# Explicitly correlated Green's function methods for calculating electron binding energies

Nakul Kushabhau Teke

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

Master of Science

in

Chemistry

Eduard Valeyev, Chair

Daniel Crawford

Diego Troya

Nicholas Mayhall

April 17, 2019

Blacksburg, Virginia

Keywords: Green's function, explicit correlation, ionization potential, electron affinity

# Explicitly correlated Green's function methods for calculating electron binding energies

Nakul Kushabhau Teke

(ABSTRACT)

Single-particle Green's function method is a direct way of calculating electron binding energy, which relies on expanding the Fock subspace in a finite single-particle basis. However, these methods suffer from slow asymptotic decay of basis set incompleteness error. An energy-dependent explicitly correlated (F12) formalism for Green's function is presented that achieves faster convergence to the basis set limit. The renormalized second-order Green's function method (NR2-F12) scales as *iterative* $\mathcal{O}(N^5)$ where $N$ is the system size. These methods are tested on small (O21) and medium sized (OAM24) organic molecules. The basis set incompleteness error in ionization potential (IP) obtained from the NR2-F12 method and aug-cc-pVDZ basis for OAM24 is 0.033 eV compared to 0.067 eV for NR2 method and aug-cc-pVQZ basis. Hence, accurate electron binding energies can be calculated at a lower cost using NR2-F12 method. For aug-cc-pVDZ basis, the electron binding energies obtained from NR2-F12 are comparable to EOM-IP-CCSD method that uses a CCSD reference and scales as *iterative* $\mathcal{O}(N^6)$.

This work received partial support from the U.S. National Science Foundation awards 1550456 and 1800348.

# Explicitly correlated Green's function methods for calculating electron binding energies

Nakul Kushabhau Teke

(GENERAL ABSTRACT)

Solving the non-relativistic time-independent Schrödinger equation is a central problem in quantum chemistry with the primary goal of finding the exact electronic wave function. Like all many-body problems, the applications of highly accurate electronic structure methods are limited to small molecules since they are computationally expensive. With scalable algorithms and parallel implementation of computer programs, the chemistry of large molecular systems can be investigated. Electron binding energies give an insight into the orbital picture of a molecule, which is manifested in chemical structure and properties of a molecule.

Green's function provides an alternative to wave function based methods to calculate ionization potential and electron affinity directly rather than solving for the wave function itself. For accurate electron binding energies, the wave function needs to be represented by large number of basis functions, which make these methods computationally expensive.

Explicitly correlated electronic structure methods are designed to produce accurate results at a smaller basis set. This work investigates the use of explicitly correlated Green's function methods to calculate electron binding energies of small and medium sized organic molecules. These results are compared to coupled cluster methods, which are known to provide accurate benchmarks in quantum chemistry.

# Acknowledgments

I thank Prof. Edward F. Valeev for taking me as a protégé and being an excellent advisor. Prof. T. Daniel Crawford, Prof. Diego Troya and Prof. Nicholas Mayhall for their insights in this work. My family and friends for their unconditional support.

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

With the turn of twenty-first century, the premise of chemical problems has expanded to focus on determination of structure, function and dynamics of biological molecules,[2–4] drug-design,[5–7] development of semiconductors for renewable energy and storage[8–11] and so on. With high performance computing at hand, the experimental probes to these problems are often augmented with computational studies giving us insights into the molecular structure, thereby refining our understanding of the material world. Meanwhile, machine learning for predicting molecular properties,[12–16] quantum information,[17] and quantum computing[18,19] have seen tremendous growth recently.

Qunatum chemistry deals with the design and implementation of efficient algorithms of electronic structure methods that are scalable and accurate. In this direction, solving the non-relativistic time-independent Schrödinger equation for molecular systems is a central problem. However, this is a many-body problem which can be solved exactly only for a

Hydrogen atom or Hydrogen like atoms that consists of an electron and a nucleus.

The many-electron problem can be solved approximately by density functional theory (DFT) or wave function methods. DFT is a popular choice due to its low cost of computation and is routinely employed to study large molecular systems.[20] Despite this advantage, the errors in DFT computations need to be determined statistically; requiring large datasets for benchmarks. Also, DFT fails to model non-covalent interaction[21] and gives substantial errors for charge-transfer excited states.[22] These deficiencies of DFT models can be attributed to delocalization errors and static correlation errors that are introduced due to inadequacies in the exchange-correlation functional.[23] On the contrary, these problems can be handled in a systematic way by wave function based methods. However, they are computationally expensive and despite being scalable, their applications are limited to medium sized molecular systems even with fragmentation and reduced scaling approaches.[24–27] Wave function methods typically provide chemically accurate results and are often used to generate accurate benchmarks.[28–31] This work is focused on wave function based electronic structure methods and their applications to calculate electron binding energy.

This chapter is organized as follows. The time-independent non-relativistic Schrödinger equation and the Born-Oppenheimer approximation are introduced in section 1.1. Mean-field methods focusing on Hartree-Fock theory, its strengths and limitations are discussed in section 1.2. The fundamental electron-correlation methods like configuration interaction (CI), perturbation theory (PT), coupled cluster (CC) theory, equation of motion (EOM) CC methods are derived in section 1.3, which is followed by a summary.

## 1.1 The Schrödinger equation

A stationary state of a quantum mechanical system is specified by a wave function that satisfies the time-independent non-relativistic Schrödinger equation:

$$\hat{H}\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{1.1.1}$$

This is an eigenvalue problem, where $\hat{H}$ is the molecular Hamiltonian operator, $|\Psi\rangle$ and $E$ are eigen functions and eigen values of the Hamiltonian operator. The Hamiltonian operator consists of kinetic energy operators for electrons and nuclei, a Coulomb attraction term for electron-nucleus pairs and a Coulomb repulsion terms for both, electron and nuclei pairs. This eigenvalue problem is difficult to solve due to the coupling between electronic and nuclear wave functions.

The problem is simplified by employing the Born-Oppenheimer (BO) also known as the clamped nuclei approximation.[32] Since nuclei are approximately three orders of magnitude heavier than the electrons, they move slowly relative to electrons. Hence, it is a reasonable approximation to separate the electronic and nuclear wave function. The positions of electrons have a parametric dependence on the positions of nuclei. Hence, for a fixed nuclear configuration, we can focus on optimizing the electronic wave function. Adiabatic BO approximation is justified for many chemical reactions. With the BO approximation, the problem reduces to optimizing the electronic wave function and the electronic Hamiltonian is given by:

$$\hat{H}_{\text{elec}} = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} + \sum_{i=1}^{N}\sum_{i>j}^{N}\frac{1}{r_{ij}}, \tag{1.1.2}$$

where the terms refer to the kinetic energy of electrons, the electron-nuclear attraction and electron-electron repulsion term, respectively for $N$ electrons and $M$ nuclei. However,

certain problems need the coupling between electronic and nuclear wave functions. Conical intersections in potential energy surfaces that explain radiationless decay of electronic states is one such example of coupling of electronic and nuclear wave funtion.[33] BO cannot be used for metals and zero-band gap semiconductors since the energy differences in their electronic states are zero.[34] The BO approximation is implicitly assumed for the following methods of determining the electronic wave function.

## 1.2 Hartree-Fock theory

Starting from the wave mechanics of Schrödinger,[35] Hartree introduced wave mechanics for atom,[36] which was the first step towards solving the non-relativistic electronic Schrödinger equation approximately. Hartree represented the total wave function as a product of independent single-particle functions (orbitals), which is popularly known as the Hartree product:

$$\Psi_{HP}(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = \chi_1(\boldsymbol{x}_1)\chi_2(\boldsymbol{x}_2)\ldots\chi_N(\boldsymbol{x}_N). \tag{1.2.1}$$

In the above equation, $\chi_i(\boldsymbol{r}_i)$ is a spin-orbital, a product of spatial and spin function, $\chi_i(\boldsymbol{x}_i) = \phi(\boldsymbol{r}_i)\omega(i)$. There is an implicit assumption that electrons are distinguishable in the Hartree product; also it does not account for the *antisymmetry principle* or correlation, since the electrons are assumed to be independent.

Fundamentally, electrons are elementary particles with half-integral spins (*fermions*),[37] and are constrained by the Pauli's exclusion principle.[38] A wave function representing fermions must be antisymmetric, i.e., the function must change sign when two particles are interchanged. Also, the anti-symmetry property with the exchange of columns is intrinsic to determinants, so it can be introduced in the electronic wave function by a single reference

wave function called the Slater determinant,[39] given by:

$$\Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{x}_1) & \chi_1(\boldsymbol{x}_2) & \ldots & \chi_1(\boldsymbol{x}_N) \\ \chi_2(\boldsymbol{x}_1) & \chi_2(\boldsymbol{x}_2) & \ldots & \chi_2(\boldsymbol{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(\boldsymbol{x}_1) & \chi_N(\boldsymbol{x}_2) & \ldots & \chi_N(\boldsymbol{x}_N) \end{vmatrix}. \tag{1.2.2}$$

In a shorthand notation, the presence of a normalization factor of $(\sqrt{N!})^{-1}$ is implied and the Slater determinant is written as: $\Psi = |\chi_i \chi_j \ldots \chi_k\rangle$, where electron $i$ is assumed to be in the corresponding spin orbital $\chi_i$. This antisymmetrized product (Slater determinant) accounts for the Coulomb repulsion between an electron and the *average field* of the rest of electrons in the system. Hence, Hartree-Fock theory is often referred to as a mean-field theory. With the introduction of a Slater determinant for electronic wave function, the problem reduces to optimizing $\chi_i$ so that they remain orthonormal to each other and minimize the Raleigh quotient:

$$E_{\text{elec}} = \frac{\langle \Psi | \hat{H}_{\text{elec}} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \tag{1.2.3}$$

According to the variational principle, the $|\Psi\rangle$ that minimizes the electronic energy ($E_{\text{elec}}$) is the best approximation to $|\Psi\rangle_{\text{exact}}$. Hence, the energy obtained from variational optimization of Hartree-Fock wave function is an upper-bound for the exact energy of the system. The electronic energy obtained from a Hartree-Fock self-consistent field (SCF) procedure in Dirac notation is given by:

$$E_{\text{HF}} = \sum_i \langle i | h | i \rangle + \frac{1}{2} \sum_{ij} \langle ij | | ij \rangle, \tag{1.2.4}$$

where $i$, $j$ correspond to occupied orbitals of the orbital basis set. In the above equation, first-term is given by:

$$\langle i | h | j \rangle = \int d\boldsymbol{x}_1 \chi_i^*(\boldsymbol{x}_1) h(\boldsymbol{r}_1) \chi_j(\boldsymbol{x}_1), \tag{1.2.5}$$

where $h(\boldsymbol{r}_1)$ constitutes the kinetic energy of single electron and the potential energy of its attaction to nuclei. These second-term in Eq. (1.2.4) is described in physicist notation and expanded as:

$$\begin{aligned} \langle ij||kl\rangle &= \langle ij|kl\rangle - \langle ij|lk\rangle\,, \\ \langle ij|kl\rangle &= \int d\boldsymbol{x}_1 d\boldsymbol{x}_2 \chi_i^*(\boldsymbol{x}_1)\chi_j^*(\boldsymbol{x}_2)\frac{1}{\boldsymbol{r}_{12}}\chi_k(\boldsymbol{x}_1)\chi_l(\boldsymbol{x}_2). \end{aligned} \tag{1.2.6}$$

To evaluate the energies, the Hartree-Fock equation is formulated as an eigenvalue problem with the Coulomb ($J$) and exchange operators ($K$) given by:

$$J(1)\chi_i(1) = \sum_j \int d\boldsymbol{r}_2 \chi_j^*(2)\frac{1}{r_{12}}\chi_j(2)\chi_i(1) \tag{1.2.7}$$

and

$$K(1)\chi_i(1) = \sum_{ij}\left[\int d\boldsymbol{r}_2 \chi_i^*(2)\frac{1}{r_{12}}\chi_j(2)\right]\chi_i(1); \tag{1.2.8}$$

and the *Fock operator* ($\hat{F}$) is given by:

$$\hat{F} = \hat{h}(1) + \sum_j \left(\hat{J}_j(1) - \hat{K}_j(1)\right), \tag{1.2.9}$$

where $\hat{h}(1)$ is the one-electron part of Hamiltonian that describes its kinetic energy and potential energy of attraction to nuclei. In Hartree-Fock theory, every electron is represented by a one-particle function, which is an atomic orbital for an atom. A linear combination of these single particle functions is used to construct molecular orbitals (MOs), which (for a mean-field method) are a reasonable approximation of the electronic structure of molecules in their equilibrium geometry and ground state.

For a closed-shell system, the spatial molecular orbitals are doubly occupied and hence the spin orbitals have the same number of $m_s = +\frac{1}{2}$ and $m_s = -\frac{1}{2}$ spin functions. To reduce the

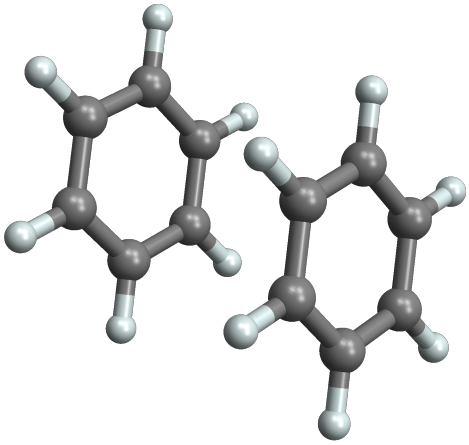


Figure 1.1: Benzene dimer in a $\pi$ stacked geometry from the S66 dataset.

computational effort, the problem is often transformed to a spatial eigenvalue equation. The exchange term exists for electrons of parallel spin and vanishes for anti-parallel spins due to spin orthogonality. Hence, for a closed shell systems, the Fock operator can be written in spin-free formalism and the Hartree-Fock energy for a closed-shell system is:

$$E_0 = 2\sum_i \langle i|\, h\, |i\rangle + \sum_{ij} \left(2\,\langle ij|ij\rangle - \langle ij|ji\rangle\right) \tag{1.2.10}$$

In a matrix form, the eigenvalue problem with the Fock ($\mathbf{F}$) and overlap ($\mathbf{S}$) matrices is formulated as:

$$\mathbf{FC} = \mathbf{SC}\boldsymbol{\epsilon}. \tag{1.2.11}$$

This eigenvalue problem is solved iteratively until self-consistency and yields the canonical orbitals ($\mathbf{C}$) and orbital energies $\boldsymbol{\epsilon}$. The total non-relativistic energy of a molecule is obtained by adding the nuclear repulsion energy to the electronic energy.

Depending on the nature of problem under investigation, Hartree-Fock approximations with

different restrictions on spin-orbitals are used. The unrestricted Hartree-Fock (*UHF*) model is used to describe bond dissociation and open-shell molecules. The restricted Hartree-Fock (*RHF*) assumes equal number electrons with $\alpha$ and $\beta$ spin and the same spatial component of the wave function (as described earlier). This is a good approximation for closed-shell molecules at an equilibrium geometry. Restricted open-shell Hartree-Fock (*ROHF*) uses the same spatial orbitals for electrons with $\alpha$ and $\beta$ spins and is designed for open-shell molecules.

Mean-field methods recover about 99% of the total electronic energy and 95% of the electronic wave function.[40] Hartree-Fock cannot be used as a stand alone theory to accurately predict chemical and spectroscopic properties since the energy differences associated with these processes are less than the error introduced by the method itself. For instance, the sequence of $\pi_{2p}$ and $\sigma_{2p}$ orbitals based on energies is fallacious for $N_2$ at Hartree-Fock level of theory. Another example is a benzene dimer. The counterpoise (CP) corrected binding energy of benzene dimer (Figure 1.1) from the S66 dataset[31] calculated using Hartree-Fock method and aug-cc-pVDZ basis set is 3.99 kcal/mol. A positive binding energy indicates that the dimer is unstable, but it is known to be stable experimentally. Hence, the binding energy is incorrect due to lack of proper treatment of electron correlation in the mean-field description of the electronic wave function. The MP2 method with aug-cc-pVDZ basis gives the CP corrected binding energy of -4.16 kcal/mol. Although the MP2 method overestimates this value, it succeeds in predicting that the complex is bound. For accurate prediction of energies and chemical properties, systematic treatment of electron correlation is essential.

## 1.3 Electron correlation methods

Electron correlation is a many-body problem in physical sense, where the 'bodies' are electrons. Most effort in electronic structure theory over the last half-century has been put into

studying and improving the description of electron correlation in molecules. Löwdin defined the electron correlation energy ($E_{corr}$) as the difference between the exact non-relativistic Born-Oppenheimer energy ($E_{nr}$) and the *exact* Hartree-Fock energy ($E_{\mathrm{HF}}$) calculated using a complete basis set:[41]

$$E_{corr} = E_{nr} - E_{\mathrm{HF}}. \tag{1.3.1}$$

However, it has beed argued that this definition is flawed since $E_{nr}$ and $E_{\mathrm{HF}}$ are not defined precisely and the choice of Hartree-Fock approximation is non trivial.[42] The Hartree-Fock wave function $|\Psi\rangle_{\mathrm{HF}}$ is a starting point for single reference wave function methods and accounts for *Fermi correlation*, i.e., the motion of electrons with *like* spin is correlated. The correlation energy is, strictly speaking, the Coulomb correlation energy.[42] The electron correlation problem and the short-falls of conventional electronic structure methods are discussed in detail in chapter 3. The methods outlined below are designed to recover the Coulomb correlation energy ($E_{\mathrm{corr}}$) that is missing from the Hartree-Fock calculation.

### 1.3.1 Configuration interaction

Configuration Interaction (CI)[43] is one of the earliest methods developed to address the electron correlation. It is a variational method to solve the non-relativistic Schrödinger equation within the BO approximation and is popular since it can be applied to excited states and open shell-molecules.[44–47] Increasing the span of our wave function improves the correlation contribution to the electronic energy. The full form of the CI wave function is written in terms of excited state determinants obtained from a reference wave function $|\phi_0\rangle$ as:

$$|\Psi\rangle_{\mathrm{CI}} = c_0 |\phi_0\rangle + \sum_{ia} c_i^a |\phi_i^a\rangle + \sum_{\substack{i<j \\ a<b}} c_{ij}^{ab} |\phi_{ij}^{ab}\rangle + \sum_{\substack{i<j<k \\ a<b<c}} c_{ijk}^{abc} |\phi_{ijk}^{abc}\rangle + \ldots, \tag{1.3.2}$$

where the terms on the right hand side in Eq. (1.3.2) are the reference state, and singly, doubly, triply substituted determinants, respectively. In this equation, $i, j, k$ refer to occupied (holes) orbitals of the orbital basis set (OBS) and $a, b, c$ are unoccupied (particles) of the OBS. For an $n$ electron system, including all excited state determinants up to $n$-tuple excitations constitute the Full CI (FCI) method. The summation restriction ensures that each configuration is included only once and Slater-Condon rules can be used to simplify the matrix elements in the above expansion.[48,49] For a $2K$ set of spin orbitals with $N$ orbitals occupied, the number of $n$-tuply excited determinants is $\binom{N}{n}\binom{2K-N}{n}$. This number grows rapidly with $N$ (the system size), hence the application of CI is often limited to small molecular systems.

Several approximations to FCI are made since FCI cannot used for large molecules. The contributions of singly-excited determinants for ground state energy is negligible compared to doubly-excited determinants, and a CI wave function $|\Psi\rangle_{\text{CID}}$ with just the doubles can be used. This is the CID method, where the excitations are truncated at doubles. The CIS method that has singly-excited determinants is useful for calculating excited state energies. CI is not size-consistent due to the truncated form of its wave function. Size consistency ensures the energy of non-interacting particles is equal to the sum of the their energies calculated separately. Also, CI is not size-extensive, which means the energy of the system does not scale correctly with the number of electrons. Alternatively, coupled cluster (CC) method is size-consistent and size-extensive due to the inclusion of disconnected clusters; and is a method of choice when numerically accurate results are required.

