g}_{kl}^{pj} + \Gamma_{ij}^{ab} \tilde{g}_{ab}^{pj} + \Gamma_{ia}^{jb} \tilde{g}_{jb}^{pa}. \quad (2.30)$$

Two systems of linear equations, Eqs. (2.19) and (2.28), describe the response of the density cumulant and orbitals to an external perturbation. The Lagrange multipliers Z_{ij}^{ab} and z_i^a obtained as the solution of these response equations constitute the orbital and the cumulant response Z-vectors¹²⁹ ($\mathbf{Z}$ and $\mathbf{z}$), which are the perturbation-independent quantities that account for the relaxation of the density cumulant and orbitals as the result of the perturbation. Careful examination of the cumulant and orbital response equations reveals that these two sets of equations are coupled in nature, with two sets of response matrix elements Z_{ij}^{ab} and z_i^a depending on each other. Thus, equations (2.19) and (2.28) must be solved iteratively, with the solution of the orbital and cumulant response equations each macroiteration; this is reminiscent of the process used to make the energy stationary with respect to $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}_2$. Analogous coupled set of orbital and amplitude response equations was previously obtained for the coupled-cluster doubles method (CCD) that utilizes projective Brueckner orbitals.^{122,130} The appearance of the coupling term (2.18) in the cumulant response equations (2.19) is the direct result of the dependence of the generalized Fock operator (2.9) on the density cumulant. Although the use of non-canonical orbitals leads to the necessity of the solution of the additional set of response equations, the orbital relaxation and the non-linear terms arising from this coupling are expected to be important. Moreover, our experience shows that the magnitude of the coupling term is generally small and the coupled response equations converge fast. Thus, most of the relaxed density is provided by the solution of the cumulant residual equations (2.14) obtained for the DCFT-06 energy.

2.3.5 DCFT-06 ENERGY GRADIENT EXPRESSION

The stationarity of the Lagrangian (2.22) with respect to all of its variables allows us to write the expression of the energy gradient with respect to a perturbation χ in a simple form:

$$\begin{aligned} \frac{dE}{d\chi} = \frac{\partial \mathcal{L}}{\partial \chi} = & \tilde{\gamma}_i^j (h_j^i)^\chi + \tilde{\gamma}_a^b (h_b^a)^\chi + \tilde{\gamma}_i^a (h_a^i)^\chi + \Gamma_{ij}^{kl} (\bar{g}_{kl}^{ij})^\chi + \Gamma_{ab}^{cd} (\bar{g}_{cd}^{ab})^\chi \\ & + \Gamma_{ij}^{ab} (\bar{g}_{ab}^{ij})^\chi + \Gamma_{ia}^{jb} (\bar{g}_{jb}^{ia})^\chi + \Gamma_{ij}^{ak} (\bar{g}_{ak}^{ij})^\chi + \Gamma_{ib}^{ac} (\bar{g}_{ac}^{ib})^\chi + \omega_p^q (S_q^p)^\chi, \end{aligned} \quad (2.31)$$

where the only perturbed quantities are the skeleton one-, antisymmetrized two-electron, and overlap derivative integrals $((h_p^q)^\chi, (\bar{g}_{pq}^{rs})^\chi$ and $(S_p^q)^\chi)$ that arise due to the changes in the atomic orbitals as the result of the perturbation. In Eq. (2.31) we additionally define the following elements of the relaxed one- and two-particle density matrices:

$$\tilde{\gamma}_i^a = z_i^a, \quad (2.32)$$

$$\Gamma_{ij}^{ak} = \frac{1}{2} P_-(ij) z_i^a \gamma_j^k, \quad (2.33)$$

$$\Gamma_{ib}^{ac} = \frac{1}{2} P_-(ac) z_i^a \gamma_b^c. \quad (2.34)$$

The coefficients ω_p^q that appear in the final term of Eq. (2.31) are elements of the symmetric energy-weighted density matrix:

$$\omega_i^a = \omega_a^i = -\frac{1}{2} (X_i^a + X_a^i) - \frac{1}{4} z_i^a (\varepsilon_i + \varepsilon_a) - \frac{1}{4} z_k^c [(\bar{g}_{cj}^{ka} + \bar{g}_{ca}^{kj}) \gamma_i^j + (\bar{g}_{cb}^{ki} + \bar{g}_{ci}^{kb}) \gamma_a^b], \quad (2.35)$$

$$\omega_i^j = -\frac{1}{2} (X_i^j + X_j^i) - \frac{1}{4} z_k^c [(\bar{g}_{cl}^{ki} + \bar{g}_{ci}^{kl}) \gamma_j^l + (\bar{g}_{cl}^{kj} + \bar{g}_{cj}^{kl}) \gamma_i^l], \quad (2.36)$$

$$\omega_a^b = -\frac{1}{2} (X_a^b + X_b^a) - \frac{1}{4} z_k^c [(\bar{g}_{cd}^{ka} + \bar{g}_{ca}^{kd}) \gamma_b^d + (\bar{g}_{cd}^{kb} + \bar{g}_{cb}^{kd}) \gamma_a^d]. \quad (2.37)$$

The analytic gradient (2.31) is evaluated by backtransforming the density matrices to the atomic orbital (AO) basis, to avoid costly transformations of the AO basis derivative integrals

for each perturbation. The DCFT-06 analytic gradients have been implemented in the PSI4¹³¹ program, and benchmarked as described in Section 2.4.

2.4 COMPUTATIONAL METHODS

Results obtained using DCFT-06 were tested against coupled cluster theory with single and double excitations (CCSD)^{13,22,23} and CCSD incorporating the perturbative treatment of triple excitations [CCSD(T)],^{132,133} using the large cc-pCVQZ core-valence basis set of Dunning.^{134,135} All electrons were correlated in all three methods. Tight convergence criteria were enforced for the energy (10^{-10} E_h) and the root mean square of the energy gradient (10^{-6} E_h/a₀). DCFT-06 harmonic vibrational frequencies were computed by numerical differentiation of analytic energy gradients.

2.5 RESULTS

Energies, equilibrium structures and harmonic vibrational frequencies were obtained at the DCFT-06 level of theory for a test set of 15 molecules: H₂, HF, H₂O, HOF, H₂O₂, HCN, HNC, NH₃, C₂H₂, C₂H₄, CH₄, N₂, CO, CO₂, CH₂O. Here, we evaluate the quality of DCFT-06 results by comparing them with CCSD, CCSD(T) and accurate empirical equilibrium structures that were derived from experimental rotational constants corrected by theoretical vibration-rotation interaction constants.¹³⁶ For H₂ and H₂O₂ the experimental equilibrium structures were used.^{137,138}

2.5.1 EQUILIBRIUM STRUCTURES

Tables 2.1 and 2.2 show the deviation in bond lengths and bond angles, relative to the empirical structures, as well as the values of the empirical geometric parameters. Not surprisingly, CCSD(T) shows the smallest mean absolute deviations (MADs) Δ_{abs} from the empirical r_e

Table 2.1: Errors in the equilibrium bond lengths (Δr_e , Å) optimized using the DCFT-06, CCSD and CCSD(T) methods with the cc-pCVQZ basis set ($\Delta r_e = r_e^{comp} - r_e^{emp}$). Results are compared to the empirical structural parameters¹³⁶ shown in the right column (r_e^{emp} , Å).

Molecule	Bond	DCFT-06	Δr_e , Å		r_e^{emp} , Å
			CCSD	CCSD(T)	Empirical
H ₂	H-H	0.0019	0.0007	0.0007	0.7411 ^a
HF	F-H	-0.0025	-0.0038	-0.0011	0.9169
H ₂ O	O-H	-0.0011	-0.0031	-0.0004	0.9575
HO ₂ F	O-H	-0.0033	-0.0056	-0.0021	0.9678
H ₂ O ₂	O-H	-0.0061	-0.0085	-0.0051	0.9670 ^a
HNC	N-H	0.0002	-0.0017	0.0010	0.9943
NH ₃	N-H	-0.0005	-0.0029	-0.0004	1.0116
C ₂ H ₂	C-H	0.0012	-0.0010	0.0008	1.0613
HCN	C-H	0.0003	-0.0019	0.0002	1.0653
C ₂ H ₄	C-H	0.0005	-0.0016	0.0002	1.0807
CH ₄	C-H	0.0008	-0.0010	0.0005	1.0859
CH ₂ O	C-H	-0.0073	-0.0084	0.0010	1.1070
N ₂	N-N	-0.0014	-0.0069	0.0004	1.0977
CO	C-O	-0.0056	-0.0067	0.0005	1.1284
CO ₂	C-O	-0.0062	-0.0070	0.0003	1.1601
CH ₂ O	C-O	-0.0037	-0.0074	-0.0004	1.2047
HCN	C-N	-0.0012	-0.0070	0.0004	1.1534
HNC	C-N	-0.0038	-0.0066	0.0006	1.1687
C ₂ H ₂	C-C	-0.0009	-0.0067	0.0000	1.2037
C ₂ H ₄	C-C	0.0033	-0.0057	0.0005	1.3307
HO ₂ F	O-F	0.0184	-0.0221	-0.0018	1.4344
H ₂ O ₂	O-O	0.0092	-0.0245	-0.0059	1.4556 ^a

^aThe H₂ and H₂O₂ bond distances were obtained directly from the experimentally derived equilibrium structures.^{137,138}

Table 2.2: Errors in the equilibrium bond angles ($\Delta\alpha_e$, $^\circ$) optimized using the DCFT-06, CCSD and CCSD(T) methods with the cc-pCVQZ basis set ($\Delta\alpha_e = \alpha_e^{comp} - \alpha_e^{emp}$). Results are compared to the empirical structural parameters¹³⁶ shown in the right column (α_e^{emp} , $^\circ$).

Molecule	Bond angle	DCFT-06	$\Delta\alpha_e$, $^\circ$		α_e^{emp} , $^\circ$
			CCSD	CCSD(T)	Empirical
H ₂ O	H–O–H	−0.15	0.01	−0.29	104.51
HOF	H–O–F	−0.07	0.69	−0.09	97.94
H ₂ O ₂	H–O–O	−2.81	−1.58	−2.33	102.32 ^a
NH ₃	H–N–H	−0.86	−0.61	−0.89	107.25
C ₂ H ₄	H–C–H	−0.13	−0.16	−0.04	117.14
CH ₂ O	H–C–H	−0.02	−0.36	−0.30	116.74

^aThe H₂O₂ bond angle was obtained directly from the experimentally derived equilibrium structure.¹³⁸

bond distances, with a value of $\Delta_{abs} = 0.0011\text{\AA}$. The MAD for DCFT-06 ($\Delta_{abs} = 0.0036\text{\AA}$) represents a significant improvement over CCSD ($\Delta_{abs} = 0.0064\text{\AA}$) and is close to CCSD(T). Only in two cases (for H₂ and C–H bond in C₂H₂) does the error in the bond length optimized with DCFT-06 exceed that of CCSD. In the cases where significant errors in the bond distances were obtained ($>0.001\text{\AA}$), DCFT-06 both over- and underestimations are observed, depending on the nature of the bond. The varying sign of the error is in contrast to CCSD, which consistently underestimates the bond distances for all cases except H₂; basis set truncation is the only significant error source for CCSD in this case. The largest deviations from the empirical structures were obtained for the HOF and H₂O₂ molecules (O–F and O–O bonds, respectively). In both cases DCFT-06 absolute errors (0.0184 $\text{\AA}$ and 0.0092 $\text{\AA}$) are significantly smaller than those of CCSD (0.0221 $\text{\AA}$ and 0.0245 $\text{\AA}$). Excluding HOF and H₂O₂ from the data set results in the reduction of the mean absolute error by $\sim 35\%$ for DCFT-06 and CCSD ($\Delta_{abs} = 0.0024\text{\AA}$ and 0.0044 $\text{\AA}$), and by more than 50% in the case of CCSD(T) ($\Delta_{abs} = 0.0005\text{\AA}$). HOF and H₂O₂ have been challenging for both theory and experiment.^{139–142} The fact that the DCFT-06 and CCSD errors for HOF and H₂O₂ are

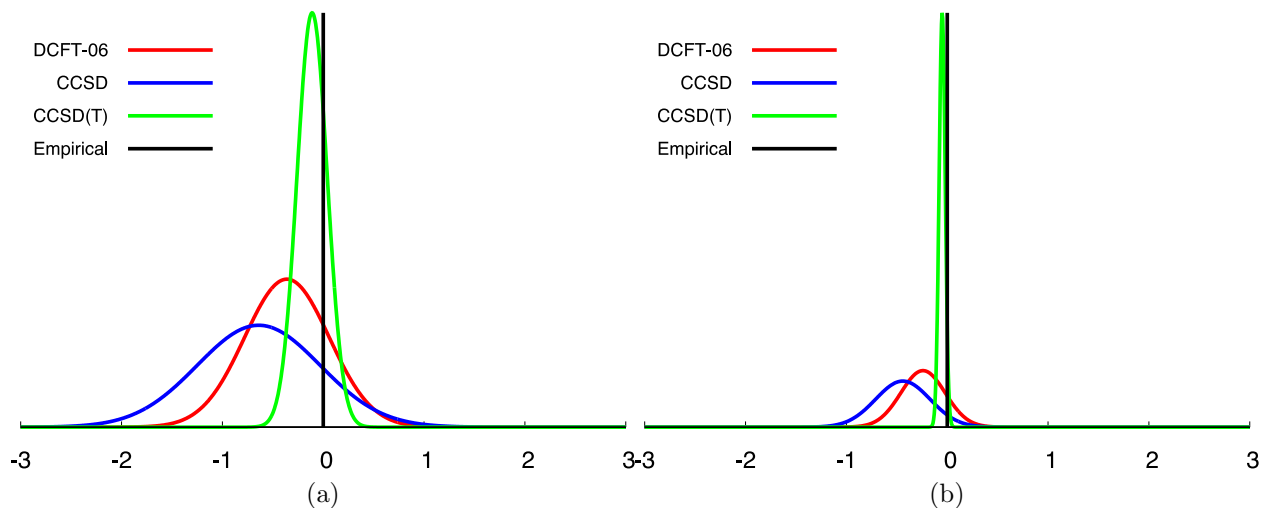


Figure 2.1: Normal distribution of errors in the bond distances (pm) predicted using the DCFT-06, CCSD and CCSD(T) methods, with respect to the empirical structures.¹³⁶ Plot (a) includes HOF and H₂O₂ in the sample of molecules, while (b) excludes them from the data set. All distributions are normalized to one.

significantly reduced at the CCSD(T) level of theory indicates that treatment of connected triple excitations is important for these molecules. The residual errors in the CCSD(T) geometries have previously raised questions about the reliability of the experimental reference structures, prompting calls for a reinvestigation of these geometrical parameters.^{136, 139}

Results presented in Table 2.1 are summarized in Figure 2.1a, which shows the plots of the normal distribution analysis of the errors in computed bond lengths. One can see from the plots that the standard deviation Δ_{std} in the DCFT-06 errors (0.0042 Å) is between that of CCSD(T) (0.0015 Å) and CCSD (0.0061 Å). Exclusion of HOF and H₂O₂ from the sample of molecules gives rise to the sharper normal distributions (Figure 2.1b) reducing the DCFT-06 and CCSD standard deviations by a factor of two ($\Delta_{std} = 0.0022$ Å and 0.0027 Å, respectively) and by a factor of five in the case of CCSD(T) ($\Delta_{std} = 0.0003$ Å).

The empirical equilibrium bond angles, and deviations for each of the theories from these, are shown in Table 2.2. For this limited data set, the DCFT-06 results are consistent with

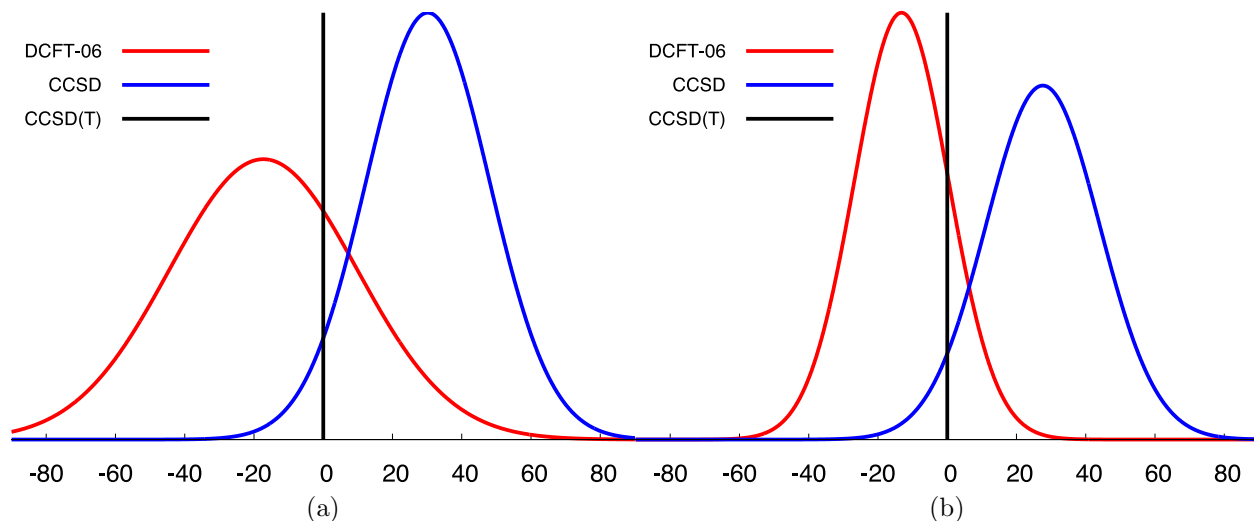


Figure 2.2: Normal distribution of errors in the harmonic vibrational frequencies (cm^{-1}) predicted using the DCFT-06 and CCSD methods, with respect to the frequencies obtained from CCSD(T). Plot (a) includes HOF and H_2O_2 in the sample of molecules, while (b) excludes them from the data set. All distributions are normalized to one.

those from CCSD and CCSD(T), and generally closer to the latter. Only in the case of H_2O_2 is the error obtained using all three methods with respect to experiment greater than 1° , which may partly be attributed to the aforementioned uncertainty in the reference structures.

2.5.2 HARMONIC VIBRATIONAL FREQUENCIES

Harmonic vibrational frequencies are presented in the supplementary material (Sec. 2.8) and summarized in Figure 2.2a. The results obtained using DCFT-06 and CCSD are compared to CCSD(T). For both methods the largest errors in computed frequencies were obtained for the HOF and H_2O_2 molecules (O–F and O–O stretching modes, respectively). In these cases, the DCFT-06 errors are significantly larger than those from CCSD; the errors for (DCFT-06, CCSD) are (177.1 cm^{-1} , 77.3 cm^{-1}) for HOF and (122.7 cm^{-1} , 67.9 cm^{-1}) for H_2O_2 , respectively, for these vibrational modes. The lack of connected triples appears to affect the

performance of both these methods; however CCSD outperforms DCFT-06, possibly due to the inclusion of disconnected higher-order excitations. For all other vibrational modes, the DCFT-06 errors are smaller or on the same order of magnitude as the CCSD errors. The MADs, Δ_{abs} , for (DCFT-06, CCSD) are just (17.3, 30.2) cm^{-1} , with respect to CCSD(T). For all molecules the values of DCFT-06 harmonic frequencies are notably lower than the corresponding CCSD values.

Figure 2.2a shows the results of the normal distribution analysis of the errors in DCFT-06 and CCSD frequencies. While the maximum of the DCFT-06 error distribution is closer to zero than in the case of CCSD, the standard deviation of DCFT-06 errors ($\Delta_{std} = 27.1 \text{ cm}^{-1}$) is larger than that of CCSD ($\Delta_{std} = 17.8 \text{ cm}^{-1}$). This is due to the already mentioned large error in the harmonic frequencies for HOF and H_2O_2 that greatly affects the statistical results in our limited test set of molecules. When HOF and H_2O_2 are excluded from the data set, the DCFT-06 normal distribution becomes much sharper (Figure 2.2b) with a Δ_{abs} of 13.2 cm^{-1} and Δ_{std} of 13.6 cm^{-1} . The distribution of CCSD errors is affected to a lesser extent ($\Delta_{abs} = 27.6 \text{ cm}^{-1}$ and $\Delta_{std} = 16.4 \text{ cm}^{-1}$) by neglecting these molecules.