### 1.3.2 Perturbation theory

Many-body perturbation theories (MBPT) were developed for applications in nuclear physics and later applied to solid-state physics. Despite its origins from Quantum Field Theory, the inception of MBPT for quantum chemistry raised skipticism since it was not compatible with the variational principle. MBPT was first seen in context of *ab initio* quantum chemistry in works of Freed.[50,51] It integrated successfully in quantum chemistry in the later decades. MBPT is reviewed extensively in quantum chemistry literature.[40,52] In MBPT, the electron correlation is treated as a perturbation to the Hartree-Fock wave function. Consider a perturbed Hamiltonian $\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)}$. The eigen functions and eigen values of a Hamiltonian $\hat{H}^{(0)}$ are calculated using Hartree-Fock method. For small perturbation, the eigenvalues of $\hat{H}(\lambda)$ must be close to the eigenvalues of $\hat{H}^{(0)}$. In Raleigh-Schrödinger Perturbation Theory (RSPT), the Hamiltonian is partitioned with inclusion of an ordering parameter $\lambda$ as:

$$\hat{H}(\lambda) = \hat{H}^{(0)} + \lambda\hat{H}^{(1)} \tag{1.3.3}$$

The eigen functions and eigen values are written as a series expansion as Eq. (1.3.4) and Eq. (1.3.5) respectively, where $E_i^{(n)}$ is the $n^{\text{th}}$ order energy. The perturbed wave functions are obtained from linear combination of zeroth order wave function.

$$|\Psi_i\rangle = \left|\phi_i^{(0)}\right\rangle + \lambda\left|\phi_i^{(1)}\right\rangle + \lambda^2\left|\phi_i^{(2)}\right\rangle + \ldots \tag{1.3.4}$$

$$E_{\text{MBPT}} = E_i^{(0)} + \lambda E_i^{(1)} + \lambda^2 E_i^{(2)} + \ldots \tag{1.3.5}$$

The corrections to the energy using second-order perturbation theory are popularly known as second-order Møller-Plesset perturbation theory (MP2).[53] In this formalism, the Hartree-Fock energy is expressed as the sum of zeroth and first order energies. The correction to this

energy occurs at the second-order of perturbation and the energy is given by:

$$E_0^{(2)} = \sum_n \frac{|\langle 0|\hat{H}^{(1)}|n\rangle|^2}{E_0^{(0)} - E_0^{(n)}} \tag{1.3.6}$$

In Eq. (1.3.6), the only terms in $|n\rangle$ that couple with $|0\rangle$ are doubly excited determinants and contribute to second-order energy. The second-order correction to energy is given by:

$$E_{\mathrm{MP2}} = \frac{1}{4}\sum_{ijab} \frac{|\langle ij||ab\rangle|^2}{\varepsilon_i + \varepsilon_j - \varepsilon_a - \varepsilon_b}. \tag{1.3.7}$$

Though some correlation is recovered in MP2, the non-truncating *partial wave* expansion leads to the slow convergence of MP2 energy to the basis set limit. As discussed in chapter 3, the basis set error can be dramatically reduced by using the explicitly correlated MP2-F12 method. In terms of percentage of correlation energy recovered, coupled cluster methods perform the best.

### 1.3.3 Coupled-cluster (CC) methods

The genesis of coupled cluster theory is in nuclear structure theory.[54,55] The earliest study of CC theory in chemical context was done by Cižek.[56,57] Over the decades, numerous methodologies and approximations have evolved. The seminal contributions include introduction of CC with singles and doubles (CCSD)[58] and with full triples (CCSDT),[59] equation of motion methods,[60–62] Green's function,[63] multi-reference method,[64] renormalization formalism,[65,66] explicitly-correlated methods,[67–69] reduced-scaling methods,[26,70] etc. Due to their widespread presence in quantum chemistry literature, the coupled-cluster methods are reviewed extensively.[40,71–73]

In CC formalism, the wave function is expressed as an exponential *ansatz* of the cluster

operator, as shown in Eq.1.3.8. The cluster operator, $\hat{T}$ is a linear combination of the product of excitation operators $\hat{\tau}_\mu$ and cluster amplitudes $t_\mu$.

$$|\Psi\rangle_{\text{CC}} = e^{\hat{T}} |\Psi\rangle_{HF} \qquad \hat{T} = \sum_{\mu} t_\mu \hat{\tau}_\mu \tag{1.3.8}$$

$$e^{\hat{T}} = 1 + \hat{T} + \frac{1}{2!}\hat{T}^2 + \frac{1}{3!}\hat{T}^3 + \ldots + \frac{1}{n!}\hat{T}^n + \ldots \tag{1.3.9}$$

Unlike CI method, the exponential *ansatz* includes disconnected terms, which include the effects of higher order excitations in the wave function without the need for calculating high order amplitudes. This makes CC methods computationally robust compared to the traditional CI methods.

The cluster operator is truncated for computational viability. Despite having a truncated wave function, the CC method is size-consistent. The wave function obtained by truncation of $\hat{T}$ at doubles is $|\Psi\rangle_{\text{CCSD}}$ given by Eq. (1.3.10) and the coupled cluster energy is obtained by projection onto a Hartree-Fock state $\langle\Psi_0|$ as Eq. (1.3.11).

$$|\Psi\rangle_{\text{CCSD}} = e^{\hat{T}_1 + \hat{T}_2} |\Psi_0\rangle \tag{1.3.10}$$

$$E_{\text{CCSD}} = \langle\Psi_0| \bar{H} |\Psi_0\rangle , \tag{1.3.11}$$

where, $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$ is the similarity transformed Hamiltonian. An $n$-electron system can have a maximum of $n$ excitations. Hence, inclusion of all terms up to $\hat{T}_n$ will correspond to a full CI wave function in the given MO basis. The $\hat{T}$ amplitudes are determined from the following amplitude equations. For instance, $\hat{T}_1$ are determined from Eq. (1.3.12) while $\hat{T}_2$ are determined from Eq. (1.3.13).

$$\langle\psi_i^a| \bar{H} |\psi_0\rangle = 0 \tag{1.3.12}$$

$$\left\langle \psi_{ij}^{ab} \right| \bar{H} \left| \psi_0 \right\rangle = 0 \tag{1.3.13}$$

The errors in reaction energies obtained from the coupled cluster methods are of the order of 1 kcal $\text{mol}^{-1}$, which is equivalent to the accuracy of such experiments.[74] This number is commonly referred to as chemical accuracy in computational chemistry. In practice, the wave function is restricted to doubles (CCD), single and double excitations (CCSD), and single, double and triple excitations(CCSDT). CCSD models were formulated by Bartlett and Purvis.[75,76] Coupled cluster singles and doubles with perturbative triples, CCSD(T) formulated by Raghavachari *et al.* is colloquially said to be the gold standard in quantum chemistry due to its superior accuracy without formulating the triples amplitudes explicitly;[77] it was recently used to study non-covalent interactions by Řezáč and Hobza.[78] The CCSDT model was formulated and implemented by Noga and Bartlett.[79] The equation of motion CC methods are formulated using second quantization and allow us to calculate excitation and electron binding energies efficiently.

### 1.3.4 Equation-of-motion (EOM) CC methods

The EOM-CC suite of methods can be used in context of excitation energies (EE), ionization potentials (IP) and electron affinity (EA). Within the EOM formalism, the excited state wave function $|\psi\rangle_{ex}$ is obtained from the ground state by action of a linear (CI-like) excitation operator, $\hat{R}$, which is expressed as a sum of single, doubles, triples and so on.

$$|\psi\rangle_{\text{ex}} = \hat{R} \left| \psi_0 \right\rangle \tag{1.3.14}$$

In the above expression, $|\psi_0\rangle$ is a ground state wave function and $\hat{R}$ can be an excitation (EE), electron detachment (IP) or electron attachment (EA) operator. The expressions for

these operators are given by:

$$\begin{aligned}\hat{R}_{\mathrm{EE}} &= r_0 + r_a^i a_i^a + \frac{1}{4} r_{ab}^{ij} a_{ij}^{ab} + \frac{1}{36} r_{abc}^{ijk} a_{ijk}^{abc} + \dots \\ \hat{R}_{\mathrm{IP}} &= r^i a_i + \frac{1}{2} r_a^{ij} a_{ij}^a + \frac{1}{12} r_{ab}^{ijk} a_{ijk}^{ab} + \dots \\ \hat{R}_{\mathrm{EA}} &= r_a a^a + \frac{1}{2} r_{ab}^i a_i^{ab} + \frac{1}{12} r_{abc}^{ij} a_{ij}^{abc} + \dots\end{aligned} \tag{1.3.15}$$

In the above expressions, the operators are written in generalized tensor notation with an implicit Einstein summation. For practical applications, these operators are truncated at double or triple excitations. The excited state energy can be obtained from ground state $|\phi\rangle$ by the action of $\hat{R}$ operator , i.e.,

$$\hat{H}\hat{R}\,|\psi_0\rangle = E_{\mathrm{ex}}\hat{R}\,|\psi_0\rangle\,. \tag{1.3.16}$$

Multiplying the Schrödinger equation by the excitation operator, we get:

$$\hat{R}\hat{H}\,|\psi_0\rangle = E_{\mathrm{gs}}\hat{R}\,|\psi_0\rangle\,. \tag{1.3.17}$$

The last two equations can be subtracted to yield a commutator with the Hamiltonian and excitation operator:

$$[\hat{H}, \hat{R}]\,|\psi_0\rangle = \omega\hat{R}\,|\psi_0\rangle\,, \tag{1.3.18}$$

where $\omega$ is the difference between the excited and ground state energy ($\omega = E_{\mathrm{ex}} - E_{\mathrm{gs}}$). Since we obtained the $|\psi_0\rangle$ as $|\psi_0\rangle = e^{\hat{T}}\,|\phi_0\rangle$, we can write:

$$[\hat{H}, \hat{R}]e^{\hat{T}}\,|\phi_0\rangle = \omega\hat{R}e^{\hat{T}}\,|\phi_0\rangle\,. \tag{1.3.19}$$

By multiplying the above equation by $e^{-\hat{T}}$ on the left, we obtain this equation in terms of the similarity transformed Hamiltonian ($\bar{H}$):

$$[\bar{H}, \hat{R}] \left|\phi_0\right\rangle = \omega \hat{R} \left|\phi_0\right\rangle . \tag{1.3.20}$$

In the matrix form, this problem is represented in terms of CI-like eigen value problem in singles, doubles and triples subspace.

$$\bar{\mathbf{H}} = \begin{bmatrix} \langle \mathbf{S}| \bar{\mathbf{H}} |\mathbf{S}\rangle & \langle \mathbf{S}| \bar{\mathbf{H}} |\mathbf{D}\rangle & \langle \mathbf{S}| \bar{\mathbf{H}} |\mathbf{T}\rangle \\ \langle \mathbf{D}| \bar{\mathbf{H}} |\mathbf{S}\rangle & \langle \mathbf{D}| \bar{\mathbf{H}} |\mathbf{D}\rangle & \langle \mathbf{D}| \bar{\mathbf{H}} |\mathbf{T}\rangle \\ \langle \mathbf{T}| \bar{\mathbf{H}} |\mathbf{S}\rangle & \langle \mathbf{T}| \bar{\mathbf{H}} |\mathbf{D}\rangle & \langle \mathbf{T}| \bar{\mathbf{H}} |\mathbf{T}\rangle \end{bmatrix} . \tag{1.3.21}$$

The matrix $\bar{\mathbf{H}}$ from the above equation is non-Hermitian, which has different left- and right-eigenvectors but the same eigenvalues. Diagonalization of $\bar{\mathbf{H}}$ yields the eigen vectors and energies of the desired states.

## 1.4 Summary

Wave function based methods provide a hierarchy of methods that are systematically improvable. These methods can be utilized based on accuracy demands, computational resources at disposal and the nature of the chemical problem. Systems with a few hundred atoms can be studied with Hartree-Fock (HF) and density functional theory (DFT) with a reasonable accuracy. The most computationally expensive step in HF is evaluation of the two electron integrals, which scale as $\mathcal{O}(N^4)$ with system size. Linear scaling computation of Hartree-Fock exchange for insulating systems was done recently.[80,81]

Canonical MP2 scales as $\mathcal{O}(N^5)$ with system size. Low-order scaling methods like linear scale

local MP2 method exploit orbital domains for electron pairs.[82] Linear scaling domain based local pair natural orbital (DLPNO) MP2 and the explicitly-correlated DLPNO-MP2-F12 form have been implemented recently.[25,27]

Coupled Cluster methods and its derivatives are used for high accuracy. CCSD, CCSD(T) and CCSDT scale as $\mathcal{O}(N^6)$, $\mathcal{O}(N^7)$ and $\mathcal{O}(N^8)$, respectively. In principle, coupled cluster methods with higher excitations can be formulated but the steep scaling with system size makes them impractical for realistic molecular systems. These methods lay a foundation for complex approximations designed for specific problems. For example, linear scaling DLPNO coupled cluster method and the explicitly-correlated DLPNO CCSD(T)$_{\overline{\text{F12}}}$ methods are designed for large systems.[83,84] Equation-of-motion coupled cluster (EOM-CC) methods are useful to calculate excitation energies, ionization potential and electron affinity. However, their requirement of a coupled cluster reference makes them feasible for small to medium sized molecules. Green's function methods, on the other hand, use a Hartree-Fock reference and low-order methods (in perturbation sense) can be used to study large molecules efficiently. Single particle Green's function methods for electron binding energies are discussed in the following chapter.

# Chapter 2

# Single Particle Green's Function

Electron propagators (EP) or Green's function (GF) methods are employed as an alternative way of solving the Schrödinger equation. Favorable scaling and the ability to eliminate the uncoupled terms in perturbation led to their wide presence in many-body theory. Electron propagators are a systematic way to determine vertical electron binding energies, photoionization intensity and electron scattering amplitudes, one-electron properties and total energies. EP methods have been used in electronic struncture studies over past few decades.[85–88] Recently, EP has been applied to study quantum mechanical system at finite temperatures,[89,90] finite systems,[91,92] molecular systems and their spectroscopies,[93,94] embedding theories,[95–97] electron correlation,[98–100] and much more.

This chapter summarizes GF formalism in context of electron binding energies and is organized as follows. Section 2.1 introduces the Koopmans' theorem and its pitfalls, the superoperator notation is briefly discussed in section 2.2, the Dyson equation and self-energy terms are explained in sections 2.3 and 2.4, respectively.

## 2.1 Koopmans' theorem

Koopmans' theorem states that in a closed-shell Hartree-Fock(HF) reference, the first ionization energy is equal to the negative of the highest occupied molecular orbital (HOMO) energy.[101] Similarly, the electron affinity is the negative of the lowest unoccupied molecular orbital (LUMO) energy. Though these are reasonable approximations, they fail for molecules with incorrect orbital states due to mean-field approximation, e.g., the $N_2$ molecule as described in section 1.2.

For a closed-shell molecule, the ionization potential (IP) is defined as the difference in total energies of the cationic and molecular states:

$$\mathrm{IP} = \left\langle {}^{N-1}\Psi_i \right| \hat{H} \left| {}^{N-1}\Psi_i \right\rangle - \left\langle {}^{N}\Psi_0 \right| \hat{H} \left| {}^{N}\Psi_0 \right\rangle \overset{\mathrm{KT}}{=} -\varepsilon_i, \tag{2.1.1}$$

where $\varepsilon_i$ is the orbital energy of HOMO from a HF calculation and $\left|{}^{N-1}\Psi_i\right\rangle$ is an approximation to $N-1$ state generated by removing an electron from orbital $i$ from $|\Psi\rangle$. Similarly, the electron attachment (EA) are given by:

$$\mathrm{EA} = \left\langle {}^{N}\Psi_0 \right| \hat{H} \left| {}^{N}\Psi_0 \right\rangle - \left\langle {}^{N+1}\Psi_a \right| \hat{H} \left| {}^{N+1}\Psi_a \right\rangle \overset{\mathrm{KT}}{=} -\varepsilon_a, \tag{2.1.2}$$

where $\varepsilon_a$ is the LUMO energy and $\left|{}^{N+1}\Psi_a\right\rangle$ is an approximate state generated from $|\Psi\rangle$. These are not HF $N \pm 1$ particle wave functions since they have the spin orbitals of $N$-particle system. In reality, the canonical orbitals of $N$, $N+1$ and $N-1$ systems are not the same. The HF energy of $N-1$ particle system with a missing $i^{\mathrm{th}}$ orbital is:

$$E_{corr}^{N-1}(i) = \langle {}^{N-1}\Psi_i | \hat{H} | {}^{N-1}\Psi_i \rangle + E_R^{N-1}(i). \tag{2.1.3}$$

The last term on the right hand side in the above equation is called *relaxation energy*. The

correlation energies of $N$ and $N-1$ particle systems can be written as:

$$\mathcal{E}_0^N = E_0^N + E_{corr}^N \tag{2.1.4}$$

and

$$\mathcal{E}_0^{N-1} = E_0^{N-1} + E_{corr}^N, \tag{2.1.5}$$

respectively. These expressions need to be evaluated to calculate exact IP given by:

$$IP = \mathcal{E}_i^{N-1} - \mathcal{E}_0^N = -\varepsilon_i + E_R^{N-1}(i) - \left(E_{corr}^N - E_{corr}^{N-1}(i)\right). \tag{2.1.6}$$

Hence the ionization potentials calculated using the Koopmans' theorem needs to be corrected by relaxation energy of $N-1$ particle systems and the difference of correlation energy of $N$ and $N-1$ particle systems. Similar arguments are valid for EA. Green's function formalism provides an elegant way to include these corrections directly post mean-field calculation. This formalism is conveniently expressed in terms of superoperator notation.

## 2.2 Superoperator notation

For electron detachment/attachment, the matrix-element of the electron-propagator matrix $\boldsymbol{G}(E)$ in the spectral form is given by:[102]

$$\begin{aligned} G_{pq}(E) &\equiv \langle\langle a_p^\dagger; a_q \rangle\rangle \\ &= \lim_{\eta\to 0} \left\{ \sum_r \frac{\langle \Psi_0^N | a_p^\dagger | \Psi_r^{N-1} \rangle \langle \Psi_r^{N-1} | a_q | \Psi_0^N \rangle}{E - E_0^N + E_r^{N-1} - i\eta} + \sum_r \frac{\langle \Psi_0^N | a_q | \Psi_r^{N+1} \rangle \langle \Psi_r^{N+1} | a_p^\dagger | \Psi_0^N \rangle}{E - E_r^{N+1} + E_0^N + i\eta} \right\}, \end{aligned} \tag{2.2.1}$$

where, $p$, $q$ refer to the orbital basis set, $|\Psi_0^N\rangle$ is the exact ground state of N-electron system and $|\Psi_r^{N+1}\rangle$, $|\Psi_r^{N-1}\rangle$ are the electron attached and electron detached states, respectively. A generalized propagator that changes the number of electrons by one is defined as:

$$E\langle\langle\mu^\dagger;\nu\rangle\rangle = \langle N,0|[\mu^\dagger,\nu]_+|N,0\rangle + \langle\langle\mu^\dagger;[\nu,H]\rangle\rangle, \tag{2.2.2}$$

where the moment expansion is given by:

$$\langle\langle a_r^\dagger;a_s\rangle\rangle = E^{-1}(a_r|a_s) + E^{-2}(a_r|\hat{H}a_s) + E^{-3}(a_r|\hat{H}^2a_s) + E^{-4}(a_r|\hat{H}^3a_s) + ... \tag{2.2.3}$$

The matrix element of electron propagator can be obtained after formal summation and written in terms of superoperator resolvent:

$$\langle\langle a_r^\dagger;a_s\rangle\rangle \equiv G_{rs}(E) = (a_r|(E\hat{I}-\hat{H})^{-1}a_s). \tag{2.2.4}$$

In the above equation, the action of the identity superoperator, $\hat{I}$, and the superoperator Hamiltonian, $\hat{H}$ are defined by: $\hat{I}X = X$, and $\hat{H}X = [X,H]_-$; and the superoperator metric for field operators products $\mu$ and $\nu$ is defined as:

$$(\mu|\nu) = \langle N|[\mu^\dagger,\nu]_+|N\rangle. \tag{2.2.5}$$

With the above notation, the second-quantized Hamiltonian is defined as:

$$H = \sum_{pq} h_{pq}a_p^\dagger a_q + \frac{1}{4}\sum_{pqrs}\langle pq|\,|rs\rangle\, a_p^\dagger a_q^\dagger a_s a_r \tag{2.2.6}$$

Eq. (2.2.4) can be expressed equivalently in a matrix form as:

$$\mathbf{G}(E) = (\mathbf{a}|(E\hat{I} - \hat{H})^{-1}\mathbf{a}). \tag{2.2.7}$$

In this equation, inner projection is used to avoid the treatment of inverse operator and Eq. (2.2.7) is written in terms of field operator products $\mathbf{u}$ as:

$$\mathbf{G}(E) = (\mathbf{a}|\mathbf{u})(\mathbf{u}|(E\hat{I} - \hat{H}\mathbf{u})^{-1}(\mathbf{u}|\mathbf{a}), \tag{2.2.8}$$

where $\mathbf{u}$ is partitioned in the primary space $\mathbf{a}$ of hole ($h$) and particle ($p$) annihilator operators and an orthogonalized space $\mathbf{f}$ with the $2h1p$ and $2p1h$ counterparts. With the above spaces, the poles of EP occur when values of $E$ are equal to the eigenvalues, $\omega$, of superoperator Hamiltonian $\hat{\mathbf{H}}$ and the eigenvalue problem is formulated as:

$$\hat{\mathbf{H}}\mathbf{U} = \mathbf{U}\omega. \tag{2.2.9}$$

For an exact superoperator Hamiltonian, the poles of EP correspond to the exact ionization potentials and electron affinities. In practice, we solve Eq. (2.2.9) for an approximate superoperator Hamiltonian. For instance, poles for second-order self-energy are recovered from the superoperator Hamiltonian,

$$\hat{\mathbf{H}} = \begin{bmatrix} (\mathbf{a}|\hat{H}\mathbf{a})^{(0)} & (\mathbf{a}|\hat{H}\mathbf{f}_3)^{(1)} \\ (\mathbf{f}_3|\hat{H}\mathbf{a})^{(1)} & (\mathbf{f}_3|\hat{H}\mathbf{f}_3)^{(0)} \end{bmatrix}, \tag{2.2.10}$$

where the subscript denotes the highest order through which the matrix element is evaluated. Rather than diagonalizing the superoperator Hamiltonian, a convenient form of using the single-particle Green's function for calculating electron binding energies is by formulating

the problem in terms of a non-local potential called self-energy ($\mathbf{\Sigma}$) and incorporating it into a Dyson equation.[103,104] In this formulation, the poles and residues are are the eigen values and eigen functions of the Dyson equation, respectively.

## 2.3 Dyson equation

The electron-propagator matrix is given by, $\mathbf{G}(E) = \mathbf{G}_0(E) + \mathbf{G}_0(E)\mathbf{\Sigma}(E)\mathbf{G}(E)$, where $\mathbf{G}_0(E)$ is the zeroth-order Green's function. Alternatively, this equation can be written in an inverse form and poles of the EP correspond to $\mathbf{G}^{-1}(E)$ going to zero. After separating the operator manifolds, Eq. (2.2.8) can be written in a complementary form as an inverse Green's function:

$$\mathbf{G}^{-1}(E) = E1 - (\mathbf{a}|\hat{H}\mathbf{a}) - (\mathbf{a}|\hat{H}\mathbf{f})[E1 - (\mathbf{f}|\hat{H}\mathbf{f})]^{-1}(\mathbf{f}|\hat{H}\mathbf{a}), \tag{2.3.1}$$

where the first two terms constitute the zeroth-order inverse propagator matrix, $\mathbf{G}_0^{-1}(E)$, and the third term is self-energy, $\mathbf{\Sigma}(E)$. Eq. (2.3.1) can be written in a compact form as:

$$\mathbf{G}^{-1}(E) = \mathbf{G}_0^{-1}(E) - \mathbf{\Sigma}(E). \tag{2.3.2}$$

The poles of $\mathbf{G}_0^{-1}(E)$ are equal to canonical Hartree-Fock orbital energies for the unperturbed EP inverse matrix given by:

$$\mathbf{G}_0^{-1}(E) = E\mathbf{1} - \epsilon. \tag{2.3.3}$$

The off-diagonal elements of the self-energy matrix in canonical basis are small and have negligible effect on the poles and Feynman-Dyson amplitudes. The derivation and approximations to the Dyson quasiparticle equation are discussed by Corzo and Ortiz.[103] The eigenvalues, which correspond to the poles of EP are obtained by solving the Dyson equa-

tion self-consistently, which in a matrix form is given by:

$$[\mathbf{F} + \mathbf{\Sigma}(E)]\mathbf{C} = E\mathbf{C} \tag{2.3.4}$$

In this equation, $\mathbf{F}$ is a Fock matrix and $\mathbf{C}$ are Dyson orbitals for cases where $\mathbf{\Sigma}(E)$ has a non-diagonal form. The Dyson equation improves the mean-field approximation by incorporating the correlation effects of both, the molecule and the ionized (or electron attached) state. Eq. (2.3.4) can be corrected *perturbatively* or solved for $E$ in a self-consistent way. The non-diagonal, renormalized approximation of the electron propagator formulated by Ortiz, is a reliable method to calculate ionization potentials among other non-diagonal EP methods.[105,106] The following section summarizes the second-order Green's function (GF2) and the non-digonal renormalized second-order (NR2) approximations to the self-energy.

## 2.4 Self-energy approximations

With Green's function imbibed in the Dyson equation, approximate formulation of the energy-dependent non-local potential called the self-energy, $\mathbf{\Sigma}(E)$, allow us to calculate ionization potential and electron affinity. The self-energy can be obtained from the coupled cluster Green's function (CCGF) approach[63,107,108] and is discussed in Chapter 6.

### 2.4.1 Second-order Green's function (GF2)

In terms of superoperator metric (Eq. 2.2.5), the matrix element of second-order self-energy is given by:

$$\begin{aligned}\mathbf{\Sigma}_{pq}(E) = (a_p|\hat{H}^{(1)}a^b_{kl})(a^a_{ij}|(E-\hat{H}^{(0)})a^b_{kl})^{-1}(a^a_{ij}|\hat{H}^{(1)}a_q) \\ + (a_p|\hat{H}^{(1)}a^j_{cd})(a^i_{ab}|(E-\hat{H}^{(0)})a^j_{cd})^{-1}(a^i_{ab}|\hat{H}^{(1)}a_q)\end{aligned} \tag{2.4.1}$$

where the zeroth-order Hamiltonian $\hat{H}^{(0)}$ is a Fock operator $\hat{F}$ and the first-order Hamiltonian $\hat{H}^{(1)} \equiv \hat{H} - \hat{H}^{(0)}$ is a fluctuation operator $\hat{W}$. $i$, $j$, ... are occupied and $a$, $b$, ... are unoccupied orbitals of orbital basis set. $p$, $q$ are general indices of the orbital basis set. In terms of spin orbitals, eq. (2.4.1) can be written as:

$$\mathbf{\Sigma}^{(2)}_{pq}(E) = \frac{1}{2}\sum_{ija}\frac{\langle pa|\,|ij\rangle\,\langle ij|\,|qa\rangle}{E+\varepsilon_a-\varepsilon_i-\varepsilon_j} + \frac{1}{2}\sum_{iab}\frac{\langle pi|\,|ab\rangle\,\langle ab|\,|qi\rangle}{E+\varepsilon_i-\varepsilon_a-\varepsilon_b} \tag{2.4.2}$$

In Eq. (2.4.2), the first term on the right hand side is a 2-hole-1-particle ($2h1p$) term that accounts for orbital relaxation effects (correlation energy of the ionized state), while the second 2-particle-1-hole ($2p1h$) term accounts for correlation effects. Due to symmetric nature of superoperator Hamiltonian, the above self-energy expression can be used for evaluating both, ionization potential and electron affinity.

A diagonal form of self-energy, $\mathbf{\Sigma}(E)$, is used in D2, D3, OVGF, P3 and P3+ approximations.[104,109] A non-diagonal form of $\mathbf{\Sigma}(E)$ is used in *2p-h* TDA, ADC(3) and NR2 approximations.[103] Third-order contributions to self-energy for IP were formulated by Born and Öhrn.[110] Numerical results show indicate that NR2 produces least errors in IP and EA among non-diagonal self-energy approximations.[109,111]

### 2.4.2 Non-diagonal renormalized second-order Green's function (NR2)

The non-diagonal renormalized second-order (NR2) self-energy was formulated by neglecting higher scaling terms from a third-order superoperator Hamiltonian,

$$\hat{\mathbf{H}} = \begin{bmatrix} (\mathbf{a}|\hat{H}\mathbf{a})^{(3)} & (\mathbf{a}|\hat{H}\mathbf{f}_3)^{(2)} \\ (\mathbf{f}_3|\hat{H}\mathbf{a})^{(2)} & (\mathbf{f}_3|\hat{H}\mathbf{f}_3)^{(1)} \end{bmatrix}. \tag{2.4.3}$$

This form of $\hat{\mathbf{H}}$ also creates terms in higher-orders and the corresponding self-energy approximation is called 3+. OVGF, 2*p*-*h* TDA methods are also derived from the above superoperator form of Hamiltonian.

#### The NR2 superoperator for IP

For calculating ionization potentials, an asymmetric superoperator metric is defined as:

$$(\mu|\nu) = \langle 0| \left[\mu^\dagger, \nu\right]_+ (1 + \hat{T}_2^{(1)}) |0\rangle , \tag{2.4.4}$$

where $\hat{T}_2^{(1)}$ is a first-order double-excitation operator. IPs are found by diagonalizing the superoperator Hamiltonian (Eq. 2.4.5), which can be done efficiently by Löwdin partitioning of the superoperator Hamiltonian. The calculation of the diagonal $\hat{\mathbf{H}}^{(1)}_{hhp,hhp}$ block is the most

expensive step in this formulation.

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{H}}^{(0)}_{h,h} & \hat{\mathbf{H}}^{(0)}_{h,p} & \hat{\mathbf{H}}^{(1)}_{h,hhp} & \hat{\mathbf{H}}^{(1)}_{h,pph} \\ \hat{\mathbf{H}}^{(0)}_{p,h} & \hat{\mathbf{H}}^{(0)}_{p,p} & \hat{\mathbf{H}}^{(1)}_{p,hhp} & \hat{\mathbf{H}}^{(1)}_{p,pph} \\ \hat{\mathbf{H}}^{(2)}_{hhp,h} & \hat{\mathbf{H}}^{(1)}_{hhp,p} & \hat{\mathbf{H}}^{(1)}_{hhp,hhp} & \hat{\mathbf{H}}^{(0)}_{hhp,pph} \\ \hat{\mathbf{H}}^{(1)}_{pph,h} & \hat{\mathbf{H}}^{(1)}_{pph,p} & \hat{\mathbf{H}}^{(0)}_{pph,hhp} & \hat{\mathbf{H}}^{(0)}_{pph,pph} \end{bmatrix} \tag{2.4.5}$$

Substituting the metric yields the following expressions for blocks of the superoperator Hamiltonian:

$$\begin{aligned}
\mathbf{\Sigma}_{ij}(E) &= (a_i|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1}-\hat{H})a^d_{mn})^{(1)}\}^{-1}(a^d_{mn}|\hat{H}a_j)^{(2)} \\
&\quad + (a_i|\hat{H}a^k_{cd})^{(1)}\{(a^k_{cd}|(E\hat{1}-\hat{H})a^l_{ef})^{(0)}\}^{-1}(a^l_{ef}|\hat{H}a_j)^{(1)} \\
\mathbf{\Sigma}_{ai}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1}-\hat{H})a^d_{mn})^{(1)}\}^{-1}(a^d_{mn}|\hat{H}a_i)^{(2)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(1)}\{(a^k_{cd}|(E\hat{1}-\hat{H})a^l_{ef})^{(0)}\}^{-1}(a^l_{ef}|\hat{H}a_i)^{(1)} \\
\mathbf{\Sigma}_{ia}(E) &= [\mathbf{\Sigma}_{ai}(E)]^* \\
\mathbf{\Sigma}_{ab}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1}-\hat{H})a^d_{mn})^{(0)}\}^{-1}(a^d_{mn}|\hat{H}a_b)^{(1)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(1)}\{(a^k_{cd}|(E\hat{1}-\hat{H})a^l_{ef})^{(0)}\}^{-1}(a^l_{ef}|\hat{H}a_b)^{(1)}.
\end{aligned} \tag{2.4.6}$$

After evaluating the blocks independently, they are symmetrized and the Dyson equation is solved until self-consistency in $E$ is achieved. Symmetrizing procedure introduces errors in fourth-order, which are neglected. For IPs, NR2 scales as iterative $\mathcal{O}(O^2V^3)$ and the largest intermediate has $OV^3$ elements.

## The NR2 superoperator for EA

For electron affinities, the asymmetric superoperator metric is defined as:

$$(\mu|\nu) = \langle 0| \, (1 + \hat{T}_2^{(1)})^\dagger [\mu^\dagger, \nu]_+ \, |0\rangle \, . \tag{2.4.7}$$

The superoperator Hamiltonian matrix for NR2-EA is derived from the 3+ Hamiltonian (Eq. 2.4.3) . The second-order matrix element appears only in the $2p1h$,$h$ block and $\hat{\mathbf{H}}$ is given by:

$$\hat{\mathbf{H}} = \begin{bmatrix} \hat{\mathbf{H}}^{(0)}_{h,h} & \hat{\mathbf{H}}^{(0)}_{h,p} & \hat{\mathbf{H}}^{(1)}_{h,hhp} & \hat{\mathbf{H}}^{(1)}_{h,pph} \\ \hat{\mathbf{H}}^{(0)}_{p,h} & \hat{\mathbf{H}}^{(0)}_{p,p} & \hat{\mathbf{H}}^{(1)}_{p,hhp} & \hat{\mathbf{H}}^{(1)}_{p,pph} \\ \hat{\mathbf{H}}^{(1)}_{hhp,h} & \hat{\mathbf{H}}^{(1)}_{hhp,p} & \hat{\mathbf{H}}^{(0)}_{hhp,hhp} & \hat{\mathbf{H}}^{(0)}_{hhp,pph} \\ \hat{\mathbf{H}}^{(2)}_{pph,h} & \hat{\mathbf{H}}^{(1)}_{pph,p} & \hat{\mathbf{H}}^{(0)}_{pph,hhp} & \hat{\mathbf{H}}^{(1)}_{pph,pph} \end{bmatrix} \tag{2.4.8}$$

NR2-EA are obtained by diagonalization of the superoperator Hamiltonian given in Eq. (2.4.8). Symmetrizing $\hat{\mathbf{H}}$ gives the following expressions for blocks of self-energy matrix $\mathbf{\Sigma}(E)$:

$$\begin{aligned}
\mathbf{\Sigma}_{ij}(E) &= (a_i|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1} - \hat{H})a^d_{mn})^{(0)}\}^{-1}(a^d_{mn}|\hat{H}a_j)^{(1)} \\
&\quad + (a_i|\hat{H}a^k_{cd})^{(2)}\{(a^k_{cd}|(E\hat{1} - \hat{H})a^l_{ef})^{(1)}\}^{-1}(a^l_{ef}|\hat{H}a_j)^{(1)} \\
\mathbf{\Sigma}_{ai}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1} - \hat{H})a^d_{mn})^{(0)}\}^{-1}(a^d_{mn}|\hat{H}a_i)^{(1)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(2)}\{(a^k_{cd}|(E\hat{1} - \hat{H})a^l_{ef})^{(1)}\}^{-1}(a^l_{ef}|\hat{H}a_i)^{(1)} \\
\mathbf{\Sigma}_{ia}(E) &= [\mathbf{\Sigma}_{ai}(E)]^* \\
\mathbf{\Sigma}_{ab}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}\{(a^c_{kl}|(E\hat{1} - \hat{H})a^d_{mn})^{(0)}\}^{-1}(a^d_{mn}|\hat{H}a_b)^{(1)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(2)}\{(a^k_{cd}|(E\hat{1} - \hat{H})a^l_{ef})^{(1)}\}^{-1}(a^l_{ef}|\hat{H}a_b)^{(1)}.
\end{aligned} \tag{2.4.9}$$

For EA, NR2 scales as iterative $\mathcal{O}(OV^4)$ and the largest intermediate has $V^4$ elements. Evaluating the blocks with $\langle ab|cd\rangle$ type intermediates for large basis set is not practical from a memory standpoint. For such cases, integral direct formulation for the intermediates can be employed.

## Summary

The IPs obtained from Hartree-Fock orbital energies are large since they neglect the orbital relaxation of the final state. Green's function methods include these effects by augmenting the Fock matrix with the energy-dependent non-local potential called self-energy ($\mathbf{\Sigma}(E)$). The self-energy can have both diagonal or non-diagonal forms. The diagonal second-order (D2) self-energy approximation and the electron binding energy can be evaluated directly from an SCF reference.

$$E = \epsilon_p + \Sigma^{(2)}_{pp}(E) \tag{2.4.10}$$

Though this approximation is reasonable, the IPs are predicted to be smaller than the experimental values. Non-diagonal energy-dependent self-energy forms give better approximates to electron binding energy compared to diagonal methods. These methods suffer from slow convergence of basis set incompleteness error (BSIE). Hence, large basis sets are required for accurate electron binding energies. Explicitly-correlated methods have shown to dramatically improve the convergence of BSIE. These methods and their formulation for self-energy are discussed in the following chapter.

# Chapter 3

# Explicitly-correlated Methods

In wave function based methods, a linear combination of Slater determinants built from an independent-particle approximation is often used as a reference. Though they model the antisymmetric properties of electrons, they fail to account for the Coulomb repulsion between electrons of *unlike* spins. Hence, they are insufficient to describe the instantaneous electron-electron interaction, which is also know as the electron correlation problem. Electron correlation, if not addressed correctly, is the single most significant contributor to errors in any quantum chemistry calculations.