2.6 CONCLUSIONS

We have presented expressions for analytic gradients for the DCFT-06 variant of density cumulant functional theory. Analogous to the evaluation of the DCFT-06 energy, the gradient requires the solution of coupled equations, for both orbital and cumulant response. Although the energy and gradient computations have the same asymptotic scaling, our experience shows that the computational cost of the latter is significantly lower than the former, as a result of the rapid convergence of the response equations. Thus the use of DCFT-06 gradients yield much additional information about the potential energy hypersurface for modest computational cost.

Our benchmark results indicate that DCFT-06 provides a well balanced description of electron correlation, yielding optimized structural parameters and harmonic vibrational frequencies, ranging in quality between those of CCSD and CCSD(T) theory. In all but two of the cases tested herein, the errors in the DCFT-06 equilibrium bond distances with respect to the accurate empirical structures are smaller than the errors of CCSD. Harmonic frequencies computed at the DCFT-06 level of theory are also in good agreement with those from CCSD(T) for the majority of molecules in the test set.

Density cumulant functional theory has a number of attractive properties. It is formally exact and, even when approximations are introduced, it is both size-consistent and size-extensive. The use of approximate conditions gives rise to a hierarchy of methods, allowing systematic improvements of the theory to be introduced. The computational scaling of the DCFT-06 method is $\mathcal{O}(N^6)$, as is that of CCSD. The former, however, has reduced nonlinearities and a Hermitian energy functional. We are currently working on more extensive benchmarking of the DCFT-06 method, and exploring more accurate variations of the theory that use refined n -representability constraints.

2.7 ACKNOWLEDGMENTS

This research was supported by the U.S. National Science Foundation, Grant CHE-1054286. The authors would like to thank Professor Werner Kutzelnigg for his fundamental contributions in the field of density cumulant functional theory and for many helpful discussions.

2.8 SUPPLEMENTARY MATERIAL

See supplementary material at <http://dx.doi.org/10.1063/1.4739423> for tables with energies, harmonic frequencies and structural parameters obtained using DCFT-06, CCSD, CCSD(T) methods.

CHAPTER 3

DENSITY CUMULANT FUNCTIONAL THEORY: THE DC-12 METHOD, AN IMPROVED DESCRIPTION OF THE ONE-PARTICLE DENSITY MATRIX*

*A. Yu. Sokolov, A. C. Simmonett, and H. F. Schaefer, J. Chem. Phys. **138**, 024107 (2013). Reproduced with permission of the American Institute of Physics.

3.1 ABSTRACT

Density *cumulant* functional theory (DCFT) is a theory that, in principle, can compute energies and properties exactly without a wavefunction. To accomplish this, the energy is expressed as an exact, known functional of the one-particle density matrix and two-particle density cumulant. The correlation contribution to the one-particle density matrix is obtained from the cumulant, to eliminate redundancy in the equations. The previous formulation of DCFT introduced this relationship in an approximate way, to obtain tractable equations. In this research it is demonstrated that the correlation contribution to the one-particle density matrix can be extracted exactly from the cumulant, with minimal computational overhead and no increase in the asymptotic cost of the theory. We present numerical results, showing the improvements resulting from this reformulation (DC-12), and offer a perturbative analysis of the new equations to compare them to their predecessors.

3.2 INTRODUCTION

The last few decades have seen important developments increasing the tractability of the electron correlation problem. Of the many accurate wavefunction-based methods available, coupled cluster theory has proven popular due to its size-extensivity, widespread applicability and inclusion of disconnected products of cluster operators. While the presence of these operator products incorporates higher-order effects, it leads to challenging nonlinear equations, which complicate the theory’s spin adaptation and efficient implementation. Motivated by this, there have been numerous attempts to simplify the coupled cluster approach by linearization,^{109–115,143} while maintaining the excellent performance of the method. The resulting pair correlation methods have been demonstrated to perform very well, but may start to degrade in the presence of increased multireference character.

Another approach to the same problem involves directly obtaining density matrices, circumventing the computation of the wavefunction. In these approaches, it is common to decompose the desired density matrix into its fully connected part – namely, its cumulant^{58,77,85–91,144–146} – and a product of lower-order density matrices. The motivation for doing so is that the cumulant is an extensive quantity, which will ultimately yield a size-extensive theory, as is the case in coupled cluster theory. Cumulants have been employed in many areas of quantum chemistry, including the solution of the contracted Schrödinger equation^{67,86,94,95} and its anti-Hermitian counterpart,^{77,147–150} canonical transformation theory,^{102–104,151–153} parametric reduced density matrix methods,^{64,105–108} multi-reference methods based on generalized normal-ordering,^{58,154–160} and density cumulant functional theory (DCFT).^{96,97}

The DCFT method was first proposed by Kutzelnigg,⁹⁶ and implemented in our group along with its analytic gradients.^{97,161} The theory proceeds by expressing the energy exactly in terms of the one-particle density matrix (OPDM) and the two-particle density cumulant. After constraining the latter to be at least approximately n -representable, *i.e.*, making it correspond to an n -electron wavefunction, the energy can be found by making this functional stationary. The choice of these n -representability constraints is flexible, allowing a hierarchy of methods. In addition to using perturbatively-derived n -representability constraints, one more approximation was made in the original formulation; the correlation contribution to the OPDM was derived from the two-particle cumulant, via a low-order power series expansion, to obtain the correct number of independent variables. In this research we show how the natural spin-orbital (NSO) basis employed in DCFT allows one to avoid such an approximation, investigate its numerical effects, and assess the quality of the currently used n -representability conditions. Hereafter, we refer to the original theory using Kutzelnigg’s approximate OPDM expansion as DC-06, while the modified, more rigorous method is denoted DC-12. Our previous work^{97,161} denoted DC-06 as DCFT-06, but here a more compact

notation is adopted for brevity, and to avoid confusion with Kohn-Sham DFT nomenclature. By comparing the results of DC-06 and DC-12, we can gauge the effect of each of these approximations in the DCFT framework, and take another step towards defining a hierarchy of methods.

3.3 THEORY

3.3.1 OVERVIEW OF DENSITY CUMULANT FUNCTIONAL THEORY

Consider a Fock space composed of an orthonormal set of spin-orbitals $\{\psi_p\}$. Begin by writing the expression of the molecular Hamiltonian $\hat{H}$ in the second-quantized form:⁵⁷

$$\hat{H} = h_p^q a_q^p + \frac{1}{2} g_{pq}^{rs} a_{rs}^{pq} , \quad (3.1)$$

where the excitation operators a_p^q and a_{pq}^{rs} are expressed as products of creation $a_p^\dagger$ and annihilation a_p operators^{13,22} as:

$$a_p^q = a_q^\dagger a_p , \quad a_{pq}^{rs} = a_r^\dagger a_s^\dagger a_q a_p \quad (3.2)$$

and h_p^q and g_{pq}^{rs} are one- and two-electron integrals:

$$h_p^q = \langle \psi_p(1) | \hat{h} | \psi_q(1) \rangle , \quad g_{pq}^{rs} = \langle \psi_p(1) \psi_q(2) | \frac{1}{r_{12}} | \psi_r(1) \psi_s(2) \rangle . \quad (3.3)$$

In Eq. (3.1) the labels $p, q, r, s \dots$ run over all spin-orbitals and the Einstein convention is used for the summation over repeated indices. Taking the expectation value of the Hamiltonian (3.1) with respect to its eigenfunction Ψ , we obtain the expression for the energy:

$$E = \langle \Psi | \hat{H} | \Psi \rangle = h_p^q \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \gamma_{rs}^{pq} , \quad (3.4)$$

where the reduced one- and two-particle density matrices, γ_1 and γ_2 , respectively, are introduced

$$\gamma_p^q = \langle \Psi | a_p^q | \Psi \rangle \quad , \quad \gamma_{pq}^{rs} = \langle \Psi | a_{pq}^{rs} | \Psi \rangle \quad (3.5)$$

along with the antisymmetrized two-electron integrals:

$$\bar{g}_{pq}^{rs} = g_{pq}^{rs} - g_{pq}^{sr} \quad . \quad (3.6)$$

The unknown parameters, γ_p^q and γ_{pq}^{rs} , in Eq. (3.4) are related to each other through a partial trace and, therefore, cannot be independently varied to obtain the stationary value of the energy. In density cumulant functional theory^{96,97} one circumvents this redundancy by introducing the two-particle density cumulant^{58,77,85–91,144–146}

$$\lambda_{pq}^{rs} = \gamma_{pq}^{rs} - \gamma_p^r \gamma_q^s + \gamma_p^s \gamma_q^r \quad (3.7)$$

and decomposing the one-particle density matrix into its best idempotent approximation¹²⁴ κ and a correction⁹⁶ τ :

$$\gamma_p^q = \kappa_p^q + \tau_p^q \quad , \quad \kappa_p^r \kappa_r^q = \kappa_p^q \quad . \quad (3.8)$$

The dependence of the non-idempotent correction τ on the density cumulant can be seen from the partial trace relation⁹⁶

$$\sum_r \lambda_{pr}^{qr} = (\gamma_1^2 - \gamma_1)_p^q \equiv d_p^q = \tau_p^r \tau_r^q - \tau_p^q + 2\tau_p^r \kappa_r^q \quad , \quad (3.9)$$

where d_p^q is defined as the deviation of γ_1 from idempotency. Inserting Eqs. (3.7) and (3.8) into Eq. (3.4) and defining the generalized Fock operator matrix

$$f_p^q = h_p^q + \bar{g}_{pr}^{qs} \gamma_s^r \quad (3.10)$$

we obtain the following expression for the energy as an exact functional of $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}_2$:

$$E = h_p^q \gamma_q^p + \frac{1}{2} \bar{g}_{pq}^{rs} \gamma_r^p \gamma_s^q + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq} = \frac{1}{2} (h_p^q + f_p^q) \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq} . \quad (3.11)$$

Variation of the energy functional (3.11) with respect to unitary transformations of $\boldsymbol{\kappa}$ results^{96,125} in the following stationarity condition:

$$[\mathbf{f}, \boldsymbol{\kappa}] = 0 . \quad (3.12)$$

Equation (3.12) can be satisfied by self-consistent diagonalization of the generalized Fock operator matrix $\mathbf{f}$ (Eq. (3.10)). The latter can be conveniently decomposed into two contributions:

$$f_p^q = (f_0)_p^q + (f')_p^q , \quad (f_0)_p^q = h_p^q + \bar{g}_{pr}^{qs} \kappa_s^r , \quad (f')_p^q = \bar{g}_{pr}^{qs} \tau_s^r , \quad (3.13)$$

where $(f_0)_p^q$ represents the Fock operator employed in Hartree-Fock theory and $(f')_p^q$ is a correction that describes the effects of electron correlation on the orbitals. While $\boldsymbol{\kappa}$ is n -representable by construction, one needs to employ a sufficient set of conditions to guarantee the n -representability of the density cumulant upon the variation of the energy with respect to $\boldsymbol{\lambda}_2$.

3.3.2 DC-06 METHOD

The simplest method of density cumulant functional theory, denoted as DC-06, arises^{96,97} from the use of approximate n -representability conditions derived from second-order perturbation theory.⁹² Inserting these conditions into Eq. (3.11), we obtain an expression for the DC-06 energy:

$$E = \frac{1}{2} (h_p^q + f_p^q) \gamma_q^p + \frac{1}{2} \bar{g}_{ij}^{ab} \lambda_{ab}^{ij} + \frac{1}{8} \lambda_{kl}^{ab} \bar{g}_{ij}^{kl} \lambda_{ab}^{ij} + \frac{1}{8} \lambda_{ij}^{cd} \bar{g}_{cd}^{ab} \lambda_{ab}^{ij} - \lambda_{ik}^{ac} \bar{g}_{jc}^{kb} \lambda_{ab}^{ij} . \quad (3.14)$$

As usual, indices $i, j, k \dots$ and $a, b, c \dots$ run over occupied and virtual spin-orbitals, respectively.

In DC-06 the non-idempotent component, $\boldsymbol{\tau}$, of the one-particle density matrix is obtained⁹⁶ from the cumulant $\boldsymbol{\lambda}_2$ by approximately solving the quadratic equation (3.9) in the natural spin-orbital basis (NSO), where $\boldsymbol{\tau}$ is diagonal. Discriminating between occupied (i') and virtual (a') NSOs (primes denote the use of the NSO basis) with elements $\kappa_{i'} = 1$ and $\kappa_{a'} = 0$, respectively, we obtain:

$$\tau_{i'}^2 + \tau_{i'} - d_{i'} = 0 , \quad (3.15)$$

$$\tau_{a'}^2 - \tau_{a'} - d_{a'} = 0 . \quad (3.16)$$

Expanding equations (3.15) and (3.16) in a Taylor series

$$\tau_{i'} = -\frac{1}{2} \left(1 - \sqrt{1 + 4d_{i'}} \right) = d_{i'} - d_{i'}^2 + 2d_{i'}^3 - \dots , \quad (3.17)$$

$$\tau_{a'} = \frac{1}{2} \left(1 - \sqrt{1 + 4d_{a'}} \right) = -d_{a'} + d_{a'}^2 - 2d_{a'}^3 + \dots , \quad (3.18)$$

the elements of $\boldsymbol{\tau}$ are approximated⁹⁶ by truncating (3.17) and (3.18) at the leading term:

$$\tau_i^j \approx d_i^j = -\frac{1}{2} \lambda_{ik}^{ab} \lambda_{ab}^{jk} , \quad (3.19)$$

$$\tau_a^b \approx -d_a^b = \frac{1}{2} \lambda_{ac}^{ij} \lambda_{ij}^{bc} , \quad (3.20)$$

where expressions for d_p^q are obtained exactly from n -representability conditions.^{92,96}

Using the approximate relations (3.19) and (3.20) between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$, the stationarity conditions of the DC-06 energy (3.14), with respect to variation of $\boldsymbol{\lambda}_2$, can be written as:^{97, 161}

$$\begin{aligned} \frac{\partial E}{\partial \lambda_{ij}^{ab}} = & P_{-}(ab) [f_a^c \lambda_{cb}^{ij}] - P_{-}(ij) [f_k^i \lambda_{ab}^{kj}] + \bar{g}_{ab}^{ij} \\ & + \frac{1}{2} \bar{g}_{kl}^{ij} \lambda_{ab}^{kl} + \frac{1}{2} \bar{g}_{ab}^{cd} \lambda_{cd}^{ij} - P_{-}(ij) P_{-}(ab) [\bar{g}_{kb}^{jc} \lambda_{ac}^{ik}] = 0 , \end{aligned} \quad (3.21)$$

where we introduce the antisymmetric permutation operator $P_{-}(ij)f(i, j) \equiv f(i, j) - f(j, i)$ that acts only on the terms in the square brackets. The residual equations (3.21) are non-linear in $\boldsymbol{\lambda}_2$, as the first two terms of Eq. (3.21) implicitly depend on $\boldsymbol{\tau}$, which is quadratic in the cumulant. In order to obtain the stationary value of the DC-06 energy, conditions (3.12) and (3.21) must be satisfied simultaneously. This can be achieved by the iterative solution of Eq. (3.21), followed by the self-consistent update of the $\boldsymbol{f}$ matrix and orbital transformation with every macroiteration. Interestingly, setting $\boldsymbol{\tau}$ to zero in expressions (3.12) and (3.21), one obtains the system of equations that define coupled electron pair approximation zero (CEPA₀).^{162, 163} The same system of equations can be obtained in many other ways; for example, by linearization of coupled cluster method with only double excitations (LCCD).^{21, 109} Thus, the DC-06 method can be viewed as a natural extension of LCCD, for which additional descriptions of electron correlation are obtained via the non-idempotency of the one-particle density matrix, and use of the generalized Fock operator that introduces orbital relaxation from the Hartree-Fock starting orbitals.

3.3.3 IMPROVED DESCRIPTION OF THE OPDM IN THE DC-12 MODEL

Let us take another look at Eq. (3.9) defining the exact relationship between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$. By considering the occupied and virtual terms separately and recalling that the $\boldsymbol{\kappa}$ matrix is diagonal with elements $\kappa_i^i = 1$ and $\kappa_a^a = 0$, we obtain two systems of coupled quadratic

equations:

$$\tau_i^j = d_i^j - \tau_i^k \tau_k^j , \quad (3.22)$$

$$\tau_a^b = -d_a^b + \tau_a^c \tau_c^b . \quad (3.23)$$

Comparing the above equations with expressions (3.19) and (3.20), one can see that approximations (3.19) and (3.20), introduced in the DC-06 method, are consistent with ignoring the coupling term $\tau_p^r \tau_r^q$ in (3.22) and (3.23). Let us now attempt to solve the non-linear Eqs. (3.22) and (3.23) exactly. While simple analytic solutions (3.17) and (3.18) can be obtained in the NSO basis, it is possible to solve Eqs. (3.22) and (3.23) iteratively using approximate Eqs. (3.19) and (3.20) as the starting guess. The iterative solution is efficient ($\mathcal{O}(N^3)$) and is expected to converge quickly, as the coupling term should be a small correction.

Making the energy (3.14) stationary with respect to variation of $\boldsymbol{\lambda}_2$, we obtain for the first term:

$$\frac{\partial}{\partial \lambda_{ij}^{ab}} \left(\frac{1}{2} (h_p^q + f_p^q) \gamma_q^p \right) = f_p^q \frac{\partial \tau_q^p}{\partial \lambda_{ij}^{ab}} = f_k^l \frac{\partial \tau_l^k}{\partial \lambda_{ij}^{ab}} + f_c^d \frac{\partial \tau_d^c}{\partial \lambda_{ij}^{ab}} , \quad (3.24)$$

where the derivatives of $\boldsymbol{\tau}$ with respect to $\boldsymbol{\lambda}_2$ are unknown and must be evaluated. Differentiating Eqs. (3.22) and (3.23), we get:

$$\frac{\partial \tau_k^l}{\partial \lambda_{ij}^{ab}} = \frac{\partial d_k^l}{\partial \lambda_{ij}^{ab}} - \frac{\partial \tau_k^m}{\partial \lambda_{ij}^{ab}} \tau_m^l - \tau_k^m \frac{\partial \tau_m^l}{\partial \lambda_{ij}^{ab}} , \quad (3.25)$$

$$\frac{\partial \tau_c^d}{\partial \lambda_{ij}^{ab}} = -\frac{\partial d_c^d}{\partial \lambda_{ij}^{ab}} + \frac{\partial \tau_c^e}{\partial \lambda_{ij}^{ab}} \tau_e^d + \tau_c^e \frac{\partial \tau_e^d}{\partial \lambda_{ij}^{ab}} . \quad (3.26)$$

Unfortunately, the coupled linear equations (3.25) and (3.26) do not have simple analytic solutions and their iterative solution is rather expensive with prohibitive $\mathcal{O}(N^7)$ scaling. We will, therefore, attempt to solve these equations in the NSO basis. Performing an eigendecomposition of $\boldsymbol{\tau}$ yields

$$\mathbf{U}^\dagger \boldsymbol{\tau} \mathbf{U} = \boldsymbol{\tau}_{NSO} , \quad (3.27)$$

where matrices $\boldsymbol{\tau}_{NSO}$ and $\mathbf{U}$ contain eigenvalues and eigenvectors of $\boldsymbol{\tau}$, respectively. Because the elements of $\boldsymbol{\tau}$ in the occupied-virtual block are zero, matrix $\mathbf{U}$, providing the transformation to the NSO basis, only mixes the orbitals within the occupied-occupied and virtual-virtual blocks, which leaves the energy unchanged. Evaluating the solution of Eqs. (3.25) and (3.26) in the NSO basis, we find

$$\frac{\partial \tau_{k'}^{l'}}{\partial \lambda_{i'j'}^{a'b'}} = \frac{1}{1 + \tau_{k'} + \tau_{l'}} \frac{\partial d_{k'}^{l'}}{\partial \lambda_{i'j'}^{a'b'}} , \quad (3.28)$$

$$\frac{\partial \tau_{c'}^{d'}}{\partial \lambda_{i'j'}^{a'b'}} = -\frac{1}{1 - \tau_{c'} - \tau_{d'}} \frac{\partial d_{c'}^{d'}}{\partial \lambda_{i'j'}^{a'b'}} , \quad (3.29)$$

where the derivatives of the deviation from idempotency $\mathbf{d}$ (3.19) and (3.20) can be written as:

$$\frac{\partial d_{k'}^{l'}}{\partial \lambda_{i'j'}^{a'b'}} = -\frac{1}{4} \lambda_{l'j'}^{a'b'} \delta_{k'}^{i'} + \frac{1}{4} \lambda_{l'i'}^{a'b'} \delta_{k'}^{j'} - \frac{1}{4} \lambda_{k'j'}^{a'b'} \delta_{l'}^{i'} + \frac{1}{4} \lambda_{k'i'}^{a'b'} \delta_{l'}^{j'} , \quad (3.30)$$

$$\frac{\partial d_{c'}^{d'}}{\partial \lambda_{i'j'}^{a'b'}} = -\frac{1}{4} \lambda_{i'j'}^{d'b'} \delta_{a'}^{c'} + \frac{1}{4} \lambda_{i'a'}^{d'b'} \delta_{b'}^{c'} - \frac{1}{4} \lambda_{c'j'}^{d'b'} \delta_{a'}^{d'} + \frac{1}{4} \lambda_{i'j'}^{c'a'} \delta_{b'}^{d'} . \quad (3.31)$$

As in the previous section, the NSO basis indices in Eqs. (3.28) – (3.31) are labeled by the primes.