Large basis sets that increase the span of unoccupied orbitals are often employed to reduce these errors, however using large basis sets for a higher scaling methods becomes impractical for large systems. On the contrary, R12/F12 methods[112–114] provide an alternative approach to recovering the correlation energy by modeling the electron-electron cusp efficiently. This eliminates the requirement of large basis sets, or extrapolation to complete basis set limit (CBS) that requires several large bases. Presence of the cusp feature in the wave function guarantees a faster convergence of orbital expansion. The cusp in wave function for small inter electronic distance is discussed in section 3.1, the Hylleraas approach used to study the

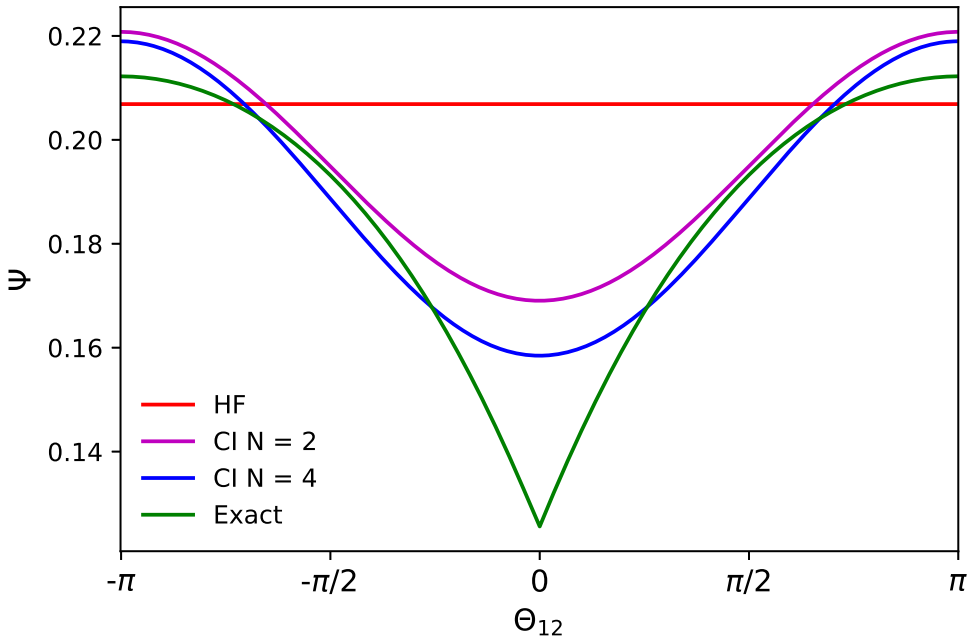



Figure 3.1: Hartree-Fock, CI and exact ground state wave functions for a Helium atom.

Helium atom is in section 3.2, modern R12/F12 methods are described in section 3.3 and the F12 correction to self-energy is discussed in section 3.4 followed by a summary.

## 3.1 The cusp in wave function

The exact wave function has a *cusp*, where the first derivative of wave function is discontinuous at a point where the two-electrons coincide. Meanwhile, the CI wave function is smooth and overestimates the probability density at the coalescence point (Figure 3.1).

The Coulomb interaction scales as the reciprocal of the distance between charged particles and diverges when the particles approach one another. However, the local energy $E = \hat{H}\Psi/\Psi$ is constant for an exact wavefunction, hence the divergence of Coulomb energy must be cancelled by the kinetic energy operator. This is achieved by including a linear dependence of inter-particle distance ($r_{12}$) in the wave function. This leads to the famous Kato's cusp

Table 3.1: Total energies and errors ($\varepsilon$) in a.u. of Helium atom calculated using CI and Hylleraas methods with increasing the angular momentum ($N$).

| | CI-like | | Hylleraas | |
|---|---|---|---|---|
| $N$ | Energy | $\varepsilon$ | Energy | $\varepsilon$ |
| 0 | -2.8350000 | 0.0687247 | -2.8350000 | 0.0687247 |
| 1 | -2.8475829 | 0.0561418 | -2.8912202 | 0.0125045 |
| 2 | -2.8929062 | 0.0108185 | -2.9034232 | 0.0003015 |
| 3 | -2.8962161 | 0.0075086 | -2.9036282 | 0.0000965 |
| 4 | -2.8986634 | 0.0050613 | -2.9037011 | 0.0000236 |
| 5 | -2.9000544 | 0.0036703 | -2.9037162 | 0.0000085 |
| 6 | -2.9009166 | 0.0028081 | -2.9037212 | 0.0000035 |
| 7 | -2.9014937 | 0.0022310 | -2.9037230 | 0.0000017 |
| 8 | -2.9019007 | 0.0018240 | -2.9037237 | 0.0000010 |

condition,[115] which models the true nature of the first derivate of the exact wave function:

$$\left[\frac{\partial\Psi\left(r_{12}\right)}{\partial r_{12}}\right]_{r_{12}\to 0} = \frac{1}{2}\Psi\left(r_{12}=0\right). \tag{3.1.1}$$

From an energy standpoint, the size and shape of the Coulomb hole around the coalescence point are important, which is inscribed by the cusp condition. This is apparent in case of a Helium atom which has been studied extensively. The Hylleraas' approach to calculating the total energy of Helium was seminal and revived decades later.

## 3.2 Hylleraas approach

The electron correlation problem is difficult to study because of the simultaneous pair-wise interactions between electrons. The Helium atom is the smallest system which exhibits such interactions. Hence, it is a good candidate for understanding electron correlation in chemical systems. An accurate approximate wave function for the Helium atom was formulated by Hylleraas as early as 1929.[116] The largest source of error in an approximate wave function

arises from our inability to model the wave function when electrons are close to each other. This problem is illustrated in the following example. The angular dependence of wave function for a Helium atom, with two electrons on a fixed radius is shown in Figure 3.1. The Hartree-Fock wave function is constant throughout and does not include the Coulomb correlation. The CI-like wave function is obtained by retaining only the even powers of $r_{12}$ term in the product of single particle wave functions. The CI-like wave function is unable to model the *cusp* feature of the exact wave function. Hence, including odd powers of $r_{12}$ term is crucial for representing the exact shape of the wave function.

For a sequence of increasing the total angular momentum, $N = n + m + l$, the total energies and errors ($\varepsilon$) of Helium atom calculated using a Hylleraas and CI-like wave functions are shown in Table 3.1. The errors are relative to the exact non-relativistic ground state energy of -2.9037247 a.u.[117] From the table, it is apparant that the Hylleraas wave function converges rapidly. This is a result of including odd powers of $r_{12}$ in the wave function expression and modern R12/F12 methods are based on these observations.

## 3.3 Modern R12/F12 method

The Hylleraas approach in its original form cannot be applied to molecules. The resolution of identity (RI) approximation was employed to factorize the many-electron problem in the new R12 methods proposed by Kutzelnigg.[112] In this approach, the canonical wave function was supplemented with the product of the reference determinant and linear $r_{12}$ factors. This approach was used to formulate the MP2-R12 method.[118,119] This approach was limited to small molecules as it used large basis sets. They assumed standard approximation (SA),[119] generalized Brillouin and the extended Brillouin condition. Auxiliary basis set (ABS),[120] complementary auxiliary basis (CABS),[121] numerical quadrature[122,123] and use of Slater-

type geminal (STG) correlation factors extended the applications of these methods to large molecular systems.

The versatility of F12 methods have led to their combination with a variety of existing methods over the last two decades. Combination with local methods like DLPNO-MP2 and linear-scaling DLPNO-CCSD(T) is recent.[27,70] Potential energy surfaces for scattering and clustering,[124] analytic energy gradients for MP2-F12 with density fitting have been formulated.[125] CCSD-F12, multi-reference F12 coupled cluster theory, explicitly-correlated Mukherjee's state-specific coupled cluster, EOM-IP-CCSD-F12 have been done by Ten-no and co-workers.[126–128] Monte Carlo GF2-F12 has been reported recently.[129] Perturbative F12 approach has been combined with orbital-optimized methods for strongly correlated systems.[130]

## 3.4 F12 corrected self-energy

In the self-energy expression (Eq. 2.4.1), the *2p1h* term is slowly convergent, similar to the MP2 pair-energy term. The $r_{12}$ dependence can be included in this term by augmenting the *Manne operator manifold*, $\mathbf{f}$, with the geminal $\mathbf{f}^{\gamma}$ operator. It is a reasonable approximation to assume that the $\mathbf{f}$ and $\mathbf{f}^{\gamma}$ operator manifolds are decoupled. This leads to additional term in the self-energy $\mathbf{\Sigma}(E)$ expression with the geminal operators $\mathbf{f}^{\gamma}$ as:[131]

$$\mathbf{\Sigma}(E) \leftarrow (\mathbf{a}|\hat{H}^{(1)}\mathbf{f}^{\gamma})(\mathbf{f}^{\gamma}|(E-\hat{H}^{(0)})\mathbf{f}^{\gamma})^{-1}(\mathbf{f}^{\gamma}|\hat{H}^{(1)}\mathbf{a}). \tag{3.4.1}$$

In the above equation, $\mathbf{f}^{\gamma} = \frac{1}{2}\tilde{R}^{\alpha\beta}_{ir}\tilde{a}^{i}_{\alpha\beta}$is the geminal field operator and $\tilde{R}$ is the anti-symmetrized matrix element of geminal correlation factor $f(r_{12})$:

$$R^{\alpha\beta}_{ir} \equiv \left\langle ir \left| \hat{Q} f\left(r_{12}\right) \right| \alpha\beta \right\rangle. \tag{3.4.2}$$

The pair-spin projections arising due to singlet and triplet cusp conditions are implicit in this expression and $\hat{Q}$ guarantees the orthogonality of geminal functions with the Hartree-Fock orbitals and double excitation amplitudes. The resolution of matrix elements yields standard intermediates from F12 theory. The first-order terms are given by:

$$(\mathbf{a}|\hat{H}^{(1)}\mathbf{f}^{\gamma}) = \frac{1}{2}\overline{g}^{\gamma\delta}_{kp}\tilde{R}^{kr}_{\gamma\delta} = \frac{1}{2}V^{kr}_{kp} \tag{3.4.3}$$

and

$$(\mathbf{f}^{\gamma}|\hat{H}^{(1)}\mathbf{a}) = \frac{1}{2}\tilde{R}^{\alpha\beta}_{is}\overline{g}^{iq}_{\alpha\beta} = \frac{1}{2}V^{iq}_{is}. \tag{3.4.4}$$

The intermediate $V$ from the above expressions is given by:

$$V^{kl}_{ij} = \frac{1}{2}\overline{g}^{\alpha\beta}_{ij}\tilde{R}^{kl}_{\alpha\beta}. \tag{3.4.5}$$

The resolvent from Eq. 4.2.6 yields the following terms:

$$\begin{aligned}\left(\mathbf{f}^{\gamma}|\hat{H}^{(0)}\mathbf{f}^{\gamma}\right) &= \frac{1}{2}\delta^{i}_{k}\tilde{R}^{\alpha\beta}_{is}F^{\gamma}_{\alpha}\tilde{R}^{kr}_{\gamma\beta} - \frac{1}{4}F^{i}_{k}\tilde{R}^{\alpha\beta}_{is}\tilde{R}^{kr}_{\alpha\beta} \\ &= \frac{1}{2}\delta^{i}_{k}B^{kr}_{is} - \frac{1}{4}F^{i}_{k}X^{kr}_{is}\end{aligned} \tag{3.4.6}$$

and

$$E\left(\mathbf{t}^{\gamma}|\mathbf{t}^{\gamma}\right) = \frac{1}{4}E\delta^{i}_{k}\tilde{R}^{\alpha\beta}_{is}\tilde{R}^{kr}_{\alpha\beta} = \frac{1}{4}E\delta^{i}_{k}X^{k}_{is}. \tag{3.4.7}$$

From the above expressions, intermediates $B$ and $X$ are given by:

$$B_{kl}^{mn} = \tilde{R}_{\alpha\gamma}^{mn} F_{\beta}^{\alpha} \tilde{R}_{kl}^{\beta\gamma} \tag{3.4.8}$$

and

$$X_{kl}^{mn} = \frac{1}{2} \tilde{R}_{\alpha\beta}^{mn} \tilde{R}_{kl}^{\alpha\beta}. \tag{3.4.9}$$

This formulation has been applied successfully for studying ionization potentials using energy dependent non-diagonal second-order Green's function (GF2-F12) and non-diagonal renormalized second-order (NR2-F12) methods. The resolvent of *2p1h* term is zeroth order for GF2-F12 hence, this F12 correction is applicable for electron affinities too. However, the resolvent is of first order for NR2-EA formulation. Hence, using the zeroth-order F12 correction leads to fallacies in the formalism and is manifested in the results as shown in chapter 5. The first order resolvent formulation leads to an additional intermediate $P$, which is also seen in CCSD-R12 formalism. The expression for intermediate $P$ is given by:

$$P_{xy}^{ow} = \frac{1}{4} \tilde{R}_{\alpha\beta}^{ow} \overline{g}_{\gamma\delta}^{\alpha\beta} \tilde{R}_{xy}^{\gamma\delta}. \tag{3.4.10}$$

The large scaling of this term makes it impractical for applications to the current system of interests and the implementation of a higher-order resolvent is left for future work.

# Summary

The *electron-electron* cusp is correctly represented by inclusion of the inter-electronic distance in the wave function. The shape of the Coulomb hole has a strong influence on the energy of the molecular system, hence the cusp needs to be included for accurate description of the electronic wave function. This accurate description of wave function reduces basis set

incompleteness errors allowing us to calculate accurate energies with smaller basis sets (less functions). In case of Green's function methods for electron binding energies, the slow convergence of double summation over virtual orbitals of orbital basis set (OBS) can be efficiently handled with an explicitly correlated self-energy correction.

# Chapter 4

# Explicitly correlated renormalized second-order Green's function for accurate ionization potentials of closed-shell molecules

Reproduced from Teke, N. K.; Pavošević, F.; Peng, C.; Valeev, E. F. *The Journal of Chemical Physics* **2019**, *150*(20), 00000., with the permission of AIP Publishing.

## 4.1 Introduction

Green's function (GF) (propagator) formalisms are the foundation of the many-body electronic structure in physics.[132–134] This is also the case in chemistry,[92,102,135] where they draw increasing attention as an alternative to wave function/operator methods[96,107,108,136–139] as well as a way to systematically improve the density functional theory.[97,140–143] Although recent chemistry literature has utilized single-particle and polarization propagators for computing total correlation energies[98,99] and as a basis for quantum mechanical embedding methods,[95,144–146] the traditional application of single-particle (electron) propagator (EP) methods were for computing ionization potential, electron affinity and photoelectron spectra.[85,86,147? ? –151] The appeal of propagator methods is their direct connection to the observables as well as their interpretive power.[104]

Conventional spectral solvers for propagator methods, utilizing expansions in a Fock subspace built from a finite single-particle basis, suffer from slow asymptotic decay of basis set errors of energies (such as ionization potentials, total correlation energies, etc.) and other properties. This is due to the inability to model efficiently localized features such as electron-electron cusps with products of single-particle basis functions. Explicitly correlated wave function methods solve this problem by introducing 2-particle basis functions that describe the cusp structure directly. The R12/F12 wave function methods realized this idea practically by approximating the many-electron integrals by resolution of the identity and other techniques following the original ideas by Kutzelnigg[152] and finalized by many others.[113,153–160] Recently the explicitly correlated F12 formalism was also applied to propagator methods. Ohnishi and Ten-no utilized MP2-F12 energies to correct the *diagonal* elements of self-energy at *unperturbed* (Koopmans) energies;[161] recently this approach was reformulated using stochastic integration formalism and applied to large systems.[129] General *nondiagonal energy-dependent* F12 correction to the second-order self-energy (GF2-F12) was proposed

by some of us.[131] The GF2-F12 method scales as iterative $\mathcal{O}(N^5)$ and shows much faster convergence to the complete basis set limit than the standard GF2 counterpart, thus requiring only a triple-zeta basis to reduce the numerical error in ionization potentials below 0.05 eV. However, the accuracy of the second-order (non-self-consistent) model of self-energy is limited and more complete models of self-energy must be utilized.

Among nondiagonal EP methods,[103,104] the renormalized second-order (NR2) approximation of EP is a promising method to calculate ionization potentials.[105] Numerical results indicate that the performance of NR2 is the best among non-diagonal EP methods at a complete basis set limit.[106] In this article, we investigate the basis set convergence of an energy dependent F12 correction to the NR2 method. IPs calculated with NR2-F12 for small and medium sized organic molecules (Figure 4.1) are compared to GF2-F12, EOM-IP-CCSD, CCSD(T)-F12 methods and a CCSD(T) reference IPs extrapolated to the CBS limit.

This article is organized as follows. Section 4.2 introduces the NR2 method and how its basis set error is reduced via an F12 correction to self-energy. The implementation details are included in section 4.3. Basis set errors of NR2-F12, their accuracy compared to other methods against a CCSD(T) benchmark, is presented in section 4.4, followed by conclusions in section 4.5.

## 4.2 Formalism

We have used the time-independent (superoperator) EP formalism in this work.[102,162] The eigenvalues, which correspond to the poles of EP are obtained by solving the Dyson quasiparticle equation self-consistently. The Dyson equation is an eigenvalue problem given by

$$[\mathbf{F} + \mathbf{\Sigma}(E)]\mathbf{C} = E\mathbf{C}. \tag{4.2.1}$$

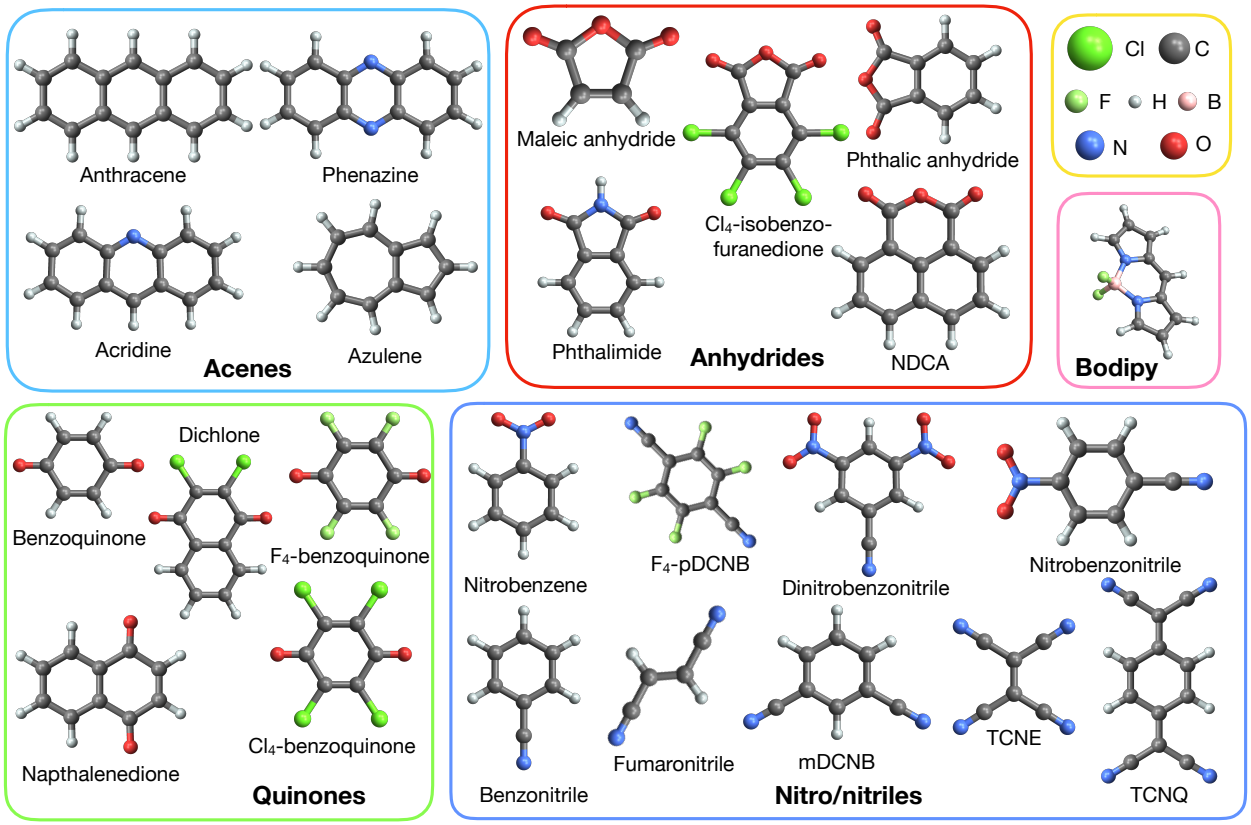



Figure 4.1: Benchmark set of 24 medium sized organic molecules (OAM24) from Ref. 1 used in this work.

In this equation, $\mathbf{\Sigma}(E)$ is the energy-dependent non-local potential called the self-energy; $\mathbf{F}$ is the Fock matrix and $\mathbf{C}$ are canonical Hartree-Fock orbitals in case of a diagonal self-energy and Dyson orbitals for a non-diagonal self-energy. The off-diagonal elements of the self-energy matrix in canonical basis are small and have negligible effect on the poles and the Dyson orbitals. The derivation and approximations to the Dyson quasiparticle equation are discussed by Corzo and Ortiz.[103]

The focus of this work is on a particular model for the self-energy operator — specifically, the Non-diagonal Renormalized 2nd-order (NR2) model — proposed by Ortiz.[105] Recently it has been demonstrated as a particularly cost-effective way to compute accurate ionization potentials and electron affinities of lowest IP/EA of organic molecules.[106] The non-Hermitized

NR2 self-energy for electron detachment (NR2-D) is represented as follows:

$$\begin{aligned}
\mathbf{\Sigma}_{ij}(E) &= (a_i|\hat{H}a^c_{kl})^{(1)}(a^c_{kl}|(E\hat{1}-\hat{H})^{-1}a^d_{mn})^{(1)}(a^d_{mn}|\hat{H}a_j)^{(2)} \\
&\quad + (a_i|\hat{H}a^k_{cd})^{(1)}(a^k_{cd}|(E\hat{1}-\hat{H})^{-1}a^l_{ef})^{(0)}(a^l_{ef}|\hat{H}a_j)^{(1)} \\
\mathbf{\Sigma}_{ai}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}(a^c_{kl}|(E\hat{1}-\hat{H})^{-1}a^d_{mn})^{(1)}(a^d_{mn}|\hat{H}a_i)^{(2)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(1)}(a^k_{cd}|(E\hat{1}-\hat{H})^{-1}a^l_{ef})^{(0)}(a^l_{ef}|\hat{H}a_i)^{(1)} \\
\mathbf{\Sigma}_{ia}(E) &= [\mathbf{\Sigma}_{ai}(E)]^* \\
\mathbf{\Sigma}_{ab}(E) &= (a_a|\hat{H}a^c_{kl})^{(1)}(a^c_{kl}|(E\hat{1}-\hat{H})^{-1}a^d_{mn})^{(0)}(a^d_{mn}|\hat{H}a_b)^{(1)} \\
&\quad + (a_a|\hat{H}a^k_{cd})^{(1)}(a^k_{cd}|(E\hat{1}-\hat{H})^{-1}a^l_{ef})^{(0)}(a^l_{ef}|\hat{H}a_b)^{(1)},
\end{aligned} \tag{4.2.2}$$

where the matrix elements of the superoperator are defined as

$$(\mu|\nu) = \langle 0|[\mu^\dagger,\nu]_+(1+\hat{T}_2^{(1)})|0\rangle, \tag{4.2.3}$$

where $|0\rangle$ is the Hartree-Fock reference state and $\hat{T}_2^{(1)}|0\rangle$ is the first-order Møller-Plesset (MP) wave function. We have used standard Einstein-summed tensor notation for products of annihilation ($a_p$) and creation ($a^p \equiv a_p^\dagger$) operators that are normal ordered with respect to the physical vacuum.[163] Indices $i, j, \ldots$ refer to the active occupied orbitals (holes), $a, b, \ldots$ — to the unoccupied (particle) orbitals represented in the orbital basis set (OBS), $p, q, r$ — to any (hole or particle) active orbital represented in OBS, and $\alpha$, $\beta$ — to the unoccupied orbitals of the complete basis set (CBS). The superscript indices on superoperator matrix elements denote the *maximum* order (in the MP sense) of the terms kept, and the operator inverse is defined by the inverse of its given basis representation. The Hermitized expressions for NR2-D were used in practice and are given by Corzo and Ortiz.[103]

The high accuracy of NR2 IPs/EAs is only revealed when the numerical error is eliminated. Unfortunately, the rate of basis set convergence of IPs/EAs is slow due to the slow con-

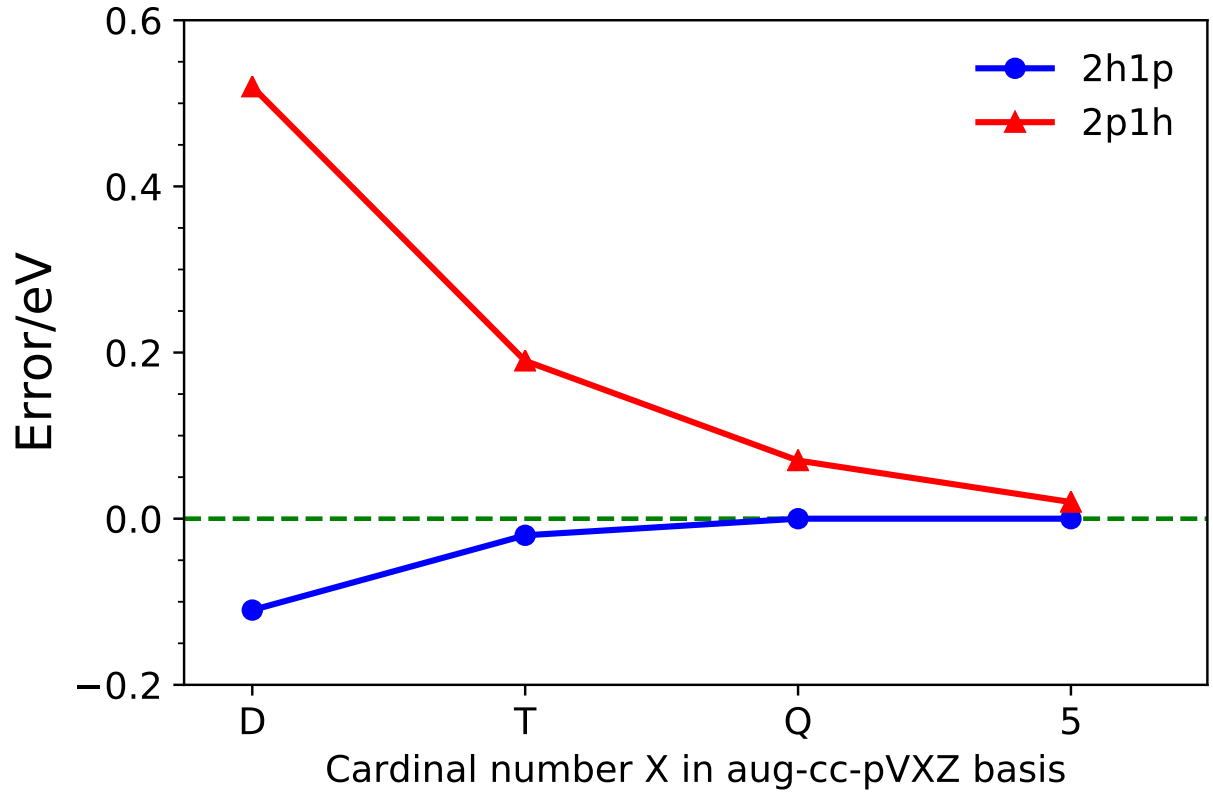



Figure 4.2: Basis set errors in $2h1p$ and $2p1h$ contributions to first ionization potential of a $N_2$ molecule evaluated with the non-iterative diagonal second-order self-energy model.

vergence of summations over unoccupied states in the (to leading order) $2p1h$ component of the self-energy. This is illustrated for the lowest-energy detachment pole of the energy-dependent diagonal second-order self-energy in Figure 4.2. It is easy to see that the basis set convergence of the $2h1p$ and $2p1h$ components (the first and second terms on the right-hand side of Eq. (4.2.2)) of the self-energy are fundamentally different. In an atom exact evaluation of the $2h1p$ contribution only requires the basis of 1-particle states to be complete through $3L_{\text{occ}}$, where $L_{\text{occ}}$ is the maximum angular momentum of an occupied orbital (this roughly explains why for $p$-elements the $2h1p$ contribution is essentially converged with a TZ basis, which is first to gain basis functions of angular momentum 3). In contrast, the summation over the particle states in the $2p1h$ term does not truncate and thus is analogous

to the slowly convergent error of truncated partial wave expansion of the atomic correlation energies.[164] To reduce the basis set error of this term some kind of basis set extrapolation is typically employed. However, a more robust solution to the basis set problem is to employ explicit correlation.

The explicitly correlated R12/F12 formalism[152,157–159] has been established as the standard solution to the basis set problem of wave function methods in which the many-electron basis includes terms with explicit dependence on the inter-electronic distances. Such formalism can be also extended to the self-energy operator, as we recently showed in the case of (non-diagonal) energy-dependent case;[131] note that our approach specifically accounts for the energy dependence of the explicitly correlated term. The energy-dependent F12 correction to self-energy is obtained by augmenting the slowly convergent $\mathbf{f} = \{a^i_{ab}\}$ operator manifold with the geminal field operator manifold, $\mathbf{f}^\gamma$, given by

$$\mathbf{f}^\gamma = \frac{1}{2}\tilde{R}^{\alpha\beta}_{ir}\tilde{a}^i_{\alpha\beta} \tag{4.2.4}$$

where $\tilde{R}$ is the antisymmetrized matrix element of the geminal correlation factor $f(r_{12})$ given by

$$R^{\alpha\beta}_{ir} \equiv \langle ir|\hat{Q}f(r_{12})|\alpha\beta\rangle. \tag{4.2.5}$$

and the projector $\hat{Q}$ ensures orthogonality of geminal and conventional $2h1p$ field operators. By neglecting the coupling blocks of the superoperator between the conventional and F12 manifolds the conventional $2p1h$ self-energy is corrected by *addition* of the F12 contribution:

$$\mathbf{\Sigma}_{\mathrm{F12}}(E) = (\mathbf{a}|\hat{H}^{(1)}\mathbf{f}^\gamma)(\mathbf{f}^\gamma|(E-\hat{H}^{(0)})\mathbf{f}^\gamma)^{-1}(\mathbf{f}^\gamma|\hat{H}^{(1)}\mathbf{a}) \tag{4.2.6}$$

Since the conventional $2p1h$ contribution to the self-energy in the NR2-D approach is identical to its GF2 counterpart, we obtain the NR2-F12 self-energy expression for the detachment

processes by simple addition of the F12 self-energy correction; note that for the NR2-A approach we expect that the F12 correction will need to be modified to account for the presence of the higher-order terms in the $2p1h$ self-energy. The $2h1p$ contribution to the NR2 or GF2 self-energy is not affected by the F12 terms, and it converges quickly to the CBS limit.

The structure of the coupling between the conventional and F12 manifolds in the GF2-F12 and NR2-D-F12 methods is completely analogous to their coupling in the MP2-F12 method (e.g. see Refs 165 and 157); namely it involved matrix elements of the Fock operator between the conventional unoccupied orbitals and their complement in the complete basis. The effect of the coupling on the MP2-F12 energies is smaller than the residual basis set errors (and the method errors) and becomes completely negligible with triple-zeta and larger basis sets,[165] and for the ionization potentials we expect the same trends. Since the 2h1p channel's contribution to the self-energy requires a triple-zeta basis anyway, the neglect of the coupling is justified.

Resolution of matrix elements of Eq.(4.2.6) yields standard F12 intermediates $V$, $X$ and $B$.[131] The derivation and programmable equations of these terms are available in literature.[157] In this work, the intermediates $V$ and $X$ are calculated using the complementary auxiliary basis set (CABS) approach.[121] The bottleneck of evaluating the F12 correction is the intermediate $B$, which in this work for efficiency was computed in approximation $D$,[27] rather than the more complete approximation $C$.[166]

With canonical Hartree-Fock orbital energy as an initial guess, the eigenvalue problem of Eq. (4.2.1) is solved iteratively until the root is converged with the energy-dependent F12 correction to self-energy included in every iteration.

## 4.3 Computational details

Distributed-memory parallel implementations of NR2 and NR2-F12 methods were developed in the Massively Parallel Quantum Chemistry package (MPQC).[167] The performance of the NR2-F12 method was assessed for a set of 21 small molecules used previously by Ortiz and co-workers (and referred to here as O21)[111] and the OAM24 benchmark set of 24 medium-sized organic electron acceptor molecules,[1] with the Cartesian geometries taken from the respective references. The frozen core approximation is invoked in all the calculations reported in this work. Standard (global) density fitting (DF) approximation for 2-electron integrals was used for efficiency throughout.

We have used the augmented correlation consistent basis sets that can be reliably extrapolated to the complete basis set (CBS) limit.[168–170] The aug-cc-pVXZ orbital basis (denoted for brevity as aXZ), with X=D,T,Q and 5, were paired with the corresponding aug-cc-pVXZ-RI basis for density fitting[171] and the aug-cc-pVXZ-OptRI basis for approximating many-electron F12 integrals.[172] The standard Slater-type correlation factor,

$$f(r_{12}) = (1 - \exp(-\gamma r_{12})/\gamma). \tag{4.3.1}$$

was used in all F12 computations, with the geminal exponent $\gamma$ set as in Ref. 131 to 1.3, 1.9 and 2.1 $a_0^{-1}$ for the aDZ, aTZ, and aQZ orbital basis sets, respectively.

All propagator F12 calculations evaluated intermediate B using approximation D;[27] for the O21 set, the difference in IPs evaluated with the F12/D approach vs the more complete F12/C approach with the aTZ basis was found to be at most 0.002 eV and 0.001 eV on average, i.e. completely negligible. The GF2 and NR2 CBS IPs for the O21 and OAM24 sets were calculated by two-point $X^{-3}$ extrapolation[173] using the {aQZ,a5Z} and {aTZ,aQZ} basis set pairs, respectively.

Table 4.1: Mean absolute (MAE) and max absolute (MaxAE) basis set errors in eV for NR2 IPs for the O21 set.

| | MAE | | MaxAE | |
|---|---|---|---|---|
| Basis set | NR2 | NR2-F12 | NR2 | NR2-F12 |
| aDZ | 0.416 | 0.049 | 0.515[a] | 0.085[b] |
| aTZ | 0.185 | 0.028 | 0.246[c] | 0.047[c] |
| aQZ | 0.087 | 0.015 | 0.119[c] | 0.037[c] |
| a5Z | 0.044 | | 0.061[c] | |

[a] $CH_3COCH_3$.
[b] $H_2CO$.
[c] HF.

The reference IPs were computed as the differences between the CCSD(T) energies of the cation and neutral species. For the O21 set, the CCSD(T) and RI-CCSD(T)$_{\overline{\text{F12}}}$ (for simplicity denoted CCSD(T)-F12)[68,174,175] reference total energies were calculated using Orca electronic structure program package.[176] The geminal correlation exponents in CCSD(T)-F12 computations were set to 1.1 and 1.2 for the aDZ and aTZ bases, respectively.[177] The Hartree-Fock and correlation energies were extrapolated separately using the {aTZ,aQZ} bases via the $e^{-X\alpha}$ and $X^{-\beta}$ schemes, respectively.[173,178] CCSD(T) CBS IPs for the OAM24 dataset were taken from Ref. 1.

The EOM-IP-CCSD calculations are performed using the parallel implementation in MPQC version 4. The aug-cc-pVXZ family of basis set were used and extrapolated using the standard two-point $X^{-3}$ formula. CBS IP of the O21 set were calculated from aTZ and aQZ basis while for OAM24, aDZ and aTZ bases were used. All IPs (in eV) calculated using GF and coupled cluster methods for O21 and OAM24 can be found in the supplementary information.

Table 4.2: Differences of IP in eV of the O21 set calculated using Green's function and coupled-cluster methods with aug-cc-pVXZ (X=D,T) basis relative to CCSD(T) IP extrapolated to the CBS limit.

| | GF2-F12 | | NR2-F12 | | CCSD(T)-F12/RI | | EOM-IP-CCSD | | | CCSD(T) | Exp. |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Molecule | aDZ | aTZ | aDZ | aTZ | aDZ | aTZ | aDZ | aTZ | CBS[1] | CBS | |
| $C_2H_2$ | -0.018 | -0.049 | 0.177 | 0.152 | -0.063 | -0.026 | -0.062 | 0.062 | 0.149 | 11.467 | 11.02 |
| $C_2H_4$ | -0.270 | -0.261 | -0.075 | -0.064 | -0.067 | -0.025 | -0.170 | -0.036 | 0.039 | 10.713 | 10.51 |
| $C_5NH_5$ | -0.248 | -0.241 | -0.012 | -0.005 | -0.069 | -0.029 | -0.230 | -0.095 | -0.017 | 9.852 | 9.60 |
| $C_6H_6$ | -0.253 | -0.248 | 0.007 | 0.022 | -0.033 | 0.011 | -0.204 | -0.070 | 0.006 | 9.421 | 9.25 |
| $CH_3CH_2CH_3$ | -0.306 | -0.288 | -0.184 | -0.162 | -0.049 | -0.013 | -0.189 | -0.056 | 0.014 | 12.171 | 11.51 |
| $CH_3CHO$ | -0.988 | -0.954 | -0.180 | -0.135 | -0.086 | -0.045 | -0.347 | -0.123 | 0.012 | 10.380 | 10.26 |
| $CH_3Cl$ | -0.280 | -0.275 | -0.170 | -0.207 | -0.058 | 0.021 | -0.310 | -0.140 | 0.024 | 11.495 | 11.29 |
| $CH_3COCH_3$ | -1.000 | -0.966 | -0.155 | -0.111 | -0.081 | -0.044 | -0.341 | -0.111 | 0.027 | 9.879 | 9.70 |
| $CH_3F$ | -0.967 | -0.957 | -0.265 | -0.242 | -0.087 | -0.039 | -0.375 | -0.194 | -0.059 | 13.477 | 13.04 |
| $CH_3OH$ | -0.941 | -0.907 | -0.199 | -0.155 | -0.084 | -0.049 | -0.385 | -0.176 | -0.044 | 11.186 | 10.94 |
| $CH_3SH$ | -0.218 | -0.185 | -0.185 | -0.171 | -0.067 | 0.004 | -0.285 | -0.083 | 0.032 | 9.554 | 9.46 |
| $CH_4$ | -0.250 | -0.240 | -0.141 | -0.128 | -0.060 | -0.018 | -0.170 | -0.042 | 0.030 | 14.415 | 13.60 |
| CO | 0.044 | 0.081 | 0.078 | 0.131 | 0.002 | 0.002 | -0.068 | 0.123 | 0.220 | 14.078 | 14.01 |
| $CO_2$ | -0.615 | -0.606 | 0.107 | 0.127 | -0.084 | -0.071 | -0.353 | -0.129 | 0.031 | 13.906 | 13.78 |
| $H_2CO$ | -1.051 | -1.015 | -0.222 | -0.173 | -0.090 | -0.040 | -0.388 | -0.167 | -0.036 | 11.017 | 10.88 |
| HCl | -0.201 | -0.195 | -0.093 | -0.135 | -0.065 | 0.031 | -0.305 | -0.131 | 0.041 | 12.845 | 12.75 |
| HCN | 0.037 | 0.018 | 0.252 | 0.245 | -0.086 | -0.037 | -0.051 | 0.096 | 0.196 | 13.734 | 13.60 |
| HCOOH | -0.988 | -0.964 | -0.110 | -0.074 | -0.084 | -0.054 | -0.341 | -0.119 | 0.028 | 11.659 | 11.51 |
| HF | -1.330 | -1.345 | -0.184 | -0.182 | -0.018 | -0.043 | -0.432 | -0.247 | -0.082 | 16.295 | 16.05 |
| $N_2$ | -0.432 | -0.406 | 0.236 | 0.262 | 0.011 | 0.017 | -0.187 | 0.025 | 0.143 | 15.579 | 15.600 |
| $H_2S$ | -0.167 | -0.130 | -0.151 | -0.137 | -0.079 | 0.010 | -0.292 | -0.079 | 0.040 | 10.504 | 10.48 |
| MAE[2] | 0.505 | 0.492 | 0.152 | 0.144 | 0.063 | 0.030 | 0.261 | 0.110 | 0.060 | 0.23[3] | |
| MaxAE[2] | 1.330 | 1.345 | 0.265 | 0.262 | 0.090 | 0.071 | 0.432 | 0.247 | 0.220 | 0.82[3] | |

[a]CBS value calculated from $X^{-3}$ extrapolation of aTZ and aQZ IP.