We refer the interested reader to the appendix for the derivation, but we are now in a position to define the intermediate $\tilde{F}_p^q$

$$\tilde{F}_i^j \equiv \frac{f_{i'}^{j'} U_{i'}^{i'} (U^\dagger)_{j'}^j}{1 + \tau_{i'} + \tau_{j'}} , \quad (3.32)$$

$$\tilde{F}_a^b \equiv \frac{f_{a'}^{b'} U_{a'}^{a'} (U^\dagger)_{b'}^b}{1 - \tau_{a'} - \tau_{b'}} , \quad (3.33)$$

to obtain a simple expression for the energy stationarity condition, which can be solved exactly via an implicit transformation to the NSO basis:

$$\begin{aligned} \frac{\partial E}{\partial \lambda_{ij}^{ab}} = & P_{-}(ab) \left[\tilde{F}_a^c \lambda_{cb}^{ij} \right] - P_{-}(ij) \left[\tilde{F}_k^i \lambda_{ab}^{kj} \right] + \bar{g}_{ab}^{ij} \\ & + \frac{1}{2} \bar{g}_{kl}^{ij} \lambda_{ab}^{kl} + \frac{1}{2} \bar{g}_{ab}^{cd} \lambda_{cd}^{ij} - P_{-}(ij) P_{-}(ab) \left[\bar{g}_{kb}^{jc} \lambda_{ac}^{ik} \right] = 0 . \end{aligned} \quad (3.34)$$

The non-linear system of equations (3.34) defines a new variation of the DC-06 method that utilizes the exact relationship between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$, according to Eqs. (3.22) and (3.23). We name the new method DC-12, as it uses the approximate n -representability conditions proposed by Kutzelnigg in 2006, together with modified terms describing the orbital relaxation and the non-idempotency of one-particle density matrix $\boldsymbol{\gamma}_1$ derived in the present research. Most importantly, both the DC-06 and DC-12 methods have the same $\mathcal{O}(N^6)$ asymptotic scaling, and, therefore, similar computational cost. The only additional work required for the solution of Eq. (3.34) in DC-12 method is the iterative solution of the Eqs. (3.22) and (3.23), diagonalization of the $\boldsymbol{\tau}$ matrix and the formation of the $\tilde{F}_p^q$ intermediates according to Eqs. (3.32) and (3.33), which have overall $\mathcal{O}(N^3)$ scaling. Moreover, our experience shows that the DC-12 residual equations (3.12) and (3.34) exhibit better convergence and numerical stability than that of DC-06, giving rise to significant computational savings. The residual equations (3.12) and (3.34) can be solved simultaneously or using a macro-iteration approach. The former often leads to fewer evaluations of the terms required to solve (3.34), and benefits from the use of partial atomic-orbital based algorithms,¹³⁰ to avoid costly full integral transformations. The two-step approach can be solved entirely in the molecular orbital basis, because only a few macro-iterations and, therefore, integral transformations are typically required. We are developing improved coupling schemes between the orbital and cumulant updates to optimize the convergence rates and to determine the best general approach to solving the equations.

Let us estimate the effect of the approximation in Eqs. (3.19) and (3.20) on the energy (3.14) by performing a perturbative analysis of the DC-06 and DC-12 residual Eqs. (3.21) and (3.34). We denote μ as a perturbation parameter, for which the limit of the wavefunction $\mu \rightarrow 0$ corresponds to a single Slater determinant. As shown by Kutzelnigg and Mukherjee,⁹² the only non-vanishing elements of $\boldsymbol{\lambda}_2$ to $\mathcal{O}(\mu)$ are $\lambda_{ij}^{ab} = (\lambda_{ab}^{ij})^*$. Thus, the approximate relationship between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$ in Eqs. (3.19) and (3.20) is correct to $\mathcal{O}(\mu^2)$. Taking advantage of the exact expressions (3.22) and (3.23) only corrects $\boldsymbol{\tau}$ at $\mathcal{O}(\mu^4)$. Now take a closer look at the first two terms of the DC-12 residual (3.34). Expanding in a power series, we obtain:

$$\begin{aligned} P_-(ab) \left[\tilde{F}_a^c \lambda_{cb}^{ij} \right] - P_-(ij) \left[\tilde{F}_k^i \lambda_{ab}^{kj} \right] &= P_-(ab) \left[f_a^c \lambda_{cb}^{ij} \right] - P_-(ij) \left[f_k^i \lambda_{ab}^{kj} \right] \\ &+ P_-(ab) \left[f_{a'}^{c'} (\tau_{a'} + \tau_{c'}) U_a^{a'} (U^\dagger)_{c'}^c \lambda_{cb}^{ij} \right] + P_-(ij) \left[f_{k'}^{i'} (\tau_{i'} + \tau_{k'}) U_k^{k'} (U^\dagger)_{i'}^i \lambda_{ab}^{kj} \right] + \mathcal{O}(\mu^5) . \end{aligned} \quad (3.35)$$

The first two terms of (3.35) correspond to the two leading contributions of the DC-06 residual Eq. (3.21). The appearance of the higher-order non-linear terms can already be noticed at $\mathcal{O}(\mu^3)$. The third-order terms of the expansion (3.35) are expected to be significant, as they are dominated by the product of the large diagonal elements of the Fock matrix f_p^p and eigenvalues of $\boldsymbol{\tau}$. We therefore conclude that the coupling terms $\tau_p^r \tau_r^q$ in Eqs. (3.22) and (3.23), neglected in DC-06, mainly affect the DC-12 energy (3.14) in two ways: (i) by correcting $\boldsymbol{\tau}$ at $\mathcal{O}(\mu^4)$ as a contribution to $\boldsymbol{\gamma}_1$ and $\boldsymbol{f}$; (ii) by entering the residual equations (3.34) and affecting $\boldsymbol{\lambda}_2$ at $\mathcal{O}(\mu^3)$. While perturbation theory suggests the latter effect to be much larger, our numerical tests indicate that the two are comparable in magnitude, with the second effect slightly dominating. This is the result of the fact that the one-electron term $\frac{1}{2}(h_p^q + f_p^q)\gamma_q^p$ of expression (3.14), containing the $\boldsymbol{\tau}$ -dependent terms, provides the dominant contribution to the energy. Thus, even a fourth-order correction of $\boldsymbol{\tau}$ affects the energy

significantly. Importantly, avoiding approximations (3.19) and (3.20) in DC-12 allows us to take full advantage of the current n -representability conditions and assess their quality, as these are the only approximation used in this method. In 2006 paper⁹⁶ Kutzelnigg mentioned that the density cumulant elements of the type λ_{ij}^{ka} and λ_{ic}^{ab} also appear at $\mathcal{O}(\mu^3)$. While we expect the magnitude of these cumulant elements to be small, and, therefore, less important than the non-linear terms in Eq. (3.35), improved n -representability conditions from higher-order perturbation theory are needed to assess their significance. This will be studied in our future work. In Section 3.5 we will further demonstrate the importance of exact treatment of the relationship between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$ by comparing the results of DC-06 and DC-12 methods.

3.4 COMPUTATIONAL METHODS

The DC-12 method was implemented in a developer’s version of the PSI4¹³¹ program. The results were benchmarked against coupled cluster theory with single and double excitations (CCSD),^{13,22,23} CCSD incorporating the perturbative treatment of triple excitations [CCSD(T)],^{132,133} and linearized coupled-cluster doubles method (LCCD).^{21,109} For the dissociation of H_2 benchmark study (Sec. 3.5.1) the orbital-optimized second- and third-order Møller-Plesset perturbation theories (OMP2 and OMP3)^{164,165} as well as orbital-optimized LCCD (OLCCD)¹⁶⁴ of Bozkaya and co-workers were used. All electrons were correlated. For all computations the correlation-consistent core-valence cc-pCVQZ basis sets of Dunning^{134,135} were used, unless stated otherwise. DC-12 optimized geometries and harmonic vibrational frequencies were obtained by numerical differentiation of the energy using the five-point formula. Tight convergence criteria were enforced for the energy (10^{-10} E_h) and the root mean square of the energy gradient ($10^{-6} \text{ E}_h/\text{a}_0$).

3.5 RESULTS

Here we present preliminary benchmark results for our new DC-12 method. The purpose of our benchmark study is three-fold. By comparing the results of the DC-06 and DC-12 methods with the LCCD method, we aim to assess the significance of the treatment of the non-idempotency of the one-particle density matrix and orbital relaxation. Testing DC-12 against DC-06 will reveal the effect of the exact relationship between the non-idempotency of the OPDM and the density cumulant. Finally, the benchmark of DC-12 will allow us to evaluate the quality of the approximate n -representability conditions used in this variation of DCFT; these are the only remaining approximations in the current formulation. Three types of molecular system will be used for the benchmark. We will first test the performance of the DC-12 method in the treatment of electron correlation in the two-electron systems (Sec. 3.5.1). Next, we will assess its quality in describing the covalent bonding in diatomic molecules with single and multiple bonds (Sec. 3.5.2). Finally, we will benchmark DC-12 description of non-bonded interactions in the argon dimer and the linear hydrogen fluoride dimer (Sec. 3.5.3).

3.5.1 TWO-ELECTRON SYSTEMS

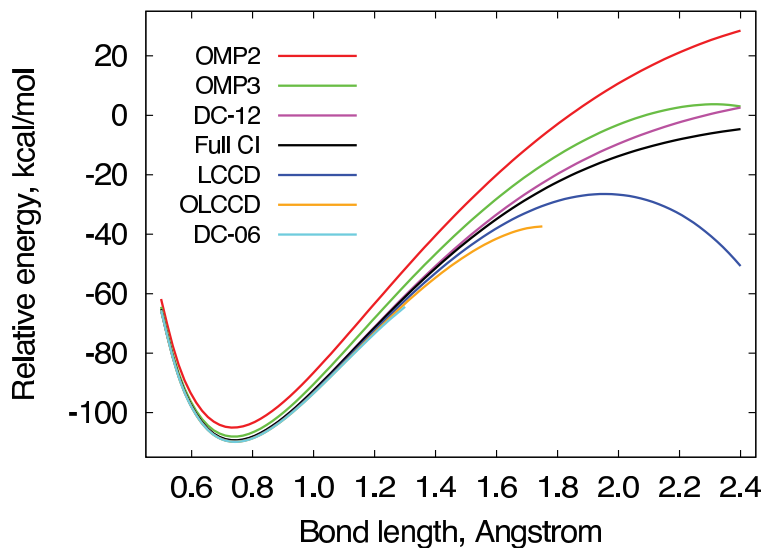
Among the most desirable properties of the electronic structure methods are size-extensivity and size-consistency, orbital invariance, and exactness for two-electron systems. While the DC-06 and DC-12 methods exhibit the first three of these properties, they are not exact for two-electron systems. In this section we test the performance of DCFT in describing the ground state properties of the He atom and H₂ molecule. Table 3.1 shows the errors in the total energies and vertical ionization potentials (IPs) of He and equilibrium H₂ computed using the LCCD, DC-06 and DC-12 methods, with respect to full CI. One can see that all three methods reproduce the exact energy with the absolute error of at most $3 \times 10^{-4} E_h$

Table 3.1: Errors in total energies (E , E_h) and vertical ionization potentials (IP, eV) of the He atom and optimized H_2 molecule computed using LCCD, DC-06, and DC-12 methods with respect to those computed with full CI. For H_2 errors in equilibrium distances (r_e , Å) and harmonic vibrational frequencies (ω_e , cm^{-1}) are also presented. For all computations, the cc-pV5Z basis set was used.

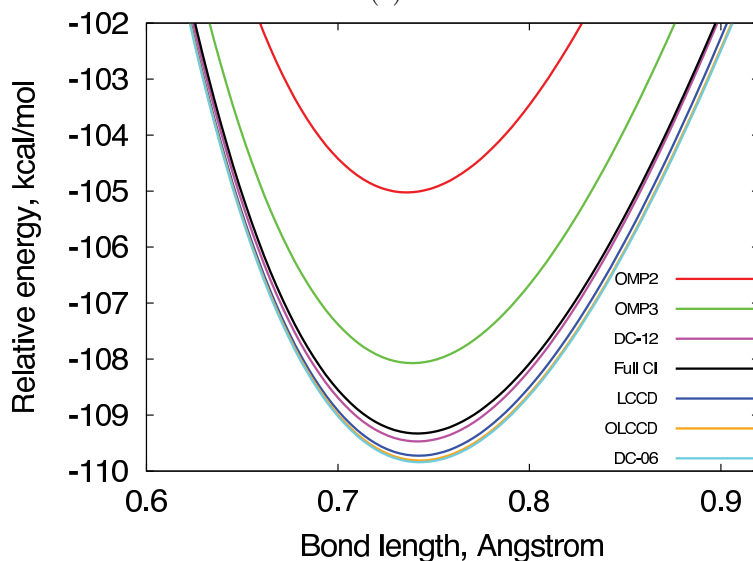
	Full CI	LCCD	DC-06	DC-12
He				
E	-2.903152	-0.000314	-0.000375	-0.000180
IP	24.578	0.009	0.010	0.005
H_2				
E	-1.174223	-0.000617	-0.000813	-0.000211
IP	16.438	0.012	0.013	0.007
r_e	0.7415	0.0007	0.0012	-0.0002
ω_e	4404.9	-13.0	-26.3	7.2

in the case of He and $8 \times 10^{-4} E_h$ in case of H_2 , which results in the error of ca. 0.01 eV in the IPs. The smallest error in the total energy ($2 \times 10^{-4} E_h$) was obtained using the DC-12 method, which indicates about a factor of two improvement over the other methods. This improvement can also be seen when comparing the equilibrium bond distances and harmonic vibrational frequencies of H_2 (r_e and ω_e , Table 3.1). In this case DC-12 again shows the best performance, with errors of only 2×10^{-4} Å and 7 cm^{-1} , respectively, compared to full CI.

Another desirable feature for a theory is consistent performance across a potential energy surface. In Figure 3.1, we test this by stretching H_2 . In addition to the three methods considered above, we also employed the OMP2, OMP3 and OLCCD methods, which are not exact for two-electron systems, but include the effect of the orbital optimization. As one can see from Figure 3.1b, in the equilibrium region (between 0.6 and 0.9 Å) DC-12 performs the best, having only a $\sim 0.15 \text{ kcal mol}^{-1}$ error, with respect to full CI. The LCCD, OLCCD, and DC-06 methods perform roughly alike, giving 0.6-0.7 kcal mol^{-1} errors. The OMP2 and OMP3 curves exhibit significant errors near equilibrium, lying above the full CI curve. Let



(a)



(b)

Figure 3.1: Potential energy curves (PECs) for the dissociation of H_2 computed using OMP2, OMP3, LCCD, OLCCD, DC-06, DC-12 and full CI methods with the cc-pV5Z basis set. Plot (a) shows the full PEC in the range of H–H bond lengths between 0.5 Å and 2.4 Å, while plot (b) shows the enlarged version of the PEC in the equilibrium region. Energies are relative to the full CI dissociation limit $E_\infty(\text{FCI}) = -0.999989 E_h$. For DC-06 and OLCCD no convergence is obtained for bond lengths longer than 1.30 Å and 1.75 Å, respectively.

us now compare the performance of all methods in the region between 1.2 and 2.4 Å (Figure 3.1a). Among the six methods tested, the curve computed with DC-12 exhibits the best agreement with the full CI curve, only showing noticeable deviation at bond lengths greater than 1.6 Å, which results in ca. 7 kcal mol⁻¹ error in the binding energy at a bond length of 2.4 Å. In contrast, the LCCD and OLCCD curves start to deviate already at 1.3 Å, leading to a very large ~ 40 kcal mol⁻¹ error at 2.4 Å for LCCD. In the case of DC-06 method the deviation can be noticed at 1.25 Å. For DC-06, no converged solution was obtained for bond distances longer than 1.3 Å, but we expect the curve to deviate even further in this region. The OMP2 and OMP3 methods underestimate the binding energy across the entire region of the dissociation curve, with OMP3 showing a significant improvement over OMP2.

Overall, the present results indicate that the improved treatment of the non-idempotency of the one-particle density matrix in the DC-12 method gives rise to significantly improved results for the two-electron systems studied herein. While in the equilibrium region of H₂ LCCD, OLCCD, DC-06 and DC-12 give similar results, DC-12 is much more reliable close to the dissociation limit, where the multi-reference effects start to play a significant role. The fact that the results of DC-12 are superior to those of orbital-optimized theories, such as OMP3 and OLCCD, indicates that the strong DC-12 performance is likely caused by the nonlinearities present in the description of one-particle correlation effects of this method, rather than orbital relaxation. Our results also suggest that the current n -representability conditions derived from second-order perturbation theory, although not exact for two-electron systems, yield a very good description of their electron correlation. This was recently anticipated by Kutzelnigg¹⁶⁶ in a model study of the two-configuration wavefunction of two-electron systems in the natural orbital basis.