[b]Mean and max absolute errors of IP obtained from GF and coupled-cluster methods are calculated with respect to CCSD(T) CBS reference.

[c]Mean and max absolute errors of CCSD(T) reference are with respect to the experimental value.

## 4.4 Results and discussion

The mean and maximum absolute basis set errors of NR2 and NR2-F12 IPs for the O21 set are listed in Table 4.1. Already with a double-zeta basis the basis set error of the explicitly correlated NR2-F12 method is almost as good as the conventional quintuple-zeta NR2 counterpart. A triple-zeta basis is sufficient to converge the IPs to better than 0.05 eV.

Errors in GF2-F12, NR2-F12, CCSD(T)-F12, and EOM-IP-CCSD IPs with respect to the CBS CCSD(T) reference, along with the experimental values, for the O21 set are reported in Table 4.2. As expected, NR2-F12 IPs are dramatically more accurate than the GF2-F12 IPs, with mean/max errors of only 0.14/0.26 eV vs 0.49/1.35 eV. However, NR2-F12 is less accurate than the more expensive CCSD(T)-F12 method. The closest counterpart to the NR2 approach is the EOM-IP-CCSD, whose cost is iterative $\mathcal{O}(N^6)$. The EOM-IP-CCSD CBS IPs are significantly more accurate than NR2-F12, with the former's mean absolute error lower by more than a factor of 2. The maximum errors of these two methods are however comparable. It is notable that the largest deviations of EOM-IP-CCSD CBS IPs from the reference CCSD(T) CBS IPs are observed for *linear* molecules; this suggests that the unphysical loss of axial symmetry in the CCSD(T) treatment of ionized states might be at fault here and it is desirable to recompute the reference IPs using EOM-IP-CC with inclusion of triples to resolve the significant discrepancies between CCSD(T) and EOM-IP-CCSD.

The mean absolute and maximum basis set errors of NR2 and NR2-F12 IPs for the OAM24 set are listed in Table 4.3. The NR2-F12 IPs with only a double-zeta basis is closer to the NR2 numerical limit than the conventional quadruple-zeta NR2 IPs. The aDZ basis is already sufficient to converge the NR2-F12 IPs to better than 0.05 eV.

Errors in GF2-F12, NR2-F12, CCSD(T)-F12, and EOM-IP-CCSD IPs with respect to the CBS CCSD(T) reference, along with the experimental values (where available), for the

Table 4.3: Mean absolute errors (MAE) and max absolute errors (MaxAE) in eV for OAM24 dataset with respect to the NR2 CBS. CBS values are calculated by extrapolating aTZ and aQZ bases.

| | MAE | | MaxAE | |
|---|---|---|---|---|
| Basis set | NR2 | NR2-F12 | NR2 | NR2-F12 |
| aDZ | 0.391 | 0.033 | 0.494[a] | 0.056[a] |
| aTZ | 0.159 | 0.015 | 0.204[a] | 0.022[b] |
| aQZ | 0.067 | | 0.086[a] | |

[a]Maleic anhydride.
[b]$F_4$-pDCNB.

OAM24 set are reported in Table 4.4. The reference CCSD(T) CBS and experimental values of IP for OAM24 dataset are used from Ref. 1. We are not aware of any EOM-IP-CCSD data for the OAM24 dataset. The molecules are categorized into four groups and Bodipy to gauge the performance of methods for each subgroup of this dataset. Compared to the O21 set, here the mean accuracy of GF2-F12 IPs is already quite high ($\sim$ 0.2 eV), however the errors are systematically larger for acenes and a particularly large error ($\sim$ 0.8 eV) is observed for benzoquinone. NR2-F12 IPs are somewhat more accurate on average ($\sim$ 0.15 eV) and have a drastically reduced maximum error. EOM-IP-CCSD IPs are again significantly more accurate on average ($\sim$ 0.08 eV) than NR2-F12, albeit at a higher cost. It is also notable that the largest errors of EOM-IP-CCSD and NR2-F12 IPs are correlated: errors greater than 0.17 eV for EOM-IP-CCSD occur for the same 3 molecules for which NR2-F12 errors exceed 0.3 eV, and the signs for the errors agree in all cases. This suggests that perhaps again unphysical polarization in the CCSD(T) treatment of cations might be at fault here and higher-order EOM-IP-CC values are highly desired.

Table 4.4: Differences of IP in eV of OAM24 dataset calculated with GF and coupled-cluster methods and aXZ (X=D,T) basis relative to CCSD(T) CBS values from Ref. 1.

| | | GF2-F12 | | NR2-F12 | | EOM-IP-CCSD | | | CCSD(T) | Exp. |
|---|---|---|---|---|---|---|---|---|---|---|
| | Molecule | aDZ | aTZ | aDZ | aTZ | aDZ | aTZ | CBS[1] | CBS | |
| Acenes | Anthracene | -0.313 | -0.309 | 0.009 | 0.024 | -0.235 | -0.101 | -0.045 | 7.52 | 7.44 |
| | Acridine | -0.347 | -0.341 | -0.025 | -0.006 | -0.247 | -0.105 | -0.045 | 8.04 | 7.8 |
| | Phenazine | -0.339 | -0.344 | -0.028 | -0.010 | -0.228 | -0.082 | -0.021 | 8.47 | 8.44 |
| | Azulene | -0.371 | -0.364 | -0.043 | -0.023 | -0.300 | -0.161 | -0.102 | 7.55 | 7.42 |
| Quinones | Benzoquinone | -0.831 | -0.807 | -0.102 | -0.077 | -0.227 | -0.001 | 0.094 | 10.27 | 10.0 |
| | Naphthalenedione | -0.192 | -0.189 | -0.039 | -0.021 | -0.158 | -0.018 | 0.041 | 9.88 | 9.5 |
| | Dichlone | -0.273 | -0.285 | 0.116 | 0.103 | -0.020 | 0.112 | 0.167 | 9.99 | 9.5 |
| | $F_4$-benzoquinone | -0.150 | -0.158 | 0.305 | 0.308 | -0.094 | 0.096 | 0.177 | 11.14 | 10.7* |
| | $Cl_4$-benzoquinone | -0.157 | -0.184 | 0.244 | 0.207 | -0.148 | 0.010 | 0.076 | 10.25 | 9.74* |
| Nitro/nitriles | Nitrobenzene | -0.129 | -0.127 | 0.124 | 0.138 | -0.138 | 0.002 | 0.061 | 10.19 | 9.94 |
| | $F_4$-pDCNB | -0.228 | -0.231 | 0.239 | 0.249 | -0.161 | 0.025 | 0.104 | 10.76 | 10.65 |
| | Dinitrobenzonitrile | 0.032 | 0.032 | 0.322 | 0.335 | -0.058 | 0.103 | 0.170 | 11.15 | N/A |
| | Nitrobenzonitrile | -0.091 | -0.090 | 0.232 | 0.248 | -0.128 | 0.030 | 0.097 | 10.62 | 10.59 |
| | Benzonitrile | -0.212 | -0.208 | 0.076 | 0.092 | -0.201 | -0.054 | 0.008 | 9.93 | 9.73 |
| | Fumaronitrile | -0.104 | -0.100 | 0.114 | 0.127 | -0.137 | 0.028 | 0.098 | 11.48 | 11.3† |
| | mDCNB | -0.173 | -0.169 | 0.127 | 0.143 | -0.186 | -0.031 | 0.034 | 10.45 | 10.2 |
| | TCNE | -0.037 | -0.035 | 0.306 | 0.317 | -0.076 | 0.099 | 0.172 | 11.99 | 11.79* |
| | TCNQ | -0.096 | -0.092 | 0.194 | 0.213 | -0.133 | 0.036 | 0.107 | 9.57 | N/A |
| Anhydrides | Maleic Anhydride | -0.113 | -0.101 | -0.090 | -0.051 | -0.306 | -0.059 | 0.044 | 11.33 | 11.07† |
| | Phthalimide | -0.065 | -0.058 | 0.110 | 0.140 | -0.057 | -0.017 | 0.000 | 10.08 | 9.90† |
| | Phthalic Anhydride | -0.131 | -0.126 | 0.147 | 0.162 | -0.149 | -0.004 | 0.056 | 10.55 | 10.1* |
| | $Cl_4$-isobenzofuranedione | -0.139 | -0.170 | 0.296 | 0.265 | -0.150 | 0.000 | 0.064 | 10.05 | 10.8 |
| | NDCA | -0.229 | -0.225 | 0.122 | 0.138 | -0.210 | -0.061 | 0.002 | 9.14 | 8.92 |
| Other | Bodipy | -0.163 | -0.156 | 0.072 | 0.097 | -0.230 | -0.076 | -0.011 | 8.07 | N/A |
| | MAE[2] | 0.205 | 0.204 | 0.145 | 0.146 | 0.166 | 0.055 | 0.075 | 0.27[3] | |
| | MaxAE[2] | 0.831 | 0.807 | 0.322 | 0.335 | 0.306 | 0.161 | 0.177 | 0.75[3] | |

*Adiabatic IP †Vertical IP

[a]CBS value calculated by the $X^{-3}$ extrapolation of aDZ and aTZ IPs.

[b]Mean and max errors relative to the CCSD(T) CBS values.

[c]Errors relative to the experimental values.

## 4.5 Conclusions

We presented a simple explicitly correlated (F12) extension of the renormalized second-order propagator (NR2) method for computing electron detachment energies. The NR2-F12 is characterized by greatly reduced basis set errors compared to the conventional NR2 method: precision of double-zeta basis NR2-F12 IPs is better than that of quadruple- or even quintuple-zeta NR2 IPs. Accuracy of NR2-F12 IPs is significantly higher than the non-renormalized (nondiagonal) GF2-F12 IPs, but is still substantially lower than that of EOM-IP-CCSD when assessed for two reference IP datasets (O21[111] and OAM24[1]). Nevertheless, NR2-F12 is a good starting point for higher-order models of superoperator/self-energy and can be an attractive alternative to EOM-IP-CCSD for well-behaved systems due to its non-iterative (rather than iterative) $\mathcal{O}(N^6)$ cost. The EOM-IP-CCSD CBS IP estimates for the O21 and OAM24 datasets point out potential inaccuracies in the reference CCSD(T) CBS IP values and call for validation against higher-order EOM-IP-CC methods. Work along these lines is in progress.

# Chapter 5

# Results and Discussions

The NR2-F12 method for calculating IPs and EAs is implemented in massively parallel quantum chemistry (MPQC) electronic structure package version 4. The integrals are evaluated in MPQC with Libint[179] and the distributed-memory implementation is facilitated by the Tiled Array framework.[180] IPs of O21 and 24 medium sized electron acceptor organic molecules (OAM24) were calculated using Green's function and coupled cluster methods and the augmented $\zeta$ correlation consistent basis sets to facilitate stable extrapolation to the basis set limit. The results were presented in Chapter 4.

The basis set incompleteness errors (BSIE) in eV of NR2 and NR2-F12 methods for O21 are shown in Figure 5.1. The rapid convergence of BSIE for NR2-F12 for OAM24 dataset can be seen in Figure 5.2. The labels in plot refer to the maximum absolute error in eV. It can be seen that the $(L_{\max}+1)^{-3}$ convergence of basis set errors of NR2 method is dramatically improved to $(L_{\max}+1)^{-7}$ for NR2-F12. The error for NR2-F12 aXZ is comparable to NR2 a(X+2)Z calculation and the largest errors in NR2-F12 IP are often seen in molecules with halogen group atoms.

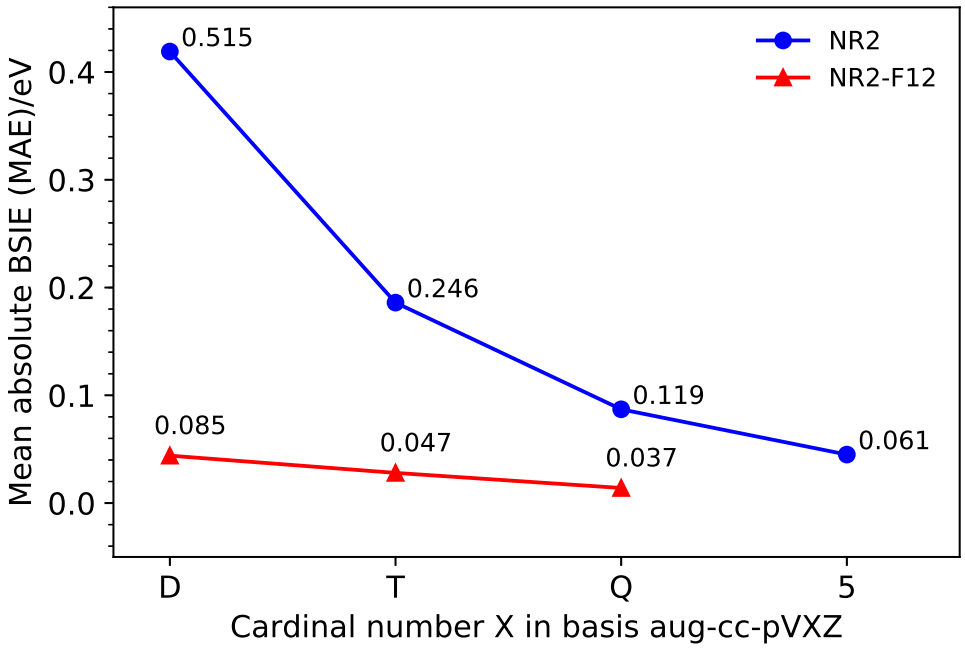



Figure 5.1: Basis set incompleteness error (MAE) of vertical IP in eV for NR2 and NR2-F12 methods for the set of 21 small molecules. Value labels are maximum absolute error (MaxAE) in eV.

For O21, the errors relative to CCSD(T)/CBS for GF2-F12 and NR2-F12 are shown in Figure 5.3. The mean error for GF2-F12/aTZ is 0.483 eV, while it is 0.054 eV for NR2-F12/aTZ. Green's function methods have the tendency to overestimate the ionization potentials. The average errors of O21 IP for CCSD(T)-F12/aTZ and EOM-IP-CCSD/CBS are 0.062 eV and -0.038 eV, respectively. On an average, CCSD(T)-F12 generally overestimates, while EOM-IP-CCSD underestimates IPs. It is evident that CCSD(T)-F12 and EOM-IP-CCSD methods give numerically better IPs compared to GF2-F12 and NR2-F12 with EOM-IP-CCSD being the most accurate among all.

The errors of explicitly-correlated Green's function methods and EOM-IP-CCSD relative to CCSD(T)/CBS reference for OAM24 are shown in Figure 5.4. The average errors for GF2-F12, NR2-F12 and EOM-IP-CCSD are 0.202 eV, -0.130 eV and -0.056 eV, respectively. For OAM24, GF2-F12 always overestimates while NR2-F12 and EOM-IP-CCSD mostly

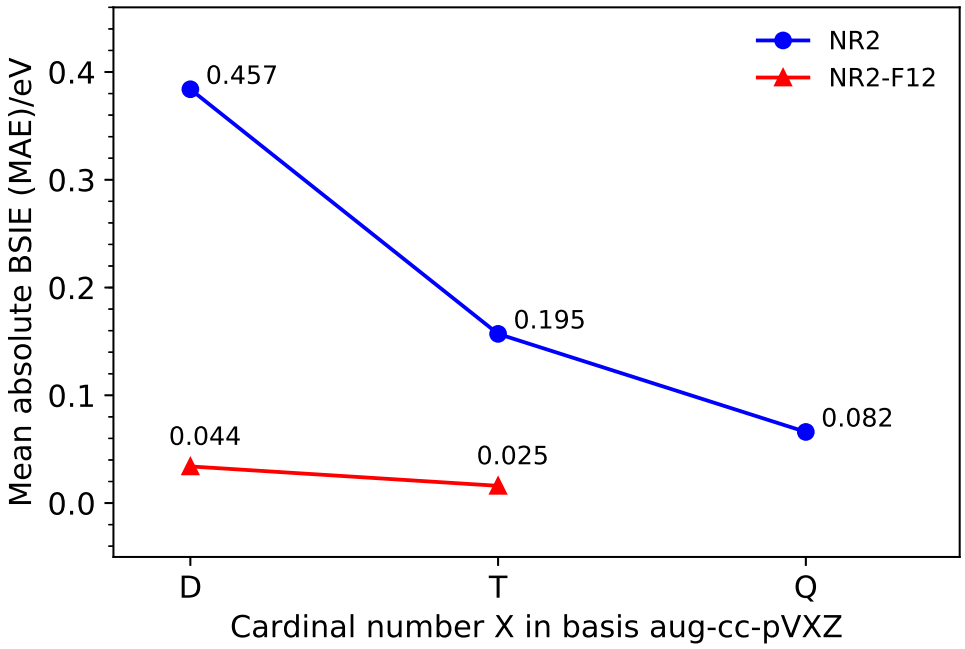
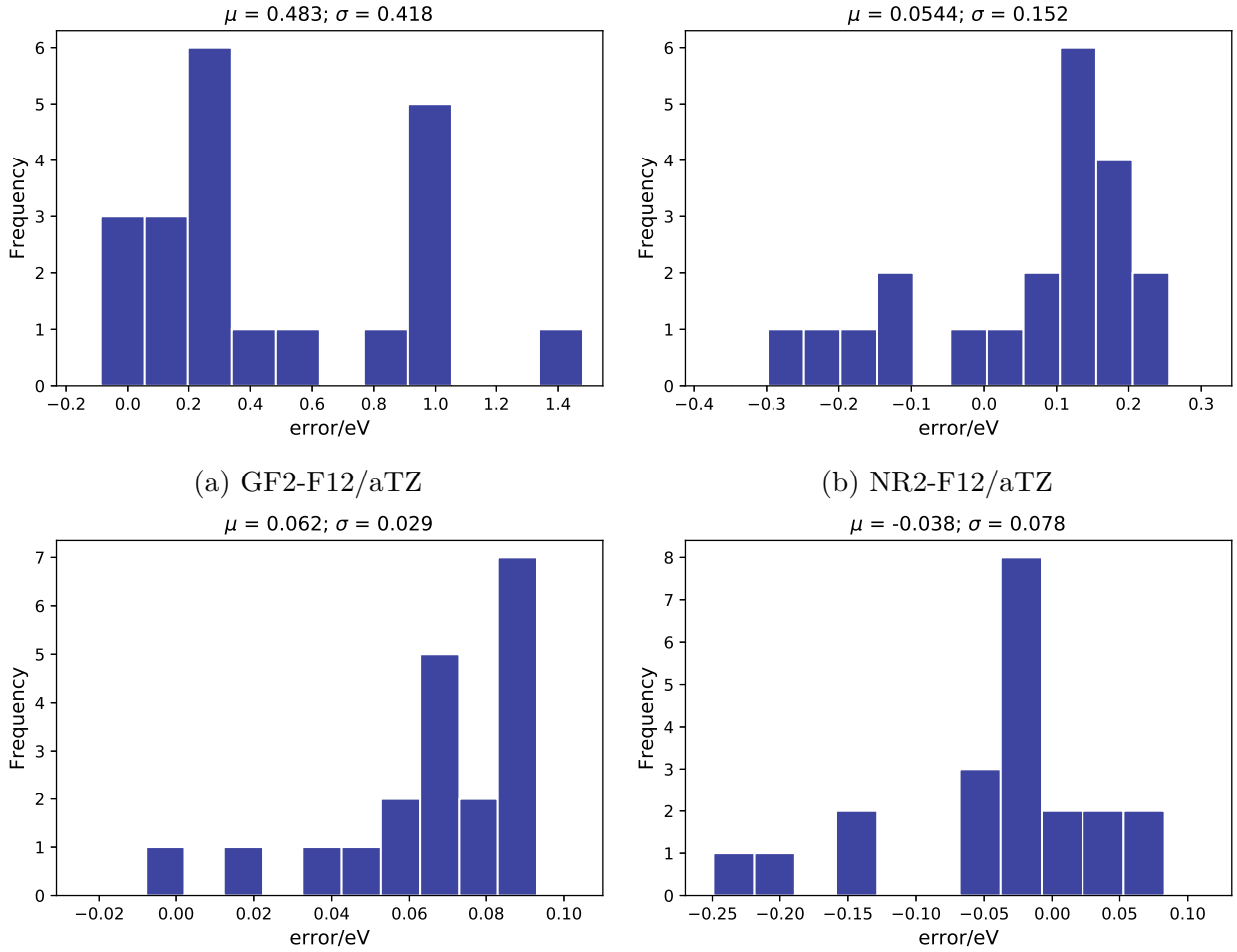



Figure 5.2: Basis set incompleteness error (MAE) of vertical IP in eV for NR2 and NR2-F12 methods for OAM24. The value labels are maximum absolute error (MaxAE) in eV.

underestimate the IP.