3.5.2 COVALENT BONDING IN DIATOMIC MOLECULES

To assess the quality of DC-12 in describing covalent bonding, the energies, equilibrium bond distances and harmonic vibrational frequencies of eight diatomic molecules with single (HF, OH⁻, F₂), double (O₂, NO⁻) and triple (N₂ and CO, plus BF) bonds computed using several methods are compared against those from CCSD(T) in Table 3.2. The total energies computed using DC-06 are in the closest agreement with the CCSD(T) energies. The new DC-12 method yields total energies that lie in between those of DC-06 and CCSD methods, often being closer to CCSD. For equilibrium bond lengths and harmonic vibrational frequencies (r_e and ω_e , Table 3.2), the results of all methods follow a similar trend for all molecules, except F₂ and NO⁻. The best agreement with the CCSD(T) results is obtained for DC-06, while the largest errors are generally observed for CCSD. Both the LCCD and DC-12 methods tend to exhibit intermediate quality between that of DC-06 and CCSD, showing significant improvement for the molecules with single bonds. Harmonic vibrational frequencies have been reported for the parametric reduced density matrix methods (pRDM)¹⁰⁸ using the cc-pVQZ basis set and the frozen-core approximation. For HF, the (M , M_2) pRDM methods give errors of (13, 18) cm⁻¹ in ω_e , which is comparable to DC-06 (16.1 cm⁻¹) and superior to CCSD (44.5 cm⁻¹) and DC-12 (40.9 cm⁻¹). In the case of F₂ the smallest error among our data is obtained with DC-12, which reproduces CCSD(T) results with only a 0.0004 Å error in r_e and 45.5 cm⁻¹ error in ω_e . This is particularly encouraging, considering the large errors of LCCD and CCSD and the fact that F₂ is known for its complicated electronic structure.¹⁶⁷ The pRDM methods perform surprisingly well for F₂, with the (M , M_2) parameterizations yielding errors of just (17, -1) cm⁻¹ in ω_e ; the corresponding errors for the multiply-bonded CO molecule are (31, 29) cm⁻¹, which are smaller than the 69.6, 89.2, and 71.5 cm⁻¹ errors observed for DC-06, DC-12, and CCSD, respectively. For NO⁻ the best performance is observed for CCSD; however, all methods underestimate the N–O distance by

Table 3.2: Errors in the total energies (E , E_h), equilibrium bond distances (r_e , Å) and harmonic vibrational frequencies (ω_e , cm^{-1}) using the LCCD, DC-06, DC-12, and CCSD methods with respect to results obtained at the CCSD(T) level of theory. For all computations, the cc-pCVQZ basis set was used. Also shown are mean absolute and standard deviations, Δ_{abs} and Δ_{std} , of the r_e and ω_e values computed with respect to the CCSD(T) results.

Molecule		LCCD	DC-06	DC-12	CCSD	CCSD(T)
HF ($X^1\Sigma^+$)	E	0.005015	0.004329	0.006481	0.008358	-100.438010
	r_e	-0.0018	-0.0014	-0.0026	-0.0027	0.9158
	ω_e	24.5	16.1	40.9	44.5	4166.3
OH ⁻ ($X^1\Sigma^+$)	E	0.006082	0.004868	0.007386	0.010883	-75.763528
	r_e	-0.0019	-0.0007	-0.0027	-0.0034	0.9633
	ω_e	18.9	-8.8	35.6	47.8	3718.9
F ₂ ($X^1\Sigma_g^+$)	E	0.007469	^a	0.012716	0.021912	-199.488097
	r_e	0.0203	^a	0.0004	-0.0230	1.4113
	ω_e	-163.1	^a	-45.5	97.6	922.4
O ₂ ($X^3\Sigma_g^-$)	E	0.013941	0.010619	0.017024	0.021090	-150.295654
	r_e	-0.0092	-0.0053	-0.0109	-0.0115	1.2058
	ω_e	56.8	18.9	72.3	85.8	1605.0
NO ⁻ ($X^3\Sigma^-$)	E	0.015219	0.012492	0.016979	0.020606	-129.860352
	r_e	-0.0140	-0.0121	-0.0144	-0.0091	1.2680
	ω_e	90.6	77.3	93.2	52.4	1427.3
N ₂ ($X^1\Sigma_g^+$)	E	0.010189	0.006906	0.014933	0.021282	-109.519838
	r_e	-0.0042	-0.0018	-0.0064	-0.0073	1.0981
	ω_e	39.7	0.4	68.8	81.0	2365.8
CO ($X^1\Sigma^+$)	E	0.011236	0.008905	0.015156	0.019585	-113.303351
	r_e	-0.0069	-0.0061	-0.0083	-0.0072	1.1289
	ω_e	76.5	69.6	89.2	71.5	2173.8
BF ($X^1\Sigma^+$)	E	0.006351	0.003226	0.010265	0.013468	-124.653886
	r_e	-0.0030	0.0002	-0.0054	-0.0039	1.2635
	ω_e	14.3	-2.3	26.7	18.7	1406.9
Δ_{abs}	r_e	0.0077	0.0040	0.0064	0.0085	—
	ω_e	60.6	27.6	59.0	62.4	—
Δ_{std}	r_e	0.0066	0.0042	0.0047	0.0066	—
	ω_e	49.7	32.1	25.2	26.1	—

^aNo convergence was obtained.

Table 3.3: Errors in the total energy (mE_h), relative to full configuration interaction, as a function of bond length (r , in Å) for the HF molecule using the 6-31G** basis set and selected correlation treatments.

r	DCFT	Coupled Cluster		pRDM ^a		
	DC-12	CCSD	CCSD(T)	K	M	M_2
0.9	3.2	2.5	0.4	3.3	2.3	2.1
1.4	6.3	4.7	0.9	4.8	3.0	2.5
1.8	17.1	9.0	0.5	6.6	3.4	2.2
2.2	44.4	14.8	-4.3	9.2	3.8	2.3
2.6	80.7	19.0	-14.8	11.3	4.2	2.8
2.8	97.9	20.3	-20.8	15.7	7.2	4.6

^aParametric reduced density matrix (pRDM) values are from Ref. 108. The basis set stated in that work is 6-311G**, but we found that the Cartesian formulation of the 6-31G** basis is needed to reproduce Hartree-Fock and full CI values therein.

ca. 0.01 Å. The overall performance of the methods in our small test set is reflected by their mean absolute deviations (MADs) of the computed r_e and ω_e values, shown in Table 3.2. Both DC-06 and DC-12 have smaller MADs, compared to CCSD and LCCD. The improved description of the one-particle correlation effects in the DC-12 method leads to the increase of its MADs relative to those of DC-06. This is most likely a consequence of a reduction in error cancellation.

A more challenging test of correlation treatments comes from stretching covalent bonds, which introduces a degree of static correlation. In Table 3.3, we demonstrate the performance of DC-12, and other methods, including pRDM methods, in describing the energy of the HF molecule as a function of its bond length. Although DC-12 can be converged for all stretched bond lengths investigated, unlike its DC-06 progenitor, the performance away from the equilibrium geometry is inferior to the other methods tested, which all contain explicit nonlinearities among the four-index unknowns. The deterioration in performance of DC-12 at highly stretched geometries is attributable to the use of n -representability constraints that

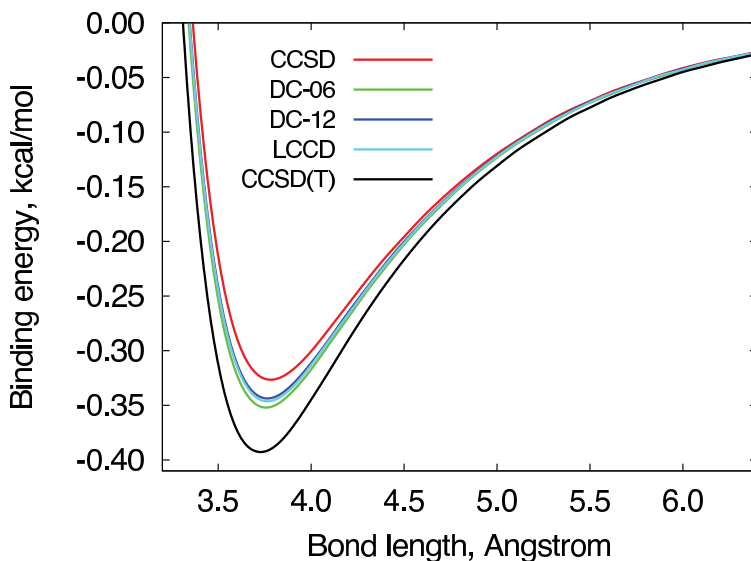


Figure 3.2: Binding potential energy of the argon dimer as a function of Ar–Ar distance computed using the CCSD, CCSD(T), DC-06, DC-12, and LCCD methods with the aug-cc-pVTZ basis set.

were derived using perturbation theory and are, therefore, appropriate for describing dynamical correlation. To obtain a theory that is even more impervious to multireference character, stronger constraints must be employed; this should also help to address the imbalance noted above.

3.5.3 NON-BONDED INTERACTIONS IN AR AND HF DIMERS

The performance of the DC-12 method in describing the non-bonded interactions was tested by computing the dissociation curves of the argon dimer Ar_2 and the linear hydrogen fluoride dimer $[\text{HF} \cdots \text{HF}]$ using the aug-cc-pVTZ basis set (Figures 3.2 and 3.3, respectively). For the HF dimer only the intermolecular $\text{F} \cdots \text{H}$ distance was varied, with each monomer’s H–F distance fixed at its CCSD(T)/cc-pCVQZ equilibrium value, to exclude the influence of the dipole-dipole interactions on the relative binding energy. As in the previous section, we use CCSD(T) for our benchmark, as it is known to describe dispersion very well.¹⁶⁸ In the case

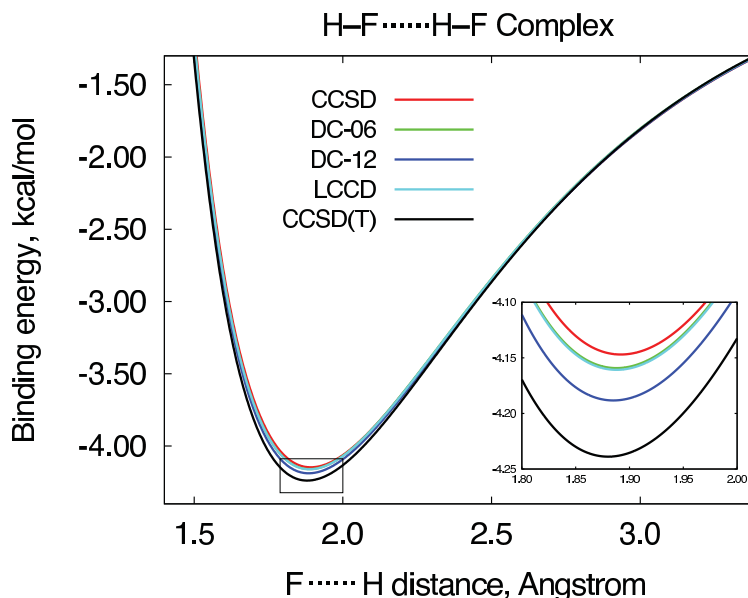


Figure 3.3: Binding energy of the interaction of two hydrogen fluoride molecules in the linear $[\text{HF} \cdots \text{HF}]$ complex as a function of $\text{F} \cdots \text{H}$ distance using the CCSD, CCSD(T), DC-06, DC-12, and LCCD methods with the aug-cc-pVTZ basis set. Each monomer $\text{H}-\text{F}$ distance was fixed at $0.9158 \text{ \AA}$.

of Ar_2 the largest error was obtained for CCSD, which underbinds the dimer by ca. $0.07 \text{ kcal mol}^{-1}$ (Figure 3.2). LCCD, DC-06 and DC-12 show similar results with the errors in the binding energy in the range of 0.04 - $0.05 \text{ kcal mol}^{-1}$. For the dissociation of the $[\text{HF} \cdots \text{HF}]$ complex the best results were obtained using the DC-12 method, with an error of $0.05 \text{ kcal mol}^{-1}$ (Figure 3.3). This is almost a factor of two smaller than the binding energy error of CCSD ($0.09 \text{ kcal mol}^{-1}$). LCCD and DC-06 show very similar results with almost overlapping binding curves, giving an error of $0.07 \text{ kcal mol}^{-1}$ (Figure 3.3).

3.6 CONCLUSIONS

We have presented the derivation and the first benchmark results of the new DC-12 variant of density *cumulant* functional theory that extracts the non-idempotent component of the

one-particle density matrix exactly from the two-particle density cumulant. By a simple transformation to the natural spin-orbital basis we arrived at an efficient formulation of the DC-12 method, which has the same $\mathcal{O}(N^6)$ asymptotic scaling as the old DC-06 method that used an approximate relationship derived from a low order power series expansion. From the perturbative analysis of the DC-12 equations we concluded that effects of the higher-order non-linear terms in the description of the correlation contribution are significant, due to the large contribution of the one-electron terms to the total energy and cumulant residual equations. In addition, our experience shows that the new DC-12 method exhibits better convergence and numerical stability than the old DC-06 formulation.

While DC-12 is not exact for two-electron systems, we have shown that it offers a very good description of the H_2 dissociation curve, both near the equilibrium region and close to the dissociation limit, as a result of the improved description of the one-particle density matrix. We also tested the performance of the DC-12 method in the description of the covalent bonding in the diatomic molecules with single and multiple bonds and the non-bonded interactions in argon and hydrogen fluoride dimers (Ar_2 and $[\text{HF}]_2$). In both cases the results of DC-12 method show a balanced quality ranging between that of the CCSD and CCSD(T) methods. The new formulation appears to outperform DC-06 for non-bonded interactions, however DC-06’s description of covalent bonds around equilibrium appears to be slightly better in some cases. This is most likely due to a reduction in scope for error cancellation, having removed one of the two major approximations made in DC-06. The challenge now is to impose more rigorous n -representability constraints to arrive at a theory that delivers high performance without relying on such balance. Research to this effect is a priority for the future.

In density cumulant functional theory, a hierarchy of size-consistent and size-extensive *ab initio* methods can be introduced by using and systematically improving the approximate conditions that constrain the density cumulant to be n -representable. Deriving the corre-

lation contribution to the one-particle density matrix exactly from the density cumulant, as done in the DC-12 method, allows us to describe both one- and two-particle correlation effects on an equal footing through the approximate n -representability conditions and is, therefore, an important step towards defining this hierarchy.

3.7 ACKNOWLEDGMENTS

This paper is dedicated to the memory of Prof. Olga V. Sizova (1947 – 2012), A.Y.S. undergraduate research advisor, colleague and friend.

This research was supported by the U.S. National Science Foundation, Grant CHE-1054286. A.Y.S. is thankful to Professor Werner Kutzelnigg for valuable discussions. A.Y.S. also thanks Professor Uğur Bozkaya for providing the OLCCD results and insightful discussions.

3.8 APPENDIX: SIMPLIFICATION OF THE EXPRESSIONS IN THE NATURAL SPIN-ORBITAL (NSO) BASIS

Using the results obtained in Eqs. (3.28) – (3.31), we can evaluate the $\boldsymbol{\tau}$ -dependent terms of the stationarity condition (3.24), by transforming them to the NSO basis. Considering the contribution for the occupied space first, we obtain:

$$\begin{aligned} f_k^l \frac{\partial \tau_l^k}{\partial \lambda_{ij}^{ab}} &= f_{k'}^{l'} \frac{\partial \tau_{l'}^{k'}}{\partial \lambda_{i'j'}^{a'b'}} U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b \\ &= -\frac{1}{2} \left(\frac{f_{i'}^{k'}}{1 + \tau_{k'} + \tau_{i'}} \lambda_{k'j'}^{a'b'} - \frac{f_{j'}^{k'}}{1 + \tau_{k'} + \tau_{j'}} \lambda_{k'i'}^{a'b'} \right) U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b . \end{aligned} \quad (3.36)$$

Realizing that the NSO basis density cumulant can be transformed back to the original basis

$$\lambda_{i'j'}^{a'b'} U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b = \lambda_{ij}^{ab} , \quad (3.37)$$

and redistributing the terms in Eq. (3.36) yields:

$$\begin{aligned}
f_k^l \frac{\partial \tau_l^k}{\partial \lambda_{ij}^{ab}} &= -\frac{1}{2} \left(\frac{f_{i'}^{k'} \lambda_{kj'}^{a'b'} (U^\dagger)_{k'}^k U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b}{1 + \tau_{k'} + \tau_{i'}} - \frac{f_{j'}^{k'} \lambda_{ki'}^{a'b'} (U^\dagger)_{k'}^k U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b}{1 + \tau_{k'} + \tau_{j'}} \right) \\
&= -\frac{1}{2} \left(\frac{f_{i'}^{k'} U_i^{i'} (U^\dagger)_{k'}^k}{1 + \tau_{k'} + \tau_{i'}} \lambda_{kj}^{ab} - \frac{f_{j'}^{k'} U_j^{j'} (U^\dagger)_{k'}^k}{1 + \tau_{k'} + \tau_{j'}} \lambda_{ki}^{ab} \right) .
\end{aligned} \tag{3.38}$$

Similarly, for the contribution in the virtual space, we obtain:

$$\begin{aligned}
f_c^d \frac{\partial \tau_d^c}{\partial \lambda_{ij}^{ab}} &= f_{c'}^{d'} \frac{\partial \tau_{d'}^{c'}}{\partial \lambda_{i'j'}^{a'b'}} U_i^{i'} U_j^{j'} (U^\dagger)_{a'}^a (U^\dagger)_{b'}^b \\
&= \frac{1}{2} \left(\frac{f_{a'}^{d'} U_a^{a'} (U^\dagger)_{d'}^d}{1 - \tau_{a'} - \tau_{d'}} \lambda_{ij}^{db} - \frac{f_{b'}^{d'} U_b^{b'} (U^\dagger)_{d'}^d}{1 - \tau_{b'} - \tau_{d'}} \lambda_{ij}^{da} \right) .
\end{aligned} \tag{3.39}$$

CHAPTER 4

ORBITAL-OPTIMIZED

DENSITY CUMULANT FUNCTIONAL THEORY*

*A. Yu. Sokolov and H. F. Schaefer, J. Chem. Phys. **139**, 204110 (2013). Reproduced with permission of the American Institute of Physics.

4.1 ABSTRACT

In density cumulant functional theory (DCFT) the electronic energy is evaluated from the one-particle density matrix and two-particle density cumulant, circumventing the computation of the wavefunction. To achieve this, the one-particle density matrix is decomposed exactly into the mean-field (idempotent) and correlation components. While the latter can be entirely derived from the density cumulant, the former must be obtained by choosing a specific set of orbitals. In the original DCFT formulation [J. Chem. Phys. **125**, 171101 (2006)] the orbitals were determined by diagonalizing the effective Fock operator, which introduces partial orbital relaxation. Here we present a new orbital-optimized formulation of DCFT where the energy is variationally minimized with respect to orbital rotations. This introduces important energy contributions and significantly improves the description of the dynamic correlation. In addition, it greatly simplifies the computation of analytic gradients, for which expressions are also presented. We offer a perturbative analysis of the new orbital stationarity conditions and benchmark their performance for a variety of chemical systems.

4.2 INTRODUCTION

The predictive power of the *ab initio* quantum chemistry relies on the ability to accurately describe electron correlation. Among many theoretical methods that tackle this problem single-reference coupled cluster (CC) theory^{13,17-23} has become especially popular due to size-extensivity and inclusion of products of disconnected cluster operators that model high-order correlation effects. While CC theory offers a well-defined hierarchy of methods that converge to the full configuration interaction (CI) limit, the error of the CC energy does not decrease quadratically with that of the wavefunction due to its non-variational nature.^{169,170} In addition, the computation of molecular properties and analytic gradients in CC theory requires evaluation of two sets of eigenvectors of the non-Hermitian similarity-transformed

Hamiltonian,^{30–33} effectively doubling the cost of the CC methods. Motivated by this, several alternative approaches have been formulated by reducing the non-linearity of the CC equations,¹⁰⁹ introducing extensivity to configuration interaction (CI) theory,^{110,112,115} exploring other formulations of the CC exponential ansatz,^{143,171–173} or directly obtaining the reduced density matrices (RDM),^{63,64} which bypasses the computation of the wavefunction.

In the RDM theory^{63,64} the electronic energy is obtained from the one- and two-particle reduced density matrices (γ_1 and γ_2):

$$\begin{aligned}\gamma_p^q &= \langle \Psi | a_p^q | \Psi \rangle \ , \quad a_p^q \equiv a_q^\dagger a_p \ , \\ \gamma_{pq}^{rs} &= \langle \Psi | a_{pq}^{rs} | \Psi \rangle \ , \quad a_{pq}^{rs} \equiv a_r^\dagger a_s^\dagger a_q a_p \ .\end{aligned}\tag{4.1}$$

Defining the Fock space molecular Hamiltonian⁵⁷ (the Einstein summation convention is implied)

$$\begin{aligned}\hat{H} &= h_p^q a_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} a_{rs}^{pq} \ , \quad h_p^q = \langle \psi_p(1) | \hat{h} | \psi_q(1) \rangle \ , \\ \bar{g}_{pq}^{rs} &= g_{pq}^{rs} - g_{pq}^{sr} \ , \quad g_{pq}^{rs} = \langle \psi_p(1) \psi_q(2) | \frac{1}{r_{12}} | \psi_r(1) \psi_s(2) \rangle \ ,\end{aligned}\tag{4.2}$$

the energy can be written as an exact functional of γ_1 and γ_2 :

$$E = \langle \Psi | \hat{H} | \Psi \rangle = h_p^q \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \gamma_{rs}^{pq} \ .\tag{4.3}$$

Eq. (4.3) provides means for obtaining the energy without explicit computation of the wavefunction. This is usually done by expressing the energy (4.3) solely in terms of γ_2 ,

$$E = \frac{1}{2} \tilde{H}_{pq}^{rs} \gamma_{rs}^{pq} \ , \quad \tilde{H}_{pq}^{rs} \equiv \frac{1}{N-1} (h_p^r \delta_q^s + h_q^s \delta_p^r) + g_{pq}^{rs} \ ,\tag{4.4}$$

and variationally minimizing^{67,86,94,95} the resulting functional (4.4), provided that the density matrices are constrained to be N -representable, *i.e.* correspond to an antisymmetric

N -electron wavefunction.^{59,60,68,75} An alternative approach was proposed in density cumulant functional theory (DCFT),^{96,97,174} where γ_2 is replaced in favor of its fully connected counterpart, the two-particle density cumulant^{58,77,85–91,93} (λ_2):

$$\lambda_{pq}^{rs} = \gamma_{pq}^{rs} - \gamma_p^r \gamma_q^s + \gamma_p^s \gamma_q^r . \quad (4.5)$$

An advantage of the cumulant representation (4.5) is that λ_2 is additively separable, which results in the size-extensive and size-consistent energy functional even when approximations are introduced. Density cumulants were used in many areas of theoretical chemistry, such as generalized normal-ordering⁵⁸ and its applications,^{154–160} contracted Schrödinger equation theory,^{67,77,86,94,95,147–150} canonical transformation theory,^{102–104,151–153} and parametric reduced density matrix methods.^{64,105–108}

We have recently presented the first implementation⁹⁷ and analytic gradients¹⁶¹ of the original DCFT formulation based on the second-order perturbative N -representability conditions (DC-06 method).⁹⁶ Our most recent work¹⁷⁴ described a new (DC-12) formulation, where the N -representability of γ_1 was improved by deriving its correlation contribution exactly from λ_2 , which introduces important higher-order terms in energy and cumulant stationarity conditions without noticeable increase of the computational cost. While incorporating these contributions improves the performance of the method for the two-electron systems, it may give rise to imbalances in the description of electron correlation for systems with many electrons.¹⁷⁴

Herein, we present a new variation of DCFT that takes advantage of the full orbital optimization. We name the orbital-optimized variants of DC-06 and DC-12 as ODC-06 and ODC-12, respectively, to indicate similarities in the description of N -representability in these methods. From the perturbative analysis of the new orbital stationarity conditions we demonstrate that the orbital optimization accounts for the important fourth-order en-

ergy components. In addition, it greatly simplifies the computation of energy gradients, since no response equations must be solved. We present analytic gradient expressions for the ODC-06 and ODC-12 methods and show the improvements of the new formulation for a variety of chemical systems, including notoriously-difficult symmetry-breaking radicals. Particularly, for the ODC-12 method, the combination of the full orbital optimization and the improved description of γ_1 gives rise to a balanced description of electron correlation for the benchmark systems studied. While the proposed DCFT methods are novel, orbital optimization has been widely used in wavefunction-based theories as an alternative to including the single excitations in the wavefunction.^{130,175–182} Recently, variational optimization of the orbitals was used in the research of Bozkaya and co-workers who developed orbital-optimized variants of Møller-Plesset (MP) perturbation theory^{164,165,183,184} and coupled electron pair approximation,¹⁸⁵ as well as CC theory.^{164,186}

4.3 THEORY

4.3.1 BRIEF OVERVIEW OF DCFT

In DCFT the exact energy is expressed in terms of one- ($\lambda_1 = \gamma_1$) and two-particle (λ_2) density cumulants. Using the γ_2 cumulant expansion (4.5), we obtain for the energy:

$$\begin{aligned} E &= h_p^q \gamma_q^p + \frac{1}{2} \bar{g}_{pq}^{rs} \gamma_r^p \gamma_s^q + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq} \\ &= \frac{1}{2} (h_p^q + f_p^q) \gamma_q^p + \frac{1}{4} \bar{g}_{pq}^{rs} \lambda_{rs}^{pq} , \end{aligned} \tag{4.6}$$

where the generalized Fock operator $\mathbf{f}$ is defined as:

$$f_p^q = h_p^q + \bar{g}_{pr}^{qs} \gamma_s^r . \tag{4.7}$$

In Eqs. (4.6) and (4.7) the exact γ_1 is used, which can be separated⁹⁶ into an idempotent¹²⁴ (Hartree-Fock-like) contribution, $\kappa = \kappa^2$, and a correction, τ :

$$\gamma_1 = \kappa + \tau . \quad (4.8)$$

The latter contribution arises due to electron correlation and can be entirely determined from the partial trace of λ_2 :

$$\sum_r \lambda_{pr}^{qr} = (\gamma_1^2 - \gamma_1)_p^q \equiv d_p^q = \tau_p^r \tau_r^q - \tau_p^q + 2\tau_p^r \kappa_r^q . \quad (4.9)$$

Using Eqs. (4.8) and (4.9), the energy (4.6) can be exactly expressed as a functional of κ and λ_2 . While the former can be obtained by choosing a specific set of orbitals, the latter needs to be evaluated by making the energy (4.6) stationary with respect to variation of λ_2 , constrained to N -representability conditions. Our earlier research^{97,174} employed a simple set of approximate conditions derived from the second-order MP perturbation theory by Kutzelnigg and Mukherjee,⁹² which leads to the following expression for the energy:

$$E = \frac{1}{2}(h_p^q + f_p^q)\gamma_q^p + \frac{1}{2}\bar{g}_{ij}^{ab}\lambda_{ab}^{ij} + \frac{1}{8}\lambda_{kl}^{ab}\bar{g}_{ij}^{kl}\lambda_{ab}^{ij} + \frac{1}{8}\lambda_{ij}^{cd}\bar{g}_{cd}^{ab}\lambda_{ab}^{ij} - \lambda_{ik}^{ac}\bar{g}_{jc}^{kb}\lambda_{ab}^{ij} , \quad (4.10)$$

where indices $i, j, k \dots$ and $a, b, c \dots$ run over occupied and virtual spin-orbitals, respectively. To make the energy functional (4.10) stationary with respect to λ_2 , the derivatives of τ need to be evaluated. In the original DC-06 formulation^{96,97} the elements of τ were approximated as:

$$\tau_i^j \approx d_i^j = -\frac{1}{2}\lambda_{ik}^{ab}\lambda_{ab}^{jk} , \quad \tau_a^b \approx -d_a^b = \frac{1}{2}\lambda_{ac}^{ij}\lambda_{ij}^{bc} . \quad (4.11)$$

In the DC-12 method the exact relationship between τ and λ_2 was used:¹⁷⁴

$$\tau_i^j = d_i^j - \tau_i^k \tau_k^j , \quad \tau_a^b = -d_a^b + \tau_a^c \tau_c^b , \quad (4.12)$$

where expressions for d_p^q are shown in (4.11). The resulting stationarity conditions were given elsewhere.^{97,174}

4.3.2 ORBITAL RELAXATION IN THE DC-06 AND DC-12 METHODS

As demonstrated in Sec. 4.3.1, in DCFT the one-particle correlation contribution $\boldsymbol{\tau}$ is entirely derived from the two-particle cumulant $\boldsymbol{\lambda}_2$ using decomposition (4.8) and, therefore, is not an independent parameter. To account for the missing one-particle (mean-field) correlation effects, orbital relaxation needs to be introduced. In the original work of Kutzelnigg⁹⁶ and our later studies^{97,174} the orbital relaxation effects were partially taken into account by minimizing the energy (4.6) with respect to unitary variations of $\boldsymbol{\kappa}$:

$$\tilde{\boldsymbol{\kappa}} = \mathbf{U}^\dagger \boldsymbol{\kappa} \mathbf{U} , \quad (4.13)$$

where the unitary matrix $\mathbf{U}$ can be defined using the exponential parametrization^{187–192}

$$\mathbf{U} = \mathbf{e}^{\mathbf{X} - \mathbf{X}^\dagger} , \quad \mathbf{U} \mathbf{U}^\dagger = \mathbf{1} . \quad (4.14)$$

Making the energy stationary with respect to parameters X_p^q for all non-redundant pairs of p and q at $\mathbf{X} = 0$ leads to a simple stationarity condition:

$$\left(\frac{\partial E}{\partial X_p^q} \right) \Big|_{\mathbf{X}=0} = [\mathbf{f}, \boldsymbol{\kappa}]_p^q = 0 . \quad (4.15)$$

Eqs. (4.13) to (4.15) can be thought of providing a set of orbitals that minimize the mean-field ($\boldsymbol{\kappa}$ -dependent) part of the DCFT energy (4.6) subject to non-local external potential $\mathbf{f}'$ due to electron correlation:

$$\mathbf{f} = \mathbf{f}_0 + \mathbf{f}' , \quad (f_0)_p^q \equiv h_p^q + \bar{g}_{pr}^{qs} \kappa_s^r , \quad (f')_p^q \equiv \bar{g}_{pr}^{qs} \tau_s^r , \quad (4.16)$$

where $\mathbf{f}_0$ is the conventional Fock operator used in Hartree-Fock (HF) theory. This procedure resembles the orbital optimization in Kohn-Sham density functional theory (KS-DFT),¹⁹³ where the approximate (usually local) exchange-correlation potential is used. The orbital stationarity condition (4.15) can be satisfied by self-consistent diagonalization of the generalized Fock operator $\mathbf{f}$. While Eq. (4.15) can be solved efficiently in the atomic-orbital (AO) basis with only $\mathcal{O}(M)^4$ effort where M is the basis set size, it has a number of disadvantages. In particular, the non-variational nature of the orbital relaxation in Eqs. (4.13) and (4.15) complicates the analytic gradient theory and molecular property computations, which require the solution of the coupled orbital-cumulant response equations.¹⁶¹ Another disadvantage will become obvious in Sec. 4.3.4, where we will demonstrate that the orbital relaxation (4.15) introduces contributions to the DCFT energy only at the fifth order in MP perturbation theory, while missing important fourth-order components.

4.3.3 VARIATIONAL ORBITAL OPTIMIZATION: THE ODC-06 AND ODC-12 METHODS

Consider a unitary transformation of $\boldsymbol{\lambda}_2$

$$\lambda_{p'q'}^{r's'} = (U^\dagger)_{p'}^p (U^\dagger)_{q'}^q \lambda_{pq}^{rs} U_r^{r'} U_s^{s'} , \quad (4.17)$$

where $\boldsymbol{\lambda}_2$ indices in the new orbital basis are labeled by the primes. Using the exponential parametrization (4.14) allows us to minimize the energy (4.6) with respect to the unitary transformation of $\boldsymbol{\kappa}$ and $\boldsymbol{\lambda}_2$ simultaneously (Eqs. (4.13) and (4.17)). The resulting stationarity condition can be written as

$$\left(\frac{\partial E}{\partial X_p^q} \right) \Big|_{\mathbf{X}=0} = (\mathbf{F} - \mathbf{F}^\dagger)_p^q = 0 , \quad (4.18)$$

where

$$F_p^q = f_p^r \gamma_r^q + \frac{1}{2} \bar{g}_{pr}^{st} \lambda_{st}^{qr} . \quad (4.19)$$

Eq. (4.18) is general for any formulation of DCFT and is equivalent to the generalized one-particle Brillouin condition (BC₁).^{77,194,195} Satisfying the stationarity condition (4.18) for all non-redundant orbital pairs yields the variationally optimized orbitals that give the lowest energy for a given set of N -representability conditions.

Employing the approximate second-order N -representability conditions⁹² and discriminating between the occupied and virtual orbitals, the explicit form of Eq. (4.19) for the orbital-optimized variants of DC-06 and DC-12 methods (ODC-06 and ODC-12, respectively) can be written as

$$F_p^i = h_p^j \gamma_j^i + 2\bar{g}_{pj}^{kl} \Gamma_{kl}^{ij} + \bar{g}_{pj}^{ab} \Gamma_{ab}^{ij} + \bar{g}_{pa}^{jb} \Gamma_{jb}^{ia} , \quad (4.20)$$

$$F_p^a = h_p^b \gamma_b^a + 2\bar{g}_{pb}^{cd} \Gamma_{cd}^{ab} + \bar{g}_{pb}^{ij} \Gamma_{ij}^{ab} + \bar{g}_{ip}^{jb} \Gamma_{jb}^{ia} , \quad (4.21)$$

where we defined the intermediate two-particle densities Γ_{pq}^{rs} :

$$\Gamma_{kl}^{ij} = \frac{1}{4} P_{-}(kl) [\gamma_k^i \gamma_l^j] + \frac{1}{8} \lambda_{kl}^{ab} \lambda_{ab}^{ij} , \quad (4.22)$$

$$\Gamma_{cd}^{ab} = \frac{1}{4} P_{-}(cd) [\gamma_c^a \gamma_d^b] + \frac{1}{8} \lambda_{cd}^{ij} \lambda_{ij}^{ab} , \quad (4.23)$$

$$\Gamma_{ij}^{ab} = \Gamma_{ab}^{ij} = \frac{1}{2} \lambda_{ij}^{ab} , \quad (4.24)$$

$$\Gamma_{jb}^{ia} = \gamma_i^j \gamma_a^b - \lambda_{bc}^{ik} \lambda_{jk}^{ac} . \quad (4.25)$$

In Eqs. (4.22) and (4.23) the antisymmetric permutation operator $P_{-}(ij)f(i,j) \equiv f(i,j) - f(j,i)$ was introduced. The orbital stationarity conditions (4.18) can be solved iteratively by computing the residual (4.18), followed by the orbital rotation (4.14) and integral transfor-

mation. Eqs. (4.22), (4.23) and (4.25) scale as $\mathcal{O}(M)^6$ with the size of the basis set, which increases the cost of the orbital update step compared to the orbital transformation in DC-06 and DC-12 methods. In addition, Eqs. (4.20) and (4.21) require the transformation of the g_{ij}^{ak} and g_{ab}^{ic} types of two-electron integrals, which are not needed for the energy (4.10) and cumulant residual computation. These costs can be reduced by computing the terms involving g_{ab}^{ic} and g_{ab}^{cd} partially in the AO basis, avoiding explicit integral transformation.¹³⁰ For further efficiency improvements it is important to bypass the explicit computation and storage of the expensive Γ_{ab}^{cd} intermediate by inserting Eq. (4.23) into Eq. (4.21) and changing the order of the contraction. Importantly, as we will demonstrate in Sec. 4.3.5, the intermediates in Eqs. (4.20) to (4.25) can be reused for the analytic gradient computation, significantly lowering the cost of the geometry optimization step. Overall, our experience shows that the ODC-06 method is only 10-30% more expensive than its DC-06 counterpart for a single geometry step (energy and gradient computation), depending on a system considered. When comparing the ODC-12 and the DC-12 methods, this efficiency difference is even smaller due to the complicated non-linear nature of the coupled response equations in the latter method.

4.3.4 PERTURBATIVE ANALYSIS OF THE ORBITAL OPTIMIZATION

Let us estimate the effect of the unitary transformation (4.17) on the DCFT energy. We consider a particular case of the wavefunction Ψ dominated by a single determinant Φ and write the Hamiltonian (4.2) in the following form:

$$\hat{H} = H_0 + \mu V = E_0 + (f_0)_p^q \tilde{a}_q^p + \mu \frac{1}{4} \bar{g}_{pq}^{rs} \tilde{a}_{rs}^{pq}, \quad (4.26)$$

where μ is a perturbation parameter, for which the limit of the wavefunction $\mu \rightarrow 0$ corresponds to a single Slater determinant, and E_0 is the energy contribution that does not

depend on λ_2

$$E_0 \equiv \frac{1}{2}(h_p^q + (f_0)_p^q)\kappa_q^p. \quad (4.27)$$

The expectation value of the normal-ordered operators $\tilde{a}_p^q$ and $\tilde{a}_{pq}^{rs}$ with respect to Φ vanishes. Eq. (4.26) is equivalent to Eq. (4.2) for $\mu = 1$. As shown by Kutzelnigg and Mukherjee,⁹² the only non-vanishing elements of λ_2 to $\mathcal{O}(\mu)$ are $\lambda_{ij}^{ab} = (\lambda_{ab}^{ij})^*$. Thus, the DCFT energy (4.6) may be expanded in the perturbative series:

$$E = E^{(0)} + E^{(1)} + E^{(2)} + E^{(3)} + \dots \quad (4.28)$$

We choose HF orbitals as our reference, in analogy with perturbative analysis in single-reference CC theory,¹³ since we are only interested in the effect of the orbital optimization relative to the zeroth-order wavefunction. For the energy contributions up to $\mathcal{O}(\mu^3)$ we obtain:

$$E^{(0)} = E_0, \quad (4.29)$$

$$E^{(1)} = 0, \quad (4.30)$$

$$E^{(2)} = (f_0)_i^j (\tau^{(2)})_j^i + (f_0)_a^b (\tau^{(2)})_b^a + \frac{1}{2}\bar{g}_{ij}^{ab}\lambda_{ab}^{ij}, \quad (4.31)$$

$$E^{(3)} = \frac{1}{8}\lambda_{kl}^{ab}\bar{g}_{ij}^{kl}\lambda_{ab}^{ij} + \frac{1}{8}\lambda_{ij}^{cd}\bar{g}_{cd}^{ab}\lambda_{ab}^{ij} - \lambda_{ik}^{ac}\bar{g}_{jc}^{kb}\lambda_{ab}^{ij}, \quad (4.32)$$

where elements $(\tau^{(2)})_p^q$ are defined in Eq. (4.11). Since all of the terms in Eqs. (4.29) to (4.32) are included in the energy expression (4.10), the DC and ODC energy is complete through the third order in MP perturbation theory.

Let us now perform a perturbative expansion of the orbital stationarity condition (4.15) used in the DC-06 and DC-12 methods. Recalling that for the HF orbitals condition $[\mathbf{f}_0, \boldsymbol{\kappa}] =$

0 is satisfied, the first non-zero contribution to the orbital residual $(R_i^a)_{\text{DC}}$ appears at $\mathcal{O}(\mu^3)$:

$$(R_i^a)_{\text{DC}}^{(3)} = \bar{g}_{iq}^{pr} (\tau^{(2)})_r^q \kappa_p^a - \kappa_i^p \bar{g}_{pq}^{ar} (\tau^{(2)})_r^q. \quad (4.33)$$

Performing a similar analysis of the new orbital stationarity condition (4.18) used in the ODC methods, we get the first non-zero contribution already at $\mathcal{O}(\mu^2)$, in agreement with the Brillouin theorem:

$$(R_i^a)_{\text{ODC}}^{(2)} = \frac{1}{2} \bar{g}_{ib}^{jk} \lambda_{jk}^{ab} - \frac{1}{2} \lambda_{bc}^{ij} \bar{g}_{aj}^{bc}. \quad (4.34)$$

The residual contributions (4.33) and (4.34) give rise to the orbital rotations that, in the case of the HF orbital reference, first appear in the energy expression by transforming the second-order energy (4.31). Expanding the unitary matrix (4.14) in a power series, we obtain for the leading contributions to the energy correction due to the orbital rotation:

$$\begin{aligned} \Delta E &= \bar{g}_{cj}^{ab} \lambda_{ab}^{ij} A_i^c + \bar{g}_{ij}^{kb} \lambda_{ab}^{ij} (A^\dagger)_k^a + \dots \\ &\equiv \bar{g}_{cj}^{ab} \lambda_{ab}^{cj} + \bar{g}_{ij}^{kb} \lambda_{kb}^{ij} + \dots \end{aligned} \quad (4.35)$$

In Eq. (4.35) the transformation matrix elements $A_p^q \equiv X_p^q - X_q^p$ must be obtained from the residuals (4.33) and (4.34), thus resulting in the fifth- and fourth-order energy contributions for the DC and ODC methods, respectively. Importantly, for the orbital-optimized DCFT methods the fourth-order energy correction (4.35) appears at the same order in perturbation theory as the non-linear terms in the exact relationship between $\boldsymbol{\tau}$ and $\boldsymbol{\lambda}_2$ (Eq. (4.12)).¹⁷⁴ As we will demonstrate in Sec. 4.5, these contributions are similarly important and must be taken into account together to obtain a balanced description of electron correlation. To assess the importance of other fourth-order energy contributions missing in the ODC-12 method, higher-order N -representability conditions need to be employed.