NR2-EA requires intermediates with $V^4$ index, making the calculations memory intensive for large bases. Hence, vertical EA were calculated for a subset of the OAM24. The basis set incompleteness errors for finite bases cannot be calculated due to unavailability of NR2 CBS calculated by extrapolating IPs calculated using large bases.

Vertical EAs were calculated for the subset using Green's function and EOM-EA-CCSD methods. The results are summarized in Table 5.1. GF2-F12 is the least expensive among these methods but produces average and maximum errors of 0.56 eV and 0.84 eV, respectively. The NR2/CBS is calculated by extrapolating EA obtained from aug-cc-pVDZ and aug-cc-pVTZ basis. The mean and max errors are about 0.20 eV and 0.30 eV. NR2-F12 with aug-cc-pVDZ basis produces mean and max errors of 0.26 eV and 0.37 eV. These errors reduce and approach the NR2 CBS limit for large bases. EOM-EA-CCSD is the most accurate




(a) GF2-F12/aTZ

(b) NR2-F12/aTZ

(c) CCSD(T)-F12/aTZ

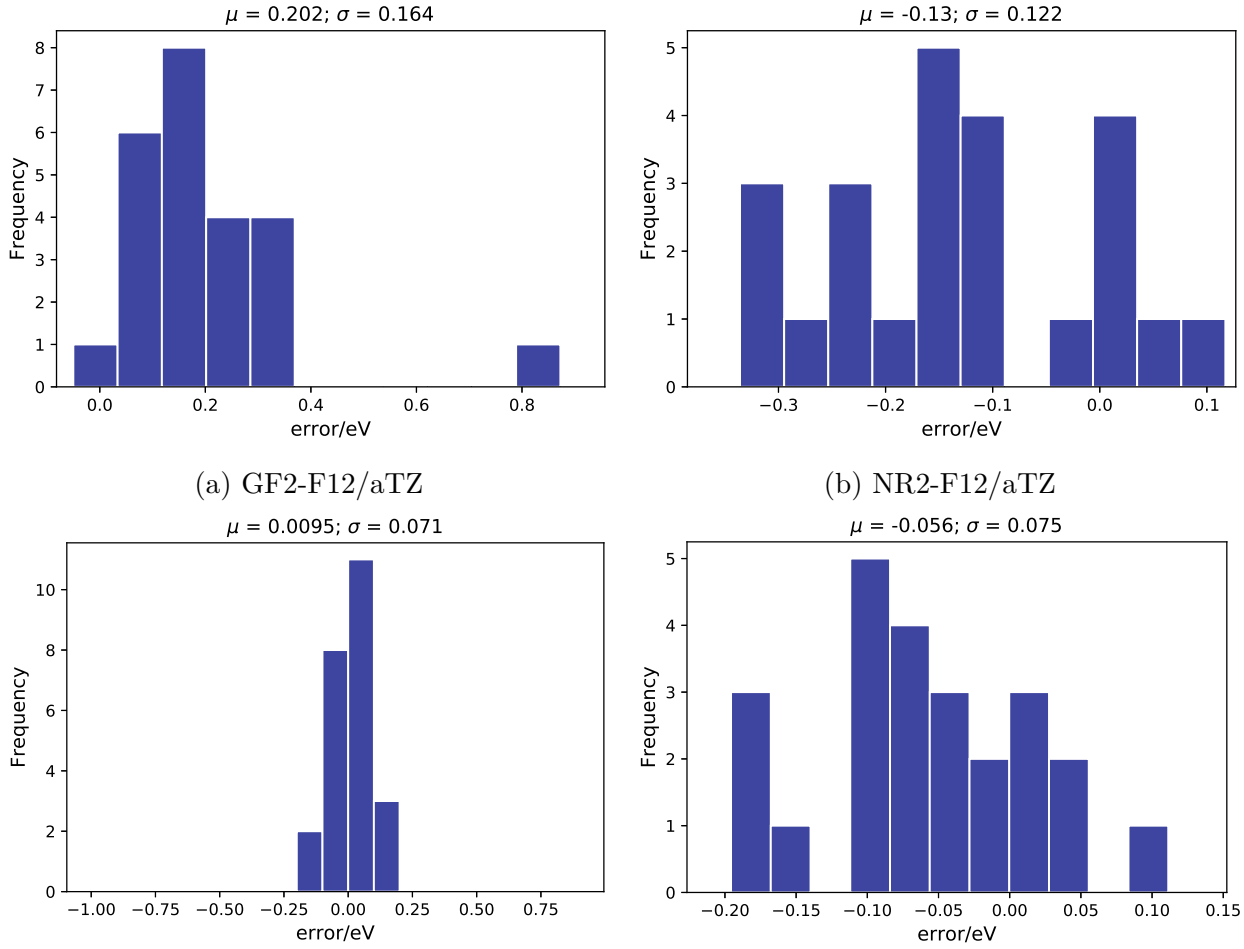
(d) EOM-IP-CCSD/CBS

Figure 5.3: Errors in vertical IP of O21 calculated using F12 methods with aTZ basis.

among these methods with least mean and max errors of 0.06 eV and 0.15 eV.

# Summary

Green's function methods provide an efficient alternative of calculating electron binding energies by introducing the energy dependent potential called self-energy. The self-energy formulation makes it possible to augment the F12 correction perturbatively. For accurate results,




(a) GF2-F12/aTZ

(b) NR2-F12/aTZ

(c) EOM-IP-CCSD/aTZ

(d) EOM-IP-CCSD/CBS

Figure 5.4: Errors in vertical IP of OAM24 calculated using Green's function F12 and EOM-IP-CCSD.

the Dyson equation has to be solved iteratively using a large basis set. However, Green's function methods suffer from the slow asymptotic convergence of basis set incompleteness error (BSIE), limiting their applications to small molecular systems. The explicitly-correlated self-energy term reduces the BSIE, thereby allowing to calculate accurate self-energy using a smaller basis. Inclusion of explicitly-correlated self-energy shows dramatic convergence to basis set limit for both second-order (GF2-F12) and renormalized second-order (NR2-F12) Green's function methods.

Table 5.1: Differences of vertical EA in eV of a subset of OAM24 calculated with GF and coupled-cluster methods relative to CCSD(T) CBS values from Ref. 1.

| | GF2-F12 | | NR2-F12 | NR2 | EOM-EA-CCSD | CCSD(T) | Exp. |
|---|---|---|---|---|---|---|---|
| | aDZ | aTZ | aDZ | CBS[1] | CBS[1] | CBS | |
| Benzoquinone | 0.435 | 0.464 | 0.150 | 0.076 | -0.029 | 1.550 | 1.85* |
| Benzonitrile | 0.472 | 0.507 | 0.184 | 0.117 | -0.065 | -0.210 | 0.26 |
| Fumaronitrile | 0.556 | 0.597 | 0.323 | 0.278 | 0.050 | 0.980 | 1.25 |
| TCNE | 0.844 | 0.883 | 0.369 | 0.307 | 0.152 | 3.050 | 3.16* |
| Maleic anhydride | 0.492 | 0.524 | 0.260 | 0.189 | 0.011 | 1.010 | 1.44 |
| MAE[2] | 0.560 | 0.595 | 0.257 | 0.193 | 0.061 | 0.32[3] | |
| MaxAE[2] | 0.844 | 0.883 | 0.369 | 0.307 | 0.152 | 0.47[3] | |

*Adiabatic EA.
[1]CBS value calculated by the $X^{-3}$ extrapolation of aDZ and aTZ EAs.
[2]Mean and max errors relative to the CCSD(T) CBS values.
[3]Errors relative to the experimental values.

The explicitly correlated GF2-F12 and NR2-F12 methods are compared to coupled cluster methods. The coupled cluster methods use a CCSD reference and therefore scale as $\mathcal{O}(N^6)$ with system size. The errors in electron binding energy calculated with coupled cluster methods at a complete basis are lower than those calculated using Green's function F12 methods and finite basis. However, EOM-IP/EA-CCSD also suffers from a slow convergence of BSIE and reducing it is crucial for obtaining accurate electron binding energies with smaller basis sets.

# Chapter 6

# Future work

The versatility of Green's function allow this method to be used in a variety of problems. The GF2-F12 and NR2-F12 methods have been tested for small and medium sized organic molecules. Traditionally, the regular and geminal field operators are considered to be decoupled. The effect of coupling between regular and geminal field operators is being formulated.

With the canonical basis, Green's function methods cannot be used to study large molecular systems. So, reduced scaling Green's function methods formulated in a local basis can be used to study extended molecular systems.

On the other hand, the EOM-CC methods have proven to be more accurate than the propagator methods. In this direction, a perturbative F12 correction to EOM-IP-CCSD can be formulated. Also, accurate benchmarks of small molecules can be calculated using massively parallel implementation of EOM-IP/EA-CCSDT.

## 6.1 Reduced scaling Green's function method

The local correlation approach of Pulay and Sæbø utilized pair atomic orbitals (PAOs).[181,182] However, for higher accuracy, the size of PAO domains and the number of basis functions per domain increase drastically and the compact representation of wave function is lost. Pair natural orbitals (PNOs), on the other hand, provide high accuracy with a compact representation of wave function amplitudes.[83,183] PNOs are widely used for local correlation approaches. Neese *et al.* formulated the local PNO-CEPA,[184] cubic-scaling PNO-MP2 has been implemented by Hättig and co-workers and PNO–local MP2 by Werner and co-workers.[185,186] The iteratively optimized PNOs (iPNO) have been shown to reduce the errors arising due to truncation of regular PNOs.[187]

It has been shown that the domain based local pair natural orbital (DLPNO) based local MP2 method described using sparse maps infrastructure can reproduce the canonical MP2 energy with minimal number of cutoff-parameters.[25] To lower basis set incompleteness errors, the DLPNO-MP2-F12 method was formulated.[27] Based on these ideas, a DLPNO-GF2-F12 method for calculating electron binding energy is in the works.

## 6.2 EOM-IP-CCSD-F12 formalism

With tests on medium sized organic molecules and small anions, it is apparent that EOM-IP-CCSD outperforms both GF2 and NR2 methods in terms of accuracy. However, it suffers from the slow convergence of basis set incompleteness errors. In this direction, Bokhan and Ten-no formulated the IP-EOM-CCSD(F12) method.[188] They used a CCSD(F12) reference state which uses the geminal $\hat{T}_2'$ operators. The cost of EOM-IP on a CCSD(F12) reference is the same as EOM-IP-CCSD for this method. However, this iterative method requires

calculating $\hat{T}_2'$ for every iteration while calculating CCSD(F12) making it computationally expensive.

To overcome this problem, it is desirable to have a formalism where the F12 correction is included perturbatively. In this direction, the second-order explicitly correlated correction to CCSD energy ($\text{CCSD(2)}_{\overline{\text{R12}}}$) was done by Valeev.[189] Based on a $\text{CCSD(2)}_{\overline{\text{R12}}}$ reference state, an IP-EOM formulation is under study.

## 6.3 Parallel implementation of EOM-IP-CCSDT

The exponential ansatz of the cluster operator includes higher order terms in the wave function making it a promising approach for direct calculation of electron detachment energies. The coupled-cluster energy equation is given by:

$$e^{-\hat{T}}\bar{H}e^{\hat{T}}\left|\Psi\right\rangle = E\left|\Psi\right\rangle. \tag{6.3.1}$$

For proper description of excited electronic states, the double excitations (*pair excitations*) have to be included in the wave function. This makes it imperative to include the triples for calculating direct IP using EOM methods accurately.[62] For calculating the triples, the operator $\hat{T}$ is truncated at triple excitations:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3. \tag{6.3.2}$$

Using coupled cluster for calculating electron excitation energies, ionization potentials and election affinities is discussed in section (1.3.4). CCSDT scales as $\mathcal{O}(N^8)$ with system size and large intermediates make it memory intensive thereby limiting applications to small molecules. These issues can be solved by state of art tensor decomposition techniques[190,191]

to reduce storage and parallel implementations for applications to moderately sized systems. Evaluation of accurate benchmarks using EOM-IP-CCSDT method is in progress.

Recently, reduced scaling EOM-IP-CCSD method has been formulated by Dutta *et al.* using DLPNO method.[192] Similarly, DLPNO method can be used in formulation of a local EOM-IP-CCSDT for applications to large molecular systems.

# Chapter 7

# Conclusion

In this work, the explicitly correlated non-diagonal renormalized second-order Green's function (NR2-F12) method is formulated and implemented for calculation of ionization potentials and electron affinities of closed shell molecules. The iterative $\mathcal{O}(N^5)$ scaling for IP and $\mathcal{O}(N^6)$ scaling for EA makes NR2-F12 a good candidate for studying electron binding energies of small and medium sized molecules. The F12 correction leads to a $(L+1)^{-7}$ convergence of BSIE compared to the traditional $(L+1)^{-3}$ rate for conventional methods.

The performance of GF2-F12 and NR2-F12 methods was compared for O21 and OAM24 datasets. Renormalization includes higher order terms in the NR2 formalism. Hence, at basis set limit, the error in IP for NR2 is less compared to GF2 relative to a CCSD(T)/CBS reference. The performance of Green's function F12 method is compared to EOM-CCSD method for IP. Owing to the lower BSIE, the errors in IP of NR2-F12 are similar to those obtained from EOM-IP-CCSD with a double $\zeta$ basis. For a large basis, EOM-IP-CCSD produces the smallest errors for a CCSD(T)/CBS reference.

NR2-F12 shows lower BSIE than conventional NR2 for EA. However, it is not as dramatic

as for those observed in NR2-F12 IP.

# Publication list

1. Teke, N. K.; Pavošević, F.; Peng, C.; Valeev, E. F. *J. Chem. Phys.* **2019**, *150* (21), 214103.

# Bibliography


[1] Richard, R. M.; Marshall, M. S.; Dolgounitcheva, O.; Ortiz, J. V.; Brédas, J.-L.; Marom, N.; Sherrill, C. D. *J. Chem. Theory Comput.* **2016**, *12*, 595–604.

[2] Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C.; IUCr, *Acta Crystallogr. Sect. D Biol. Crystallogr.* **2002**, *58*, 1948–1954.

[3] Sheldrick, G. M. *Acta Crystallogr. Sect. A Found. Adv.* **2015**, *71*, 3–8.

[4] Tang, Y.; Chen, M.; Zhou, L.; Ma, J.; Li, Y.; Zhang, H.; Shi, Z.; Xu, Q.; Zhang, X.; Gao, Z.; Zhao, Y.; Cheng, Y.; Jiao, S.; Zhou, Z. *Cell Discov.* **2019**, *5*, 3.

[5] Hajduk, P. J.; Greer, J. *Nat. Rev. Drug Discov.* **2007**, *6*, 211–219.

[6] Ferreira, L.; dos Santos, R.; Oliva, G.; Andricopulo, A.; Ferreira, L. G.; Dos Santos, R. N.; Oliva, G.; Andricopulo, A. D. *Molecules* **2015**, *20*, 13384–13421.

[7] Rodrigues, T.; Reker, D.; Schneider, P.; Schneider, G. *Nat. Chem.* **2016**, *8*, 531–541.

[8] Huynh, W. U. *Science* **2002**, *295*, 2425–2427.

[9] Kamat, P. V. *J. Phys. Chem. C* **2008**, *112*, 18737–18753.

[10] Jiang, J.; Li, Y.; Liu, J.; Huang, X.; Yuan, C.; Lou, X. W. D. *Adv. Mater.* **2012**, *24*, 5166–5180.

[11] Heo, J. H.; Im, S. H.; Noh, J. H.; Mandal, T. N.; Lim, C.-S.; Chang, J. A.; Lee, Y. H.; Kim, H.-j.; Sarkar, A.; Nazeeruddin, M. K.; Grätzel, M.; Seok, S. I. *Nat. Photonics* **2013**, *7*, 486–491.

[12] Rupp, M.; Tkatchenko, A.; Müller, K.-R.; von Lilienfeld, O. A. *Phys. Rev. Lett.* **2012**, *108*, 058301.

[13] Botu, V.; Ramprasad, R. *Int. J. Quantum Chem.* **2015**, *115*, 1074–1083.

[14] Li, Z.; Kermode, J. R.; De Vita, A. *Phys. Rev. Lett.* **2015**, *114*, 096405.

[15] Hansen, K.; Biegler, F.; Ramakrishnan, R.; Pronobis, W.; von Lilienfeld, O. A.; Müller, K.-R.; Tkatchenko, A. *J. Phys. Chem. Lett.* **2015**, *6*, 2326–2331.

[16] Coe, J. P. *J. Chem. Theory Comput.* **2018**, *14*, 5739–5749.

[17] Cory, D. et al. *Fortschritte der Phys.* **2000**, *48*, 875–907.

[18] Kane, B. E. *Nature* **1998**, *393*, 133–137.

[19] Lanyon, B. P.; Whitfield, J. D.; Gillett, G. G.; Goggin, M. E.; Almeida, M. P.; Kassal, I.; Biamonte, J. D.; Mohseni, M.; Powell, B. J.; Barbieri, M.; Aspuru-Guzik, A.; White, A. G. *Nat. Chem.* **2010**, *2*, 106–111.

[20] Jones, R. O. *Rev. Mod. Phys.* **2015**, *87*, 897–923.

[21] Černý, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2005**, *7*, 1624–1626.

[22] Dreuw, A.; Head-Gordon, M. *J. Am. Chem. Soc.* **2004**, *126*, 4007–4016.

[23] Cohen, A. J.; Mori-Sanchez, P.; Yang, W. *Science* **2008**, *321*, 792–794.

[24] Gordon, M. S.; Fedorov, D. G.; Pruitt, S. R.; Slipchenko, L. V. *Chem. Rev.* **2012**, *112*, 632–672.

[25] Pinski, P.; Riplinger, C.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2015**, *143*, 034108.

[26] Riplinger, C.; Pinski, P.; Becker, U.; Valeev, E. F.; Neese, F. *J. Chem. Phys.* **2016**, *144*, 024109.

[27] Pavošević, F.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. *J. Chem. Phys.* **2016**, *144*, 144109.

[28] Feller, D.; Dixon, D. A. *J. Chem. Phys.* **2001**, *115*, 3484–3496.

[29] Tajti, A.; Szalay, P. G.; Császár, A. G.; Kállay, M.; Gauss, J.; Valeev, E. F.; Flowers, B. A.; Vázquez, J.; Stanton, J. F. *J. Chem. Phys.* **2004**, *121*, 11599–11613.