4.3.5 ANALYTIC GRADIENTS FOR THE ODC-06 AND ODC-12 METHODS

To formulate the analytic gradient expressions for the ODC-06 and ODC-12 methods it is convenient to use the method of Lagrange multipliers.^{126–128} We define the Lagrangian functional

$$\mathcal{L} = E + \omega_p^q (S_q^p - \delta_q^p) , \quad (4.36)$$

where the DCFT energy E (Eq. (4.10)) is augmented by the orthonormality constraint for the orbitals with the corresponding Lagrange multipliers ω_p^q . In Eq. (4.36) S_q^p are the overlap integrals and δ_p^q is the Kronecker delta. Inserting the energy (4.10) into Eq. (4.36) and regrouping the terms, we obtain:

$$\mathcal{L} = h_i^j \gamma_j^i + h_a^b \gamma_b^a + \bar{g}_{ij}^{kl} \Gamma_{kl}^{ij} + \bar{g}_{ab}^{cd} \Gamma_{cd}^{ab} + \bar{g}_{ij}^{ab} \Gamma_{ab}^{ij} + \bar{g}_{ia}^{jb} \Gamma_{jb}^{ia} + \omega_p^q (S_q^p - \delta_q^p) . \quad (4.37)$$

The Lagrangian functional (4.37) is stationary with respect to variations of $\boldsymbol{\lambda}_2$ and orbital rotations at convergence. Thus, no response contributions need to be evaluated. From the orbital stationarity condition the Lagrange multipliers ω_p^q may be easily evaluated as:

$$\omega_p^q = -F_p^q , \quad (4.38)$$

where the elements F_p^q are defined in Eqs. (4.20) and (4.21). This allows us to obtain the expression for the derivative of the ODC-06 and ODC-12 energy with respect to an external perturbation χ :

$$\begin{aligned} \frac{dE}{d\chi} = \frac{\partial \mathcal{L}}{\partial \chi} &= (h_i^j)^\chi \gamma_j^i + (h_a^b)^\chi \gamma_b^a + (\bar{g}_{ij}^{kl})^\chi \Gamma_{kl}^{ij} + (\bar{g}_{ab}^{cd})^\chi \Gamma_{cd}^{ab} \\ &+ (\bar{g}_{ij}^{ab})^\chi \Gamma_{ab}^{ij} + (\bar{g}_{ia}^{jb})^\chi \Gamma_{jb}^{ia} - (S_p^q)^\chi F_q^p , \end{aligned} \quad (4.39)$$

where $(h_p^q)^\chi$, $(\bar{g}_{pq}^{rs})^\chi$ and $(S_p^q)^\chi$ are the skeleton one-, antisymmetrized two-electron, and overlap derivative integrals. Conveniently, all of the intermediates necessary for Eq. (4.39) are available at the end of the ODC energy computation (Eqs. (4.20) to (4.25)), which makes the evaluation of the energy first derivatives very efficient. To avoid the costly transformation of the AO-basis two-electron derivative integrals for each perturbation, the contributions in Eq. (4.39) are evaluated by backtransforming the densities γ_p^q , Γ_{pq}^{rs} , and F_p^q to the AO basis.

4.4 COMPUTATIONAL DETAILS

The ODC-06 and ODC-12 methods were implemented in the developer’s version of PSI4¹³¹ along with their analytic gradients. The results were benchmarked against coupled cluster theory with single and double excitations (CCSD)^{13,22,23} and CCSD incorporating the perturbative treatment of triple excitations [CCSD(T)].^{132,133} To gauge the effect of the variational orbital optimization, the results were compared to those of the DC-06^{96,97,161} and DC-12 methods.¹⁷⁴ In addition, coupled electron pair approximation zero (CEPA₀)^{162,163} and orbital-optimized variants of CEPA₀ (OCEPA₀)¹⁸⁵ and third-order MP theory (OMP3)^{165,184} were used for the benchmark. All computations were performed with PSI4 and employed correlations of all electrons. For all computations the correlation-consistent core-valence cc-pCVQZ basis sets of Dunning^{134,135} were used, unless stated otherwise. Tight convergence criteria were enforced for the energy (10^{-10} E_h) and the root mean square of the energy gradient (10^{-6} E_h/a₀). For all methods harmonic vibrational frequencies were computed by numerical differentiation of analytic energy gradients.

4.5 RESULTS AND DISCUSSION

4.5.1 H₂ DISSOCIATION

We begin by evaluating the performance of the ODC-06 and ODC-12 methods for the description of the H₂ dissociation. Figure 4.1 shows the equilibrium region of the H₂ potential energy curve (PEC) computed using four DCFT methods, CEPA₀, OCEPA₀, and full CI. All approximate methods yield PECs lower in energy than the full CI curve. The best agreement with full CI is obtained for the DC-12 and ODC-12 methods with absolute errors in the equilibrium binding energy (ΔBE_e) of 0.13 and 0.24 kcal mol⁻¹, respectively. The coupled electron pair theories CEPA₀ and OCEPA₀ show intermediate performance ($\Delta BE_e = 0.39$ and 0.48 kcal mol⁻¹, respectively), while the largest errors are produced by DC-06 and ODC-06 ($\Delta BE_e = 0.51$ and 0.62 kcal mol⁻¹). A similar situation is observed for the equilibrium bond distances and harmonic vibrational frequencies of H₂ (r_e , ω_e), where DC-12

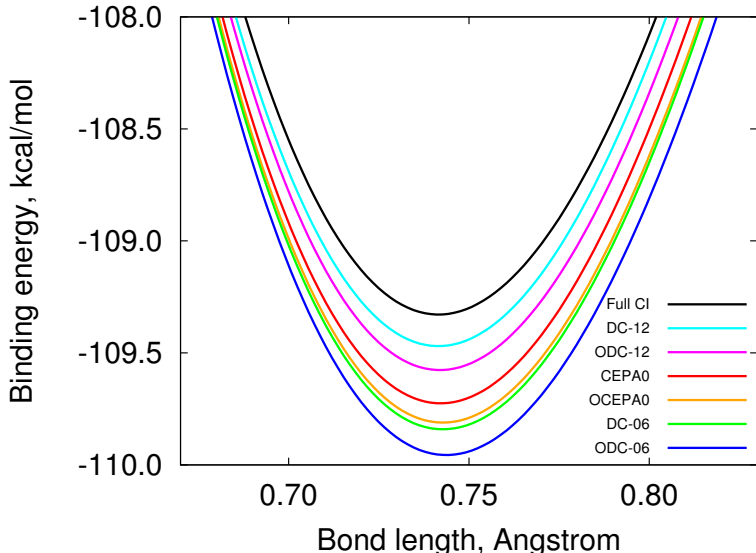


Figure 4.1: Potential energy curve for the equilibrium region of H₂ computed using seven methods with the cc-pV5Z basis set. Energies are relative to the full CI dissociation limit $E_\infty(\text{FCI}) = -0.999989 E_h$.

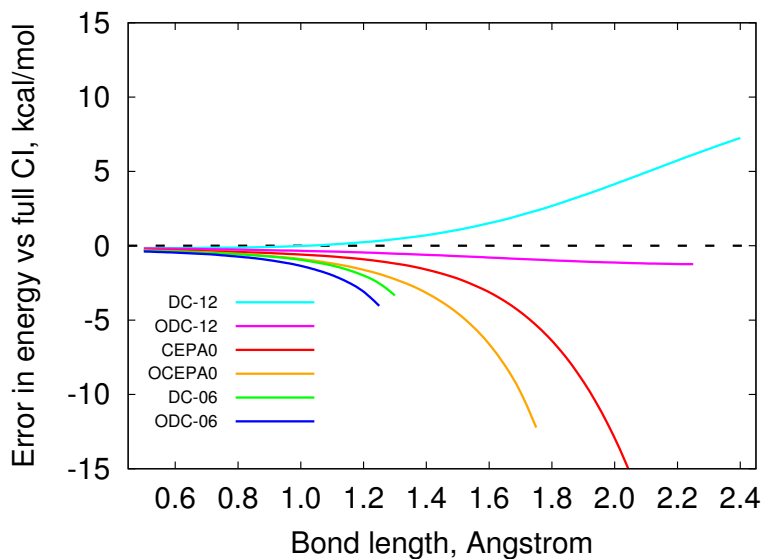


Figure 4.2: Error in the total energy (kcal mol^{-1}), relative to full CI, as a function of H_2 bond length (between 0.5 and 2.4 Å) computed using six methods with the cc-pV5Z basis set. The full CI reference is depicted with a black dashed line. The results are shown only for the H_2 bond length regions where convergence could be obtained.

and ODC-12 again show the best performance with $(\Delta r_e; \Delta \omega_e)$ errors of only $(-0.0002 \text{ Å}; 7 \text{ cm}^{-1})$ and $(0.0004 \text{ Å}; -7 \text{ cm}^{-1})$, respectively.

The performance of the methods for the stretched H_2 bond distances is shown in Figure 4.2, where the error in the H_2 total energy, relative to full CI, is plotted. Among the six methods tested, DC-06, CEPA_0 and their orbital-optimized variants start to deviate from the full CI curve after 1.1 Å, giving rise to large errors in energy for the stretched H_2 distances. For DC-06, ODC-06 and OCEPA_0 the deterioration in performance is accompanied by numerical instabilities, which lead to convergence problems. Combining the better N -representability of the one-particle density matrix (γ_1) and the orbital optimization in the ODC-12 method greatly improves the description of H_2 at the dissociation, leading to only $\sim 1.3 \text{ kcal mol}^{-1}$ energy error at 2.25 Å.

Overall, our results suggest that the orbital optimization and the improved N -representability of γ_1 are similarly important for the description of H_2 potential energy curve, which supports the perturbation theory arguments given in Sec. 4.3.4. In particular, the variational optimization of the orbitals lowers the H_2 total energy, giving rise to a binding energy much lower than the full CI for the stretched H_2 distances, where multireference effects start to play role. Improving the description of γ_1 N -representability raises the binding energy, significantly reducing the errors.

4.5.2 EQUILIBRIUM STRUCTURES AND HARMONIC VIBRATIONAL FREQUENCIES

The performance of the ODC-06 and ODC-12 methods for predicting of the equilibrium structures and harmonic vibrational frequencies was benchmarked for a set of 15 molecules. Tables 4.1 and 4.2 show the errors in the computed equilibrium bond lengths and bond angles, respectively, relative to the accurate empirical equilibrium (r_e) structures that were derived by Bak and co-workers¹³⁶ from experimental rotational constants corrected by theoretical vibration-rotation interaction constants. The results of Table 4.1 are summarized in Figure 4.3a, where the normal distribution of errors in the computed r_e structures are compared between the four DCFT methods, CCSD, and CCSD(T) with the cc-pCVQZ basis set. The best agreement with experiment is obtained for the CCSD(T) method, which yields a sharp normal distribution with the mean absolute deviation (Δ_{abs}) of only 0.0011 Å and the standard deviation (Δ_{std}) of 0.0015 Å. Among the DCFT variants, the ODC-12 method yields the best performance with $\Delta_{abs} = 0.0030$ Å and $\Delta_{std} = 0.0024$ Å (Figure 4.3a), which significantly outperforms CCSD ($\Delta_{abs} = 0.0064$ Å, $\Delta_{std} = 0.0061$ Å). When compared to coupled pair theories (Figure 4.3b), the performance of ODC-12 is superior to CEPA₀ ($\Delta_{abs} = 0.0035$ Å, $\Delta_{std} = 0.0026$ Å) and has a smaller Δ_{std} than that of OCEPA₀ ($\Delta_{abs} = 0.0028$ Å, $\Delta_{std} = 0.0035$ Å). Both DC-12 and ODC-06 methods exhibit broad normal distributions, indicating the imbalanced description of electron correlation.

Table 4.1: Errors in the equilibrium bond lengths (Δr_e , pm) optimized using several methods with the cc-pCVQZ basis set ($\Delta r_e = r_e^{comp} - r_e^{emp}$). Results are compared to the empirical structural parameters¹³⁶ shown in the right column (r_e^{emp} , pm).

Molecule	Bond	Δr_e , pm								r_e^{emp} , pm
		DC-06	DC-12	ODC-06	ODC-12	CEPA ₀	OCEPA ₀	CCSD	CCSD(T)	Empirical
H ₂	H–H	0.19	0.05	0.26	0.11	0.13	0.19	0.07	0.07	74.11 ^a
HF	F–H	–0.25	–0.37	–0.15	–0.28	–0.29	–0.21	–0.38	–0.11	91.69
H ₂ O	O–H	–0.11	–0.27	–0.03	–0.19	–0.18	–0.10	–0.31	–0.04	95.75
HO ^b	O–H	–0.33	–0.52	^b	–0.40	–0.43	–0.28	–0.56	–0.21	96.78
H ₂ O ₂	O–H	–0.61	–0.80	–0.47	–0.70	–0.70	–0.58	–0.85	–0.51	96.70 ^a
HNC	N–H	0.02	–0.12	0.11	–0.05	–0.05	0.04	–0.17	0.10	99.43
NH ₃	N–H	–0.05	–0.22	0.02	–0.16	–0.13	–0.05	–0.29	–0.04	101.16
C ₂ H ₂	C–H	0.12	–0.03	0.16	0.00	0.06	0.08	–0.10	0.08	106.13
HCN	C–H	0.03	–0.12	0.09	–0.08	–0.03	0.02	–0.19	0.02	106.53
C ₂ H ₄	C–H	0.05	–0.09	0.07	–0.07	0.00	0.01	–0.16	0.02	108.07
CH ₄	C–H	0.08	–0.06	0.13	–0.01	0.03	0.08	–0.10	0.05	108.59
CH ₂ O	C–H	–0.73	–0.82	–0.75	–0.80	–0.74	–0.73	–0.84	0.10	110.70
N ₂	N–N	–0.14	–0.60	0.16	–0.39	–0.38	–0.16	–0.69	0.04	109.77
CO	C–O	–0.56	–0.78	–0.24	–0.52	–0.64	–0.38	–0.67	0.05	112.84
CO ₂	C–O	–0.62	–0.78	–0.35	–0.54	–0.67	–0.43	–0.70	0.03	116.01
CH ₂ O	C–O	–0.37	–0.74	0.33	–0.37	–0.52	–0.10	–0.74	–0.04	120.47
HCN	C–N	–0.12	–0.62	0.25	–0.38	–0.37	–0.12	–0.70	0.04	115.34
HNC	C–N	–0.38	–0.67	–0.07	–0.43	–0.51	–0.26	–0.66	0.06	116.87
C ₂ H ₂	C–C	–0.09	–0.59	0.29	–0.35	–0.34	–0.10	–0.67	0.00	120.37
C ₂ H ₄	C–C	0.33	–0.40	1.07	–0.11	–0.11	0.21	–0.57	0.05	133.07
HO ^b	O–F	1.84	–1.35	^b	–0.08	–0.60	1.61	–2.21	–0.18	143.44
H ₂ O ₂	O–O	0.92	–1.49	3.27	–0.70	–0.77	0.45	–2.45	–0.59	145.56 ^a

^aThe H₂ and H₂O₂ bond distances were obtained directly from the experimentally derived equilibrium structures.^{137,138}

^bNo convergence was obtained in the equilibrium region.

Table 4.2: Errors in the equilibrium bond angles ($\Delta\alpha_e$, $^\circ$) optimized using several methods with the cc-pCVQZ basis set ($\Delta\alpha_e = \alpha_e^{comp} - \alpha_e^{emp}$). Results are compared to the empirical structural parameters¹³⁶ shown in the right column (α_e^{emp} , $^\circ$).

Molecule	Bond angle	$\Delta\alpha_e$, $^\circ$								α_e^{emp} , $^\circ$
		DC-06	DC-12	ODC-06	ODC-12	CEPA ₀	OCEPA ₀	CCSD	CCSD(T)	Empirical
H ₂ O	H–O–H	−0.15	0.00	−0.26	−0.10	−0.09	−0.18	0.01	−0.29	104.51
HO ₂ F	H–O–F	−0.07	0.62	^a	0.05	0.45	−0.41	0.69	−0.09	97.94
H ₂ O ₂	H–O–O	−2.81	−1.94	−3.61	−2.23	−2.20	−2.62	−1.58	−2.33	102.32 ^b
NH ₃	H–N–H	−0.86	−0.68	−0.95	−0.76	−0.78	−0.87	−0.61	−0.89	107.25
C ₂ H ₄	H–C–H	−0.13	−0.14	−0.07	−0.12	−0.13	−0.11	−0.16	−0.04	117.14
CH ₂ O	H–C–H	−0.02	−0.17	0.08	−0.18	−0.11	−0.11	−0.36	−0.30	116.74

^aNo convergence was obtained in the equilibrium region.

^bThe H₂O₂ bond angle was obtained directly from the experimentally derived equilibrium structure.¹³⁸

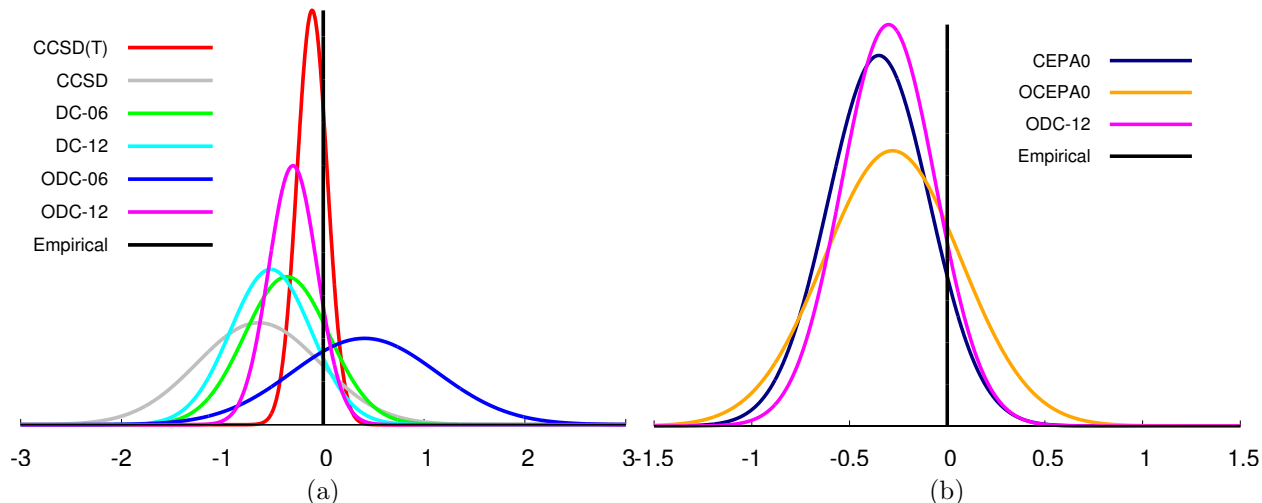


Figure 4.3: Normal distribution of errors in the computed bond distances (r_e , pm) with respect to the empirical structures.¹³⁶ For all computations the cc-pCVQZ basis set was used. Plot (a) compares the performance of four DCFT methods with that of CCSD and CCSD(T), while (b) compares the ODC-12 method with CEPA₀ and its orbital-optimized variant (OCEPA₀). All distributions are normalized to one.