[30] Yan Zhao,; Núria González-García,; Truhlar, D. G. **2005**,

[31] Rez, J.; Riley, K. E.; Hobza, P. *J. Chem. Theory Comput* **2011**, *7*, 2427–2438.

[32] Born, M.; Oppenheimer, R. *Ann. Phys.* **1927**, *389*, 457–484.

[33] Worth, G. A.; Cederbaum, L. S. *Annu. Rev. Phys. Chem* **2004**, *55*, 127–58.

[34] Pisana, S.; Lazzeri, M.; Casiraghi, C.; Novoselov, K. S.; Geim, A. K.; Ferrari, A. C.; Mauri, F. *Nat. Mater.* **2007**, *6*, 198–201.

[35] Schrödinger, E. *Ann. Phys.* **1926**, *384*, 361–376.

[36] Hartree, D. R. *Math. Proc. Cambridge Philos. Soc.* **1928**, *24*, 89.

[37] Gerlach, W.; Stern, O. *Zeitschrift für Phys.* **1922**, *9*, 349–352.

[38] Pauli, W. *Zeitschrift für Phys.* **1925**, *31*, 765–783.

[39] Slater, J. C. *Phys. Rev.* **1951**, *81*, 385–390.

[40] Shavitt, I.; Bartlett, R. J. *Many - Body Methods Chem. Phys. MBPT Coupled-Cluster Theory*; 2009.

[41] Löwdin, P.-O. *Phys. Rev.* **1955**, *97*, 1474–1489.

[42] Kutzelnigg, W. *THEORY OF ELECTRON CORRELATION*; 2003.

[43] Paldus, J. *J. Chem. Phys.* **1974**, *61*, 5321–5330.

[44] Head-Gordon, M.; Maurice, D.; Oumi, M. *Chem. Phys. Lett.* **1995**, *246*, 114–121.

[45] Kállay, M.; Gauss, J. *J. Chem. Phys.* **2004**, *121*, 9257–9269.

[46] Dreuw, A.; Head-Gordon, M. *Chem. Rev.* **2005**, *105*, 4009–4037.

[47] Eriksen, J. J.; Gauss, J. *J. Chem. Theory Comput.* **2018**, *14*, 5180–5191.

[48] Slater, J. C. *Phys. Rev.* **1929**, *34*, 1293–1322.

[49] Condon, E. U. *Phys. Rev.* **1930**, *36*, 1121–1133.

[50] Freed, K. F. *Phys. Rev.* **1968**, *173*, 1–24.

[51] Freed, K. F. *Annu. Rev. Phys. Chem.* **1971**, *22*, 313–346.

[52] Kutzelnigg, W. *Int. J. Quantum Chem.* **2009**, *109*, 3858–3884.

[53] Møller, C.; Plesset, M. S. *Phys. Rev.* **1934**, *46*, 618–622.

[54] Coester, F. *Nucl. Phys.* **1958**, *7*, 421.

[55] Coester, F. *Nucl. Phys.* **1960**, *17*, 477.

[56] Čížek, J. *Adv. Chem. Phys.*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2007; pp 35–89.

[57] Čížek, J. *J. Chem. Phys.* **1966**, *45*, 4256–4266.

[58] Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[59] Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

[60] Nooijen, M.; Bartlett, R. J. *J. Chem. Phys.* **1997**, *106*, 6441–6448.

[61] Stanton, J. F.; Bartlett, R. J. *J. Chem. Phys.* **1993**, *98*, 7029–7039.

[62] Musiał, M. *Mol. Phys.* **2010**, *108*, 2921–2931.

[63] Nooijen, M.; Snijders, J. G. *Int. J. Quantum Chem.* **1993**, *48*, 15–48.

[64] Mahapatra, U. S.; Datta, B.; Mukherjee, D. *J. Chem. Phys.* **1999**, *110*, 6171–6188.

[65] Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2000**, *113*, 18–35.

[66] Kowalski, K.; Piecuch, P. *J. Chem. Phys.* **2004**, *120*, 1715–1738.

[67] Adler, T. B.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2007**, *127*, 221106.

[68] Valeev, E. F.; Daniel Crawford, T. *J. Chem. Phys.* **2008**, *128*, 244113.

[69] Noga, J.; Kutzelnigg, W.; Klopper, W. *Chem. Phys. Lett.* **1992**, *199*, 497–504.

[70] Pavošević, F.; Neese, F.; Valeev, E. F. *J. Chem. Phys.* **2014**, *141*, 054106.

[71] Bartlett, R. J. *Ann. Rev. Phys. Chem* **1981**, *32*, 359–401.

[72] Crawford, T. D.; Schaefer, H. F. *Rev. Comput. Chem.*; John Wiley & Sons, Inc., 2000; Chapter 2, pp 33–136.

[73] Bartlett, R. J.; Musiał, M. *Rev. Mod. Phys.* **2007**, *79*, 291–352.

[74] deB Darwent, B. *Bond dissociation energies in simple molecules*; 1970.

[75] Bartlett, R. J.; Purvis, G. D. *Int. J. Quantum Chem.* **1978**, *14*, 561–581.

[76] Purvis, G. D.; Bartlett, R. J. *J. Chem. Phys.* **1982**, *76*, 1910–1918.

[77] Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. *Chem. Phys. Lett.* **1989**, *157*, 479–483.

[78] Řezáč, J.; Hobza, P. *J. Chem. Theory Comput.* **2013**, *9*, 2151–2155.

[79] Noga, J.; Bartlett, R. J. *J. Chem. Phys.* **1987**, *86*, 7041–7050.

[80] Schwegler, E.; Challacombe, M. *J. Chem. Phys.* **1996**, *105*, 2726–2734.

[81] Dziedzic, J.; Hill, Q.; Skylaris, C.-K. *J. Chem. Phys.* **2013**, *139*, 214103.

[82] Schütz, M.; Hetzer, G.; Werner, H.-J. *J. Chem. Phys.* **1999**, *111*, 5691–5705.

[83] Riplinger, C.; Neese, F. *J. Chem. Phys.* **2013**, *138*, 034106.

[84] Pavošević, F.; Peng, C.; Pinski, P.; Riplinger, C.; Neese, F.; Valeev, E. F. *J. Chem. Phys.* **2017**, *146*, 174108.

[85] Pickup, B. T.; Goscinski, O. *Mol. Phys.* **1973**, *26*, 1013–1035.

[86] von Niessen, W.; Schirmer, J.; Cederbaum, L. *Comput. Phys. Reports* **1984**, *1*, 57–125.

[87] Danovich, D. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 377–387.

[88] Ortiz, J. V. *J. Chem. Phys.* **2000**, *112*, 56–68.

[89] Zubarev, D. N. *Sov. Phys. Uspekhi* **1960**, *3*, 320.

[90] Dahlen, N. E.; van Leeuwen, R. *J. Chem. Phys.* **2005**, *122*, 164102.

[91] Schmitteckert, P. *J. Phys. Conf. Ser.* **2010**, *220*, 012022.

[92] Linderberg, J.; Ohrn, Y. *Proc. R. Soc. London. Ser. A. Math. Phys. Sci.* **1965**, *285*, 445–456.

[93] Deleuze, M. S.; Trofimov, A. B.; Cederbaum, L. S. *J. Chem. Phys.* **2001**, *115*, 5859–5882.

[94] Deleuze, M. S. *J. Phys. Chem. A* **2004**, *108*, 9244–9259.

[95] Lan, T. N.; Kananenka, A. A.; Zgid, D. *J. Chem. Phys.* **2015**, *143*, 241102.

[96] Rusakov, A. A.; Zgid, D. *J. Chem. Phys.* **2016**, *144*, 054106.

[97] Kananenka, A. A.; Zgid, D. *J. Chem. Theory Comput.* **2017**, *13*, 5317–5331.

[98] Ortiz, J. V. *Int. J. Quantum Chem.* **1992**, *44*, 1–11.

[99] Ortiz, J. V. *J. Chem. Phys.* **1995**, *103*, 5630–5639.

[100] Ortiz, J. *Adv. Quantum Chem.*; 1999; Vol. 35; pp 33–52.

[101] Koopmans, T. *Physica* **1934**, *1*, 104–113.

[102] Linderberg, J.; Öhrn, Y. *Propagators in Quantum Chemistry*; Theoretical chemistry; Wiley, 2004.

[103] Corzo, H. H.; Ortiz, J. V. *Adv. Quantum Chem.*; Elsevier Inc, 2017; Vol. 74; pp 267–298.

[104] Ortiz, J. V. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2013**, *3*, 123–142.

[105] Ortiz, J. V. *J. Chem. Phys.* **1998**, *108*, 1008–1014.

[106] Dolgounitcheva, O.; Díaz-Tinoco, M.; Zakrzewski, V. G.; Richard, R. M.; Marom, N.; Sherrill, C. D.; Ortiz, J. V. *J. Chem. Theory Comput.* **2016**, *12*, 627–637.

[107] Bhaskaran-Nair, K.; Kowalski, K.; Shelton, W. A. *J. Chem. Phys.* **2016**, *144*, 144101.

[108] Peng, B.; Kowalski, K. *J. Chem. Theory Comput.* **2018**, *14*, 4335–4352.

[109] Díaz-Tinoco, M.; Corzo, H. H.; Ortiz, J. V. *J. Chem. Theory Comput.* **2018**, *14*, 5881–5895.

[110] Born, G.; Kurtz, H. A.; Öhrn, Y. *J. Chem. Phys.* **1978**, *68*, 74.

[111] Corzo, H. H.; Galano, A.; Dolgounitcheva, O.; Zakrzewski, V. G.; Ortiz, J. V. *J. Phys. Chem. A* **2015**, *119*, 8813–8821.

[112] Kutzelnigg, W. *Theor. Chim. Acta* **1985**, *68*, 445–469.

[113] Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. *Int. Rev. Phys. Chem.* **2006**, *25*, 427–468.

[114] Ten-no, S.; Noga, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 114–125.

[115] Kato, T. *Commun. Pure Appl. Math.* **1957**, *10*, 151–177.

[116] Hylleraas, E. A. *Zeitschrift für Phys.* **1929**, *54*, 347–366.

[117] Kinoshita, T. *Phys. Rev.* **1959**, *115*, 366–374.

[118] Klopper, W.; Kutzelnigg, W. *J. Phys. Chem.* **1990**, *94*, 5625–5630.

[119] Klopper, W.; Kutzelnigg, W. *J. Chem. Phys.* **1991**, *94*, 2020–2030.

[120] Klopper, W.; Samson, C. C. M. *J. Chem. Phys.* **2002**, *116*, 6397–6410.

[121] Valeev, E. F. *Chem. Phys. Lett.* **2004**, *395*, 190–195.

[122] Ten-no, S. *J. Chem. Phys.* **2004**, *121*, 117.

[123] Ten-no, S. *J. Chem. Phys.* **2007**, *126*, 014108.

[124] Ajili, Y.; Hammami, K.; Jaidane, N. E.; Lanza, M.; Kalugina, Y. N.; Lique, F.; Hochlaf, M. *Phys. Chem. Chem. Phys.* **2013**, *15*, 10062.

[125] Győrffy, W.; Knizia, G.; Werner, H.-J. *J. Chem. Phys.* **2017**, *147*, 214101.

[126] Noga, J.; Kedžuch, S.; Šimunek, J.; Ten-no, S. *J. Chem. Phys.* **2008**, *128*, 174103.

[127] Kedžuch, S.; Demel, O.; Pittner, J.; Ten-no, S.; Noga, J. *Chem. Phys. Lett.* **2011**, *511*, 418–423.

[128] Demel, O.; Kedžuch, S.; Švaňa, M.; Ten-no, S.; Pittner, J.; Noga, J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4753.

[129] Johnson, C. M.; Doran, A. E.; Ten-no, S. L.; Hirata, S. *J. Chem. Phys.* **2018**, *149*, 174112.

[130] Kats, D.; Tew, D. P. *J. Chem. Theory Comput.* **2019**, *15*, 13–17.

[131] Pavošević, F.; Peng, C.; Ortiz, J. V.; Valeev, E. F. *J. Chem. Phys.* **2017**, *147*, 121101.

[132] Thouless, D. J. *The quantum mechanics of many-body systems*; Academic Press: New York, 1961.

[133] Economou, E. N. *Green's functions in quantum physics*; Springer, 1983; Vol. 3.

[134] Mattuck, R. D. *A guide to Feynman diagrams in the many-body problem*; McGraw-Hill, 1967.

[135] Öhrn, Y.; Linderberg, J. *Phys. Rev.* **1965**, *139*, A1063–A1068.

[136] Phillips, J. J.; Zgid, D. *J. Chem. Phys.* **2014**, *140*, 241101.

[137] Hirata, S.; Hermes, M. R.; Simons, J.; Ortiz, J. V. *J. Chem. Theory Comput.* **2015**, *11*, 1595–1606.

[138] Hirata, S.; Doran, A. E.; Knowles, P. J.; Ortiz, J. V. *J. Chem. Phys.* **2017**, *147*, 044108.

[139] Movassagh, R.; Strang, G.; Tsuji, Y.; Hoffmann, R. *J. Math. Phys.* **2017**, *58*, 033505.

[140] Janesko, B. G.; Henderson, T. M.; Scuseria, G. E. *J. Chem. Phys.* **2009**, *130*, 081105.

[141] Ma, Y.; Rohlfing, M.; Molteni, C. *J. Chem. Theory Comput.* **2010**, *6*, 257–265.

[142] Knight, J. W.; Wang, X.; Gallandi, L.; Dolgounitcheva, O.; Ren, X.; Ortiz, J. V.; Rinke, P.; Körzdörfer, T.; Marom, N. *J. Chem. Theory Comput.* **2016**, *12*, 615–626.

[143] Heßelmann, A. *Phys. Rev. A* **2017**, *95*, 062513.

[144] Maier, T.; Jarrell, M.; Pruschke, T.; Hettler, M. H. *Rev. Mod. Phys.* **2005**, *77*, 1027–1080.

[145] Kotliar, G.; Savrasov, S. Y.; Haule, K.; Oudovenko, V. S.; Parcollet, O.; Marianetti, C. A. *Rev. Mod. Phys.* **2006**, *78*, 865–951.

[146] Lan, T. N.; Zgid, D. *J. Phys. Chem. Lett.* **2017**, *8*, 2200–2205.

[147] Cederbaum, L. S. *Theor. Chim. Acta* **1973**, *31*, 239–260.

[148] Cederbaum, L. S. *J. Phys. B At. Mol. Phys.* **1975**, *8*, 290–303.

[149] Cederbaum, L. S.; Domcke, W. *Adv. Chem. Phys.* **1977**, *36*, 205–344.

[150] Cederbaum, L. S.; Domcke, W.; Schirmer, J.; von Niessen, W. *Phys. Scr.* **1980**, *21*, 481–491.

[151] von Niessen, W. *J. Chem. Phys.* **1999**, *67*, 4124.

[152] Kutzelnigg, W. *Theor. Chim. Acta* **1985**, *68*, 445–469.

[153] Gill, P. M.; Werner, H. J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 3318.

[154] Shiozaki, T.; Valeev, E. F.; Hirata, S. *Annu. Rep. Comput. Chem.*; 2009; pp 131–148.

[155] Tew, D. P.; Klopper, W. *Mol. Phys.* **2010**, *108*, 315–325.

[156] Hättig, C.; Tew, D. P.; Köhn, A. *J. Chem. Phys.* **2010**, *132*, 231102.

[157] Kong, L.; Bischoff, F. A.; Valeev, E. F. *Chem. Rev.* **2012**, *112*, 75–107.

[158] Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. *Chem. Rev.* **2012**, *112*, 4–74.

[159] Ten-no, S.; Noga, J. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 114–125.

[160] Ten-no, S. *Theor. Chem. Acc.* **2012**, *131*, 1070.

[161] Ohnishi, Y.-y.; Ten-no, S. *J. Comput. Chem.* **2016**, *37*, 2447–2453.

[162] Goscinski, O.; Lukman, B. *Chem. Phys. Lett.* **1970**, *7*, 573–576.

[163] Kutzelnigg, W.; Mukherjee, D. *J. Chem. Phys.* **1997**, *107*, 432–449.

[164] Kutzelnigg, W.; Morgan, J. D. *J. Chem. Phys.* **1992**, *96*, 4484–4508.

[165] May, A. J.; Valeev, E.; Polly, R.; Manby, F. R. *Phys. Chem. Chem. Phys.* **2005**, *7*, 2710.

[166] Kedžuch, S.; Milko, M.; Noga, J. *Int. J. Quantum Chem.* **2005**, *105*, 929–936.

[167] Peng, C.; Calvin, J. A.; Pavošević, F.; Zhang, J.; Valeev, E. F. *J. Phys. Chem. A* **2016**, *120*, 10231–10244.

[168] Dunning, T. H. *J. Chem. Phys.* **1989**, *90*, 1007–1023.

[169] Kendall, R. A.; Dunning, T. H.; Harrison, R. J. *J. Chem. Phys.* **1992**, *96*, 6796–6806.

[170] Woon, D. E.; Dunning, T. H. *J. Chem. Phys.* **1993**, *98*, 1358–1371.

[171] Weigend, F.; Köhn, A.; Hättig, C. *J. Chem. Phys.* **2002**, *116*, 3175–3183.

[172] Yousaf, K. E.; Peterson, K. A. *Chem. Phys. Lett.* **2009**, *476*, 303–307.

[173] Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. *Chem. Phys. Lett.* **1998**, *286*, 243–252.

[174] Valeev, E. F. *Phys Chem Chem Phys* **2008**, *10*, 106–113.

[175] Zhang, J.; Valeev, E. F. *J. Chem. Theory Comput.* **2012**, *8*, 3175–3186.

[176] Neese, F. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2018**, *8*, e1327.

[177] Peterson, K. A.; Adler, T. B.; Werner, H.-J. *J. Chem. Phys.* **2008**, *128*, 084102.

[178] Neese, F.; Valeev, E. F. *J. Chem. Theory Comput.* **2011**, *7*, 33–43.

[179] Valeev, E. F. Libint: A library for the evaluation of molecular integrals of many-body operators over Gaussian functions. 2017.

[180] Calvin, J. A.; Lewis, C. A.; Valeev, E. F. *Proc. 5th Work. Irregul. Appl. Archit. Algorithms - IA3 '15* **2015**, 1–8.

[181] Pulay, P. *Chem. Phys. Lett.* **1983**, *100*, 151–154.

[182] Sæbø, S.; Pulay, P. *Chem. Phys. Lett.* **1985**, *113*, 13–18.

[183] Sparta, M.; Neese, F. *Chem. Soc. Rev.* **2014**, *43*, 5032–5041.

[184] Neese, F.; Wennmohs, F.; Hansen, A. *J. Chem. Phys.* **2009**, *130*, 114108.

[185] Krause, C.; Werner, H.-J. *Phys. Chem. Chem. Phys.* **2012**, *14*, 7591.

[186] Werner, H.-J.; Knizia, G.; Krause, C.; Schwilk, M.; Dornbach, M. *J. Chem. Theory Comput.* **2015**, *11*, 484–507.

[187] Clement, M. C.; Zhang, J.; Lewis, C. A.; Yang, C.; Valeev, E. F. *J. Chem. Theory Comput.* **2018**, *14*, 4581–4589.

[188] Bokhan, D.; Ten-no, S. *J. Chem. Phys.* **2010**, *132*, 021101.

[189] Valeev, E. F. *Phys. Chem. Chem. Phys.* **2008**, *10*, 106–113.

[190] Kolda, T. G.; Bader, B. W. *SIAM Rev.* **2009**, *51*, 455–500.

[191] Schutski, R.; Zhao, J.; Henderson, T. M.; Scuseria, G. E. *J. Chem. Phys.* **2017**, *147*, 184113.

[192] Dutta, A. K.; Saitow, M.; Riplinger, C.; Neese, F.; Izsák, R. *J. Chem. Phys.* **2018**, *148*, 244101.