Figures 4.4a and 4.4b show normal distributions of errors in the harmonic vibrational frequencies (ω_e) predicted using the orbital-optimized DCFT methods, relative to those of CCSD(T). The computed ω_e values are presented in the supplementary material (Sec. 4.8). Out of six methods tested, the ODC-12 method exhibits the best agreement with CCSD(T) ($\Delta_{abs} = 12.0 \text{ cm}^{-1}$, $\Delta_{std} = 11.7 \text{ cm}^{-1}$), with Δ_{abs} smaller than that of CCSD ($\Delta_{abs} = 30.2 \text{ cm}^{-1}$, $\Delta_{std} = 17.8 \text{ cm}^{-1}$) by more than a factor of two (Figure 4.4a). All other DCFT methods yield broad normal distributions with large ω_e errors ($\Delta\omega_e$) obtained for the O–F and O–O stretching modes in the electron-dense HOF and H₂O₂ molecules. In addition, the DC-12 method exhibits an anomalously large $\Delta\omega_e$ of 252 cm^{-1} for the degenerate bending mode of HCN. All of these errors are greatly reduced in ODC-12, which produces $\Delta\omega_e$ of only 23 cm^{-1} and 11 cm^{-1} for the O–F and O–O stretches in HOF and H₂O₂, respectively (cf. $\Delta\omega_e$ of 177 and 122 cm^{-1} for DC-06). Large $\Delta\omega_e$ values for the O–F and O–O stretches

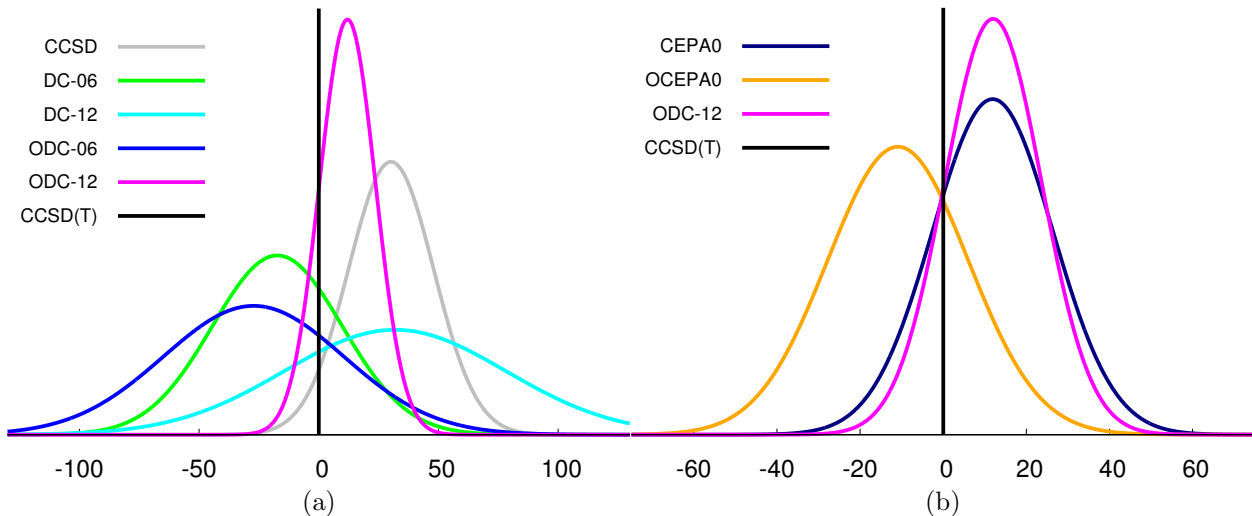


Figure 4.4: Normal distribution of errors in the computed harmonic vibrational frequencies (ω_e , cm^{-1}) with respect to the frequencies obtained from CCSD(T). For all computations the cc-pCVQZ basis set was used. Plot (a) compares the performance of four DCFT methods with that of CCSD, while (b) compares the ODC-12 method with CEPA₀ and its orbital-optimized variant (OCEPA₀). All distributions are normalized to one.

were also obtained for the OCEPA₀ method ($\Delta\omega_e = 112$ and 71 cm^{-1} , respectively) leading to the broader normal distribution ($\Delta_{abs} = 10.9 \text{ cm}^{-1}$, $\Delta_{std} = 16.9 \text{ cm}^{-1}$), compared to the ODC-12 method (Figure 4.4b). CEPA₀ yields similar performance to ODC-12, but with a wider normal distribution ($\Delta_{abs} = 11.9 \text{ cm}^{-1}$, $\Delta_{std} = 14.5 \text{ cm}^{-1}$), as shown in Figure 4.4b.

4.5.3 SPIN-SYMMETRY-BREAKING RADICALS

Finally, we assess the effects of orbital optimization for a challenging set of diatomic radicals that exhibit spin-symmetry breaking (NO, CN, OF, CO⁺, O₂⁺, N₂⁺, and F₂⁺). These systems have been shown to be difficult for conventional electronic structure methods based on the Hartree-Fock orbitals (e.g., MP2, CCSD).^{165,199–202} Figures 4.5a and 4.5b demonstrate the percent errors in the equilibrium bond distances and harmonic vibrational frequencies (Δr_e , $\Delta\omega_e$, %), relative to experiment,^{137,196–198} computed using DC-12, ODC-12, CCSD

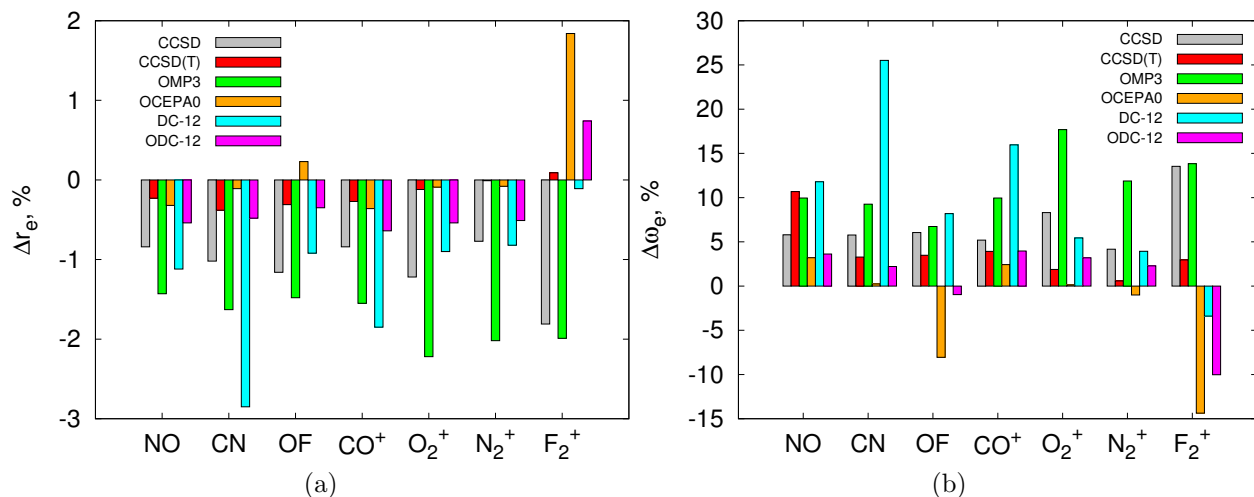


Figure 4.5: Signed percent errors of the computed equilibrium bond distances, Δr_e (plot (a)), and harmonic vibrational constants, $\Delta \omega_e$ (plot (b)), for seven spin-symmetry-breaking radicals, relative to experimental values.^{137, 196–198} The cc-pCVQZ basis set was used.

and CCSD(T). In addition, the orbital-optimized OCEPA₀ and OMP3 methods were used. For the equilibrium structures (Figure 4.5a), the best results are shown by CCSD(T) with $\Delta_{abs} = 0.20$ % and $\Delta_{std} = 0.13$ %. The importance of orbital optimization can be easily seen by comparing the performance of the DC-12 and ODC-12 methods: the large errors of the former ($\Delta_{abs} = 1.22$ %, $\Delta_{std} = 0.88$ %) are reduced by more than a factor of two, on average, for the latter ($\Delta_{abs} = 0.54$ %, $\Delta_{std} = 0.12$ %). OCEPA₀ yields smaller mean absolute deviation ($\Delta_{abs} = 0.44$ %) than ODC-12, but has a much larger standard deviation ($\Delta_{std} = 0.63$ %). For the harmonic constants (Figure 4.5b), both CCSD(T) and ODC-12 show the same $\Delta_{abs} = 3.8$ % value, while ODC-12 fortuitously produces smaller standard deviation ($\Delta_{std} = 2.9$ % and 3.2 % for ODC-12 and CCSD(T), respectively). The performance of OCEPA₀ is slightly worse ($\Delta_{abs} = 4.2$ %, $\Delta_{std} = 5.3$ %), whereas the worst results are shown by OMP3 ($\Delta_{abs} = 11.3$ %, $\Delta_{std} = 3.6$ %) and DC-12 ($\Delta_{abs} = 10.6$ %, $\Delta_{std} = 8.0$ %).

4.6 CONCLUSIONS

In this work we presented a new formulation of density cumulant functional theory (DCFT) that employs variational optimization of the orbitals. Combining the orbital optimization with the approximate N -representability conditions used in the older DCFT variants (DC-06 and DC-12)^{96,97,174} allowed us to formulate the new ODC-06 and ODC-12 methods. The orbital-optimized DCFT methods have two important advantages. Firstly, the variational optimization of the orbitals results in the simple and efficient expressions for the analytic first derivatives of the energy, since no response contributions need to be obtained. This allows for the efficient evaluation of the first-order molecular properties, as the density matrices obtained at the end of the energy computation are fully relaxed. Secondly, as we demonstrated from the perturbative analysis of the orbital stationarity conditions, the orbital optimization introduces important energy contributions that appear at the fourth order in Møller-Plesset perturbation theory, thus greatly improving the description of electron correlation. Our benchmark results for the ODC-06 and ODC-12 methods support this conclusion. In particular, the ODC-12 method, which combines the orbital optimization and the improved description of the one-particle density matrix N -representability (another fourth-order effect), exhibits consistently good performance for all benchmark systems studied, including covalently bound closed-shell molecules, spin-symmetry-breaking open-shell diatomics, as well as H_2 at the equilibrium and near dissociation. Our results suggest that in the equilibrium region the performance of the ODC-12 method is similar to that of the coupled-pair theories (such as CEPA₀ and OCEPA₀), while it is more stable in more complicated situations, where multireference effects become significant. We plan the more extensive benchmark of the DCFT methods and the formulation of more accurate N -representability conditions.

4.7 ACKNOWLEDGMENTS

This research was supported by the U.S. National Science Foundation, Grant CHE-1054286. A.Y.S. thanks Professor Werner Kutzelnigg, Professor Marcel Nooijen, and Dr. Christian Kollmar for insightful discussions. A.Y.S. also thanks Professor Uğur Bozkaya for many insights concerning the OCEPA₀ method.

4.8 SUPPLEMENTARY MATERIAL

See supplementary material at <http://dx.doi.org/10.1063/1.4833138> for tables with energies, harmonic frequencies and structural parameters obtained using DC-06, DC-12, ODC-06, ODC-12, CEPA₀, OCEPA₀, CCSD and CCSD(T) methods.

CHAPTER 5

CONCLUSIONS

The concept of the density cumulant, introduced in 1997, has already had a great influence on electronic structure theory. Density cumulants naturally appear in the generalized normal ordering, which allowed to formulate multi-reference methods for the treatment of static and explicit electron correlation. Density cumulants also play an important role in the reduced density matrix (RDM) theory providing a physical basis for the reconstruction of the high-order density matrices in the solution of the contracted Schrödinger equation. While in all of these developments the cumulant expansion of RDMs is primarily used as a tool for the simplification of the quantum-mechanical equations, in DCFT the density cumulant appears as the central variable of the theory. By expressing the energy in terms of λ_2 from the start, in DCFT we arrive at a density matrix approach to electron correlation that has a close connection with the popular wavefunction-based theories, such as Møller-Plesset perturbation theory (MPPT) and coupled-cluster theory.

In this work we have developed a variety of modifications (namely, DC-12, ODC-06, and ODC-12) of the original DCFT method (DC-06) proposed by Kutzelnigg in 2006. While all of these methods have a similar description of the density cumulant N -representability correct to second order in MPPT, the treatment of the one-particle density matrix N -representability and the orbital relaxation effects is different. In particular, we have demonstrated that the

improved description of the one-particle density matrix and the full orbital optimization are equally important, and, when taken into account together, greatly enhance the quality of the theory, preserving a simple coupled-pair-like nature of the equations. This result is fundamentally important, since the significance of the higher-order one-particle correlation effects has not been addressed in the literature. These one-particle density matrix contributions do not appear in the traditional coupled-cluster theory, while naturally arising in the DC-12 and ODC-12 methods in a very compact and efficient form.

Another important step in the development of DCFT presented herein is the DCFT analytic gradient theory. In Chapter 2 we have demonstrated that in order to obtain the DC-06 analytic gradients a coupled set of orbital-cumulant response equations need to be solved. While the response coupling between the orbitals and the cumulant can be efficiently taken into account, in Chapter 4 we have shown that no response equations need to be solved for the orbital-optimized DCFT methods. The development of the DCFT analytic gradients widely expanded the applicability of the theory and allowed us to provide important benchmark results for a variety of chemical systems with large basis sets. All of the methods presented herein were implemented in the open-source PSI4 quantum chemistry program and are already available to the scientific community free of charge. The use of the DCFT methods for the computation of the first-order molecular properties is particularly appealing, since they provide the relaxed one- and two-particle density matrices automatically at the end of the energy evaluation. The orbital-optimized DCFT formulation can be especially useful for the prediction of the molecular optical rotation properties as it does not suffer from the gauge invariance problem.

The work presented herein provides a foundation for the future development of DCFT. The challenge now is to develop higher-level DCFT formulations with the improved N -representability conditions, aiming for constructing a hierarchy of methods. Several areas of research are the priority for the future. Of particular importance are the generalizations of

the theory capable of describing strongly correlated systems, such as excited states, degenerate (or near-degenerate) ground states, bond breaking. In such systems the assumption that the one-particle density matrix can be well described by a single idempotent contribution κ no longer holds and need to be avoided. Another task is to develop DCFT methods for the treatment of higher-order correlation effects, such as three- and four-electron correlations. The research in this direction is ongoing. Finally, the applicability of the DCFT methods developed in this work needs to be further expanded by performing their efficient implementations capable of describing large molecular systems. The efficiency of the current DCFT formulations will greatly benefit from the numerical techniques such as density fitting, Cholesky decomposition, tensor decompositions and local approximations. These techniques can provide orders of magnitude of computational savings without noticeable loss in accuracy and are becoming popular in the quantum chemistry community nowadays.

BIBLIOGRAPHY

- ¹ T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic Structure Theory* (John Wiley & Sons, Ltd., New York, 2000).
- ² K. Raghavachari and J. B. Anderson, J. Phys. Chem. **100**, 12960 (1996).
- ³ E. Wigner and F. Seitz, Phys. Rev. **46**, 509 (1934).
- ⁴ W. Kutzelnigg, G. Del Re, and G. Berthier, Phys. Rev. **172**, 49 (1968).
- ⁵ P. Löwdin, Adv. Chem. Phys. **2**, 207 (1959).
- ⁶ E. Schrödinger, Phys. Rev. **28**, 1049 (1926).
- ⁷ M. Born and R. Oppenheimer, Ann. Phys. **389**, 457 (1927).
- ⁸ R. McWeeny, Proc. R. Soc. Lond. A **253**, 242 (1959).
- ⁹ R. McWeeny, Rev. Mod. Phys. **32**, 335 (1960).
- ¹⁰ J. Slater, Phys. Rev. **31**, 333 (1928).
- ¹¹ J. Slater, Phys. Rev. **34**, 1293 (1929).
- ¹² A. Szabo and N. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory* (Macmillan, New York, 1982).

- ¹³ I. Shavitt and R. J. Bartlett, *Many-Body Methods in Chemistry and Physics* (Cambridge University Press, Cambridge, UK, 2009).
- ¹⁴ J. Pople, R. Seeger, and R. Krishnan, *Int. J. Quant. Chem.* **12**, 149 (1977).
- ¹⁵ R. Krishnan, H. B. Schlegel, and J. A. Pople, *J. Chem. Phys.* **72**, 4654 (1980).
- ¹⁶ C. D. Sherrill and H. F. Schaefer, *Adv. Quant. Chem.* **34**, 143 (1999).
- ¹⁷ F. Coester, *Nucl. Phys.* **7**, 421 (1958).
- ¹⁸ F. Coester and H. Kümmel, *Nucl. Phys.* **17**, 477 (1960).
- ¹⁹ J. Čížek, *J. Chem. Phys.* **45**, 4256 (1966).
- ²⁰ R. J. Bartlett and G. D. Purvis, *Int. J. Quant. Chem.* **14**, 561 (1978).
- ²¹ R. J. Bartlett, *Annu. Rev. Phys. Chem.* **32**, 359 (1981).
- ²² T. D. Crawford and H. F. Schaefer, *Rev. Comp. Chem.* **14**, 33 (2000).
- ²³ R. J. Bartlett and M. Musiał, *Rev. Mod. Phys.* **79**, 291 (2007).
- ²⁴ Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, *J. Chem. Phys.* **81**, 5906 (1984).
- ²⁵ J. Noga and R. J. Bartlett, *J. Chem. Phys.* **86**, 7041 (1987).
- ²⁶ G. E. Scuseria and H. F. Schaefer, *Chem. Phys. Lett.* **152**, 382 (1988).
- ²⁷ J. D. Watts and R. J. Bartlett, *J. Chem. Phys.* **93**, 6104 (1990).
- ²⁸ S. A. Kucharski and R. J. Bartlett, *Theor. Chem. Acc.* **80**, 387 (1991).
- ²⁹ S. A. Kucharski and R. J. Bartlett, *J. Chem. Phys.* **97**, 4282 (1992).

- ³⁰ A. C. Scheiner, G. E. Scuseria, J. E. Rice, T. J. Lee, and H. F. Schaefer, J. Chem. Phys. **87**, 5361 (1987).
- ³¹ E. A. Salter, G. W. Trucks, and R. J. Bartlett, J. Chem. Phys. **90**, 1752 (1989).
- ³² J. Gauss, J. F. Stanton, and R. J. Bartlett, J. Chem. Phys. **95**, 2623 (1991).
- ³³ J. Gauss, W. J. Lauderdale, J. F. Stanton, J. D. Watts, and R. J. Bartlett, Chem. Phys. Lett. **182**, 207 (1991).
- ³⁴ C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).
- ³⁵ D. Cremer, WIREs Comput. Mol. Sci. **1**, 509 (2011).
- ³⁶ R. J. Bartlett and D. M. Silver, Int. J. Quant. Chem. **8**, 271 (1974).
- ³⁷ J. Binkley and J. Pople, Int. J. Quant. Chem. **9**, 229 (1975).
- ³⁸ M. Head-Gordon, J. A. Pople, and M. J. Frisch, Chem. Phys. Lett. **153**, 503 (1988).
- ³⁹ J. A. Pople, J. S. Binkley, and R. Seeger, Int. J. Quant. Chem. **10**, 1 (1976).
- ⁴⁰ R. J. Bartlett and I. Shavitt, Chem. Phys. Lett. **50**, 190 (1977).
- ⁴¹ R. Krishnan and J. A. Pople, Int. J. Quant. Chem. **14**, 91 (1978).
- ⁴² R. Krishnan, M. J. Frisch, and J. A. Pople, J. Chem. Phys. **72**, 4244 (1980).
- ⁴³ M. L. Leininger, W. D. Allen, H. F. Schaefer, and C. D. Sherrill, J. Chem. Phys. **112**, 9213 (2000).
- ⁴⁴ R. J. Bartlett and G. D. Purvis, III, Phys. Scr. **21**, 255 (1980).
- ⁴⁵ M. Nooijen, K. R. Shamasundar, and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).

- ⁴⁶ J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, *Int. J. Quantum Chem.* **14**, 545 (1978).
- ⁴⁷ J. Goldstone, *Proc. R. Soc. Lond. A* **239**, 267 (1957).
- ⁴⁸ B. Brandow, *Rev. Mod. Phys.* **39**, 771 (1967).
- ⁴⁹ D. Mukherjee, R. K. Moitra, and A. Mukhopadhyay, *Mol. Phys.* **30**, 1861 (1975).
- ⁵⁰ I. Lindgren, *Int. J. Quant. Chem.* **14**, 33 (1978).
- ⁵¹ G. Hose and U. Kaldor, *J. Phys. B: At. Mol. Opt. Phys.* **12**, 3827 (1979).
- ⁵² K. Husimi, *Proc. Phys.-Math. Soc. Japan* **22**, 264 (1940).
- ⁵³ P. O. Löwdin, *Phys. Rev.* **97**, 1474 (1955).
- ⁵⁴ R. McWeeny, *Proc. R. Soc. Lond. A* **232**, 114 (1955).
- ⁵⁵ P. O. Löwdin, *Annu. Rev. Phys. Chem.* **11**, 107 (1960).
- ⁵⁶ E. Davidson, *Reduced Density Matrices in Quantum Chemistry* (Academic Press, New York, 1976).
- ⁵⁷ W. Kutzelnigg, *J. Chem. Phys.* **77**, 3081 (1982).
- ⁵⁸ W. Kutzelnigg and D. Mukherjee, *J. Chem. Phys.* **107**, 432 (1997).
- ⁵⁹ A. J. Coleman, *Rev. Mod. Phys.* **35**, 668 (1963).
- ⁶⁰ C. Garrod and J. K. Percus, *J. Math. Phys.* **5**, 1756 (1964).
- ⁶¹ A. J. Coleman, *J. Math. Phys.* **6**, 1425 (1965).
- ⁶² H. Kummer, *J. Math. Phys.* **8**, 2063 (1967).

- ⁶³ D. A. Mazziotti, ed., *Reduced-Density-Matrix Mechanics: With Applications to Many-Electron Atoms and Molecules, A Special Volume of Advances in Chemical Physics*, vol. 134 (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).
- ⁶⁴ D. A. Mazziotti, Chem. Rev. **112**, 244 (2012).
- ⁶⁵ C. Garrod and M. Rosina, J. Math. Phys. **10**, 1855 (1969).
- ⁶⁶ C. Garrod, M. V. Mihailovic, and M. Rosina, J. Math. Phys. **16**, 868 (1975).
- ⁶⁷ D. A. Mazziotti, J. Chem. Phys. **116**, 1239 (2002).
- ⁶⁸ F. Weinhold, J. Chem. Phys. **47**, 2298 (1967).
- ⁶⁹ A. J. Coleman, J. Math. Phys. **13**, 214 (1972).
- ⁷⁰ A. Coleman, Rep. Math. Phys. **4**, 113 (1973).
- ⁷¹ A. J. Coleman, Int. J. Quant. Chem. **11**, 907 (1977).
- ⁷² A. J. Coleman, Int. J. Quant. Chem. **13**, 67 (1978).
- ⁷³ J. E. Harriman, Phys. Rev. A **17**, 1249 (1978).
- ⁷⁴ J. E. Harriman, Phys. Rev. A **17**, 1257 (1978).
- ⁷⁵ E. R. Davidson, Chem. Phys. Lett. **246**, 209 (1995).
- ⁷⁶ M. Nakata, B. J. Braams, M. Fukuda, J. K. Percus, M. Yamashita, and Z. Zhao, J. Chem. Phys. **125**, 244109 (2006).
- ⁷⁷ D. Mukherjee and W. Kutzelnigg, J. Chem. Phys. **114**, 2047 (2001).
- ⁷⁸ M. Nakata, H. Nakatsuji, M. Ehara, M. Fukuda, K. Nakata, and K. Fujisawa, J. Chem. Phys. **114**, 8282 (2001).

- ⁷⁹ D. A. Mazziotti, Phys. Rev. A **65**, 062511 (2002).
- ⁸⁰ M. Nakata, B. J. Braams, K. Fujisawa, M. Fukuda, J. K. Percus, M. Yamashita, and Z. Zhao, J. Chem. Phys. **128**, 164113 (2008).
- ⁸¹ B. Verstichel, H. van Aggelen, D. Van Neck, P. W. Ayers, and P. Bultinck, Comput. Phys. Commun. **182**, 2025 (2011).
- ⁸² M. Nakata and K. Yasuda, Phys. Rev. A **80**, 042109 (2009).
- ⁸³ H. van Aggelen, B. Verstichel, P. Bultinck, D. Van Neck, P. W. Ayers, and D. L. Cooper, J. Chem. Phys. **132**, 114112 (2010).
- ⁸⁴ B. Verstichel, H. van Aggelen, D. Van Neck, P. W. Ayers, and P. Bultinck, J. Chem. Phys. **132**, 114113 (2010).
- ⁸⁵ D. A. Mazziotti, Chem. Phys. Lett. **289**, 419 (1998).
- ⁸⁶ D. A. Mazziotti, Phys. Rev. A **57**, 4219 (1998).
- ⁸⁷ W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **110**, 2800 (1999).
- ⁸⁸ J. M. Herbert, Int. J. Quantum Chem. **107**, 703 (2006).
- ⁸⁹ J. M. Herbert and J. E. Harriman, Adv. Chem. Phys. **134**, 261 (2007).
- ⁹⁰ L. Kong and E. F. Valeev, J. Chem. Phys. **134**, 214109 (2011).
- ⁹¹ M. Hanauer and A. Köhn, Chem. Phys. **401**, 50 (2012).
- ⁹² W. Kutzelnigg and D. Mukherjee, J. Chem. Phys. **120**, 7350 (2004).
- ⁹³ L. Kong, Int. J. Quantum Chem. **111**, 3541 (2011).
- ⁹⁴ F. Colmenero and C. Valdemoro, Phys. Rev. A **47**, 979 (1993).

- ⁹⁵ H. Nakatsuji and K. Yasuda, Phys. Rev. Lett. **76**, 1039 (1996).
- ⁹⁶ W. Kutzelnigg, J. Chem. Phys. **125**, 171101 (2006).
- ⁹⁷ A. C. Simmonett, J. J. Wilke, H. F. Schaefer, and W. Kutzelnigg, J. Chem. Phys. **133**, 174122 (2010).
- ⁹⁸ G. K.-L. Chan and M. Head-Gordon, J. Chem. Phys. **118**, 8551 (2003).
- ⁹⁹ G. K.-L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).
- ¹⁰⁰ G. K.-L. Chan and T. Van Voorhis, J. Chem. Phys. **122**, 204101 (2005).
- ¹⁰¹ G. K.-L. Chan and S. Sharma, Annu. Rev. Phys. Chem. **62**, 465 (2011).
- ¹⁰² T. T. Yanai and G. K.-L. Chan, J. Chem. Phys. **124**, 194106 (2006).
- ¹⁰³ T. T. Yanai and G. K.-L. Chan, J. Chem. Phys. **127**, 104107 (2007).
- ¹⁰⁴ E. Neuscamman, T. T. Yanai, and G. K.-L. Chan, Int. Rev. Phys. Chem. **29**, 231 (2010).
- ¹⁰⁵ C. Kollmar, J. Chem. Phys. **125**, 084108 (2006).
- ¹⁰⁶ A. E. DePrince and D. A. Mazziotti, Phys. Rev. A **76**, 042501 (2007).
- ¹⁰⁷ D. A. Mazziotti, Phys. Rev. Lett. **101**, 253002 (2008).
- ¹⁰⁸ D. A. Mazziotti, Phys. Rev. A **81**, 062515 (2010).
- ¹⁰⁹ A. G. Taube and R. J. Bartlett, J. Chem. Phys. **130**, 144112 (2009).
- ¹¹⁰ F. Wennmohs and F. Neese, Chem. Phys. **343**, 217 (2008).
- ¹¹¹ F. Neese, A. Hansen, F. Wennmohs, and S. Grimme, Acc. Chem. Res. **42**, 641 (2009).
- ¹¹² F. Neese, F. Wennmohs, and A. Hansen, J. Chem. Phys. **130**, 114108 (2009).

- ¹¹³ A. Hansen, D. G. Liakos, and F. Neese, J. Chem. Phys. **135**, 214102 (2011).
- ¹¹⁴ C. Kollmar and F. Neese, J. Chem. Phys. **135**, 084102 (2011).
- ¹¹⁵ C. Kollmar and F. Neese, Mol. Phys. **108**, 2449 (2010).
- ¹¹⁶ B. R. Brooks, W. D. Laidig, P. Saxe, J. D. Goddard, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **72**, 4652 (1980).
- ¹¹⁷ Y. Osamura, Y. Yamaguchi, and H. F. Schaefer, J. Chem. Phys. **77**, 383 (1982).
- ¹¹⁸ J. A. Pople, R. Krishnan, H. B. Schlegel, and J. S. Binkley, Int. J. Quantum Chem. **16**, 225 (1979).
- ¹¹⁹ P. Jørgensen and J. Simons, J. Chem. Phys. **79**, 334 (1983).
- ¹²⁰ G. Fitzgerald, R. Harrison, W. D. Laidig, and R. J. Bartlett, Chem. Phys. Lett. **117**, 433 (1985).
- ¹²¹ M. Kállay, J. Gauss, and P. G. Szalay, J. Chem. Phys. **119**, 2991 (2003).
- ¹²² R. Kobayashi, N. C. Handy, R. D. Amos, G. W. Trucks, M. Frisch, and J. A. Pople, J. Chem. Phys. **95**, 6723 (1991).
- ¹²³ A. G. Taube and R. J. Bartlett, J. Chem. Phys. **128**, 164101 (2008).
- ¹²⁴ W. Kutzelnigg and V. H. Smith, J. Chem. Phys. **41**, 896 (1964).
- ¹²⁵ S. Polezzo, Theor. Chem. Acc. **38**, 211 (1975).
- ¹²⁶ T. Helgaker and P. Jørgensen, Adv. Quant. Chem. **19**, 183 (1988).
- ¹²⁷ P. G. Szalay, Int. J. Quant. Chem. **55**, 151 (1995).
- ¹²⁸ P. Jørgensen and T. Helgaker, J. Chem. Phys. **89**, 1560 (1988).

- ¹²⁹ N. C. Handy and H. F. Schaefer, J. Chem. Phys. **81**, 5031 (1984).
- ¹³⁰ C. Hampel, K. A. Peterson, and H.-J. Werner, Chem. Phys. Lett. **190**, 1 (1992).
- ¹³¹ J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams, N. J. Russ, M. L. Leininger, C. L. Janssen, E. T. Seidl, W. D. Allen, H. F. Schaefer, R. A. King, E. F. Valeev, C. D. Sherrill, and T. D. Crawford, WIREs Comput. Mol. Sci. **2**, 556 (2012).
- ¹³² K. Raghavachari, G. W. Trucks, J. A. Pople, and M. Head-Gordon, Chem. Phys. Lett. **157**, 479 (1989).
- ¹³³ J. F. Stanton, Chem. Phys. Lett. **281**, 130 (1997).
- ¹³⁴ T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).
- ¹³⁵ D. E. Woon and T. H. Dunning, J. Chem. Phys. **103**, 4572 (1995).
- ¹³⁶ K. L. Bak, J. Gauss, P. Jørgensen, J. Olsen, T. Helgaker, and J. F. Stanton, J. Chem. Phys. **114**, 6548 (2001).
- ¹³⁷ K. P. Huber and G. H. Herzberg, in *Constants of Diatomic Molecules* (Van Nostrand–Reinhold, New York, 1979).
- ¹³⁸ G. Pelz, K. Yamada, and G. Winnewisser, J. Mol. Spectrosc. **159**, 507 (1993).
- ¹³⁹ T. Helgaker, J. Gauss, P. Jørgensen, and J. Olsen, J. Chem. Phys. **106**, 6430 (1997).
- ¹⁴⁰ B. Ramachandran, N. S. Vegesna, and K. A. Peterson, J. Phys. Chem. A **107**, 7938 (2003).
- ¹⁴¹ R. Shepard, G. S. Kedziora, H. Lischka, I. Shavitt, T. Müller, P. G. Szalay, M. Kállay, and M. Seth, Chem. Phys. **349**, 37 (2008).

- ¹⁴² D. S. Hollman and H. F. Schaefer, J. Chem. Phys. **136**, 084302 (2012).
- ¹⁴³ W. Kutzelnigg, in *Recent Progress in Coupled Cluster Methods*, edited by P. Čársky, J. Paldus, and J. Pittner (Springer Science+Business Media, 2010), pp. 299–356.
- ¹⁴⁴ D. A. Mazziotti, Int. J. Quantum Chem. **70**, 557 (1998).
- ¹⁴⁵ W. Kutzelnigg and D. Mukherjee, Chem. Phys. Lett. **317**, 567 (2000).
- ¹⁴⁶ A. E. DePrince and D. A. Mazziotti, J. Chem. Phys. **127**, 104104 (2007).
- ¹⁴⁷ D. A. Mazziotti, Phys. Rev. Lett. **97**, 143002 (2006).
- ¹⁴⁸ D. A. Mazziotti, J. Chem. Phys. **126**, 184101 (2007).
- ¹⁴⁹ D. A. Mazziotti, Phys. Rev. A **75**, 022505 (2007).
- ¹⁵⁰ C. Valdemoro, L. M. Tel, D. R. Alcoba, and E. Pérez-Romero, Theor. Chem. Acc. **118**, 503 (2007).
- ¹⁵¹ E. Neuscamman, T. T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **130**, 124102 (2009).
- ¹⁵² T. T. Yanai, Y. Kurashige, E. Neuscamman, and G. K.-L. Chan, J. Chem. Phys. **132**, 024105 (2010).
- ¹⁵³ E. Neuscamman, T. T. Yanai, and G. K.-L. Chan, J. Chem. Phys. **132**, 024106 (2010).
- ¹⁵⁴ M. Torheyden and E. F. Valeev, J. Chem. Phys. **131**, 171103 (2009).
- ¹⁵⁵ L. Kong, K. R. Shamasundar, O. Demel, and M. Nooijen, J. Chem. Phys. **130**, 114101 (2009).
- ¹⁵⁶ L. Kong and E. F. Valeev, J. Chem. Phys. **133**, 174126 (2010).
- ¹⁵⁷ T. Shiozaki, G. Knizia, and H.-J. Werner, J. Chem. Phys. **134**, 034113 (2011).

- ¹⁵⁸ D. Datta, L. Kong, and M. Nooijen, J. Chem. Phys. **134**, 214116 (2011).
- ¹⁵⁹ L. Kong and E. F. Valeev, J. Chem. Phys. **135**, 214105 (2011).
- ¹⁶⁰ M. Hanauer and A. Köhn, J. Chem. Phys. **137**, 131103 (2012).
- ¹⁶¹ A. Y. Sokolov, J. J. Wilke, A. C. Simmonett, and H. F. Schaefer, J. Chem. Phys. **137**, 054105 (2012).
- ¹⁶² W. Meyer, J. Chem. Phys. **58**, 1017 (1973).
- ¹⁶³ S. Koch and W. Kutzelnigg, Theor. Chem. Acc. **59**, 387 (1981).
- ¹⁶⁴ U. Bozkaya, J. M. Turney, Y. Yamaguchi, H. F. Schaefer, and C. D. Sherrill, J. Chem. Phys. **135**, 104103 (2011).
- ¹⁶⁵ U. Bozkaya, J. Chem. Phys. **135**, 224103 (2011).
- ¹⁶⁶ W. Kutzelnigg, Chem. Phys. **401**, 119 (2012).
- ¹⁶⁷ W. D. Laidig, P. Saxe, and R. J. Bartlett, J. Chem. Phys. **86**, 887 (1987).
- ¹⁶⁸ M. O. Sinnokrot and C. D. Sherrill, J. Phys. Chem. A **110**, 10656 (2006).
- ¹⁶⁹ W. Kutzelnigg, Theor. Chem. Acc. **80**, 349 (1991).
- ¹⁷⁰ W. Kutzelnigg, Mol. Phys. **94**, 65 (1998).
- ¹⁷¹ P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).
- ¹⁷² B. Cooper and P. J. Knowles, J. Chem. Phys. **133**, 234102 (2010).
- ¹⁷³ F. A. Evangelista, J. Chem. Phys. **134**, 224102 (2011).
- ¹⁷⁴ A. Y. Sokolov, A. C. Simmonett, and H. F. Schaefer, J. Chem. Phys. **138**, 024107 (2013).

- ¹⁷⁵ K. A. Brueckner, Phys. Rev. **96**, 508 (1954).
- ¹⁷⁶ R. K. Nesbet, Phys. Rev. **109**, 1632 (1958).
- ¹⁷⁷ G. E. Scuseria and H. F. Schaefer, Chem. Phys. Lett. **142**, 354 (1987).
- ¹⁷⁸ N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).
- ¹⁷⁹ K. Raghavachari, J. A. Pople, E. S. Replogle, M. Head-Gordon, and N. C. Handy, Chem. Phys. Lett. **167**, 115 (1990).
- ¹⁸⁰ G. E. Scuseria, Chem. Phys. Lett. **226**, 251 (1994).
- ¹⁸¹ C. D. Sherrill, A. Krylov, E. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 4171 (1998).
- ¹⁸² A. Krylov, C. D. Sherrill, E. F. C. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 10669 (1998).
- ¹⁸³ U. Bozkaya and C. D. Sherrill, J. Chem. Phys. **138**, 184103 (2013).
- ¹⁸⁴ U. Bozkaya, J. Chem. Phys. **139**, 104116 (2013).
- ¹⁸⁵ U. Bozkaya and C. D. Sherrill, J. Chem. Phys. **139**, 054104 (2013).
- ¹⁸⁶ U. Bozkaya and H. F. Schaefer, J. Chem. Phys. **136**, 204114 (2012).
- ¹⁸⁷ H. Primas, Rev. Mod. Phys. **35**, 710 (1963).
- ¹⁸⁸ D. J. Klein, J. Chem. Phys. **61**, 786 (1974).
- ¹⁸⁹ J. Linderberg and Y. Öhrn, Int. J. Quant. Chem. **12**, 161 (1977).
- ¹⁹⁰ E. Dalgaard and P. Jørgensen, J. Chem. Phys. **69**, 3833 (1978).

- ¹⁹¹ B. Kirtman, J. Chem. Phys. **75**, 798 (1981).
- ¹⁹² R. Shepard, Adv. Chem. Phys. **69**, 63 (1987).
- ¹⁹³ W. Kohn and L. J. Sham, Phys. Rev. A **140**, 1133 (1965).
- ¹⁹⁴ W. Kutzelnigg, Chem. Phys. Lett. **64**, 383 (1979).
- ¹⁹⁵ W. Kutzelnigg, Int. J. Quantum Chem. **18**, 3 (1980).
- ¹⁹⁶ R. Tuckett, A. Dale, D. Jaffey, P. Jarrett, and T. Kelly, Mol. Phys. **49**, 475 (1983).
- ¹⁹⁷ D. T. Cramb, M. C. L. Gerry, F. W. Dalby, and I. Ozier, Chem. Phys. Lett. **178**, 115 (1991).
- ¹⁹⁸ C. E. Miller and B. J. Drouin, J. Mol. Spectrosc. **205**, 312 (2001).
- ¹⁹⁹ E. F. C. Byrd, C. D. Sherrill, and M. Head-Gordon, J. Phys. Chem. A **105**, 9736 (2001).
- ²⁰⁰ G. J. O. Beran, S. R. Gwaltney, and M. Head-Gordon, Phys. Chem. Chem. Phys. **5**, 2488 (2003).
- ²⁰¹ R. C. Lochan and M. Head-Gordon, J. Chem. Phys. **126**, 164101 (2007).
- ²⁰² W. Kurlancheek and M. Head-Gordon, Mol. Phys. **107**, 1223 (2009